When Does Power Disparity Help or Hurt Group Performance?

Murat Tarakci
Erasmus University

Patrick J. F. Groenen
Erasmus University

Lindred L. Greer
Stanford University

Power differences are ubiquitous in social settings. However, the question of whether groups with higher or lower power disparity achieve better performance has thus far received conflicting answers. To address this issue, we identify 3 underlying assumptions in the literature that may have led to these divergent findings, including a myopic focus on static hierarchies, an assumption that those at the top of hierarchies are competent at group tasks, and an assumption that equality is not possible. We employ a multimethod set of studies to examine these assumptions and to understand when power disparity will help or harm group performance. First, our agent-based simulation analyses show that by unpacking these common implicit assumptions in power research, we can explain earlier disparate findings—power disparity benefits group performance when it is dynamically aligned with the power holder’s task competence, and harms group performance when held constant and/or is not aligned with task competence. Second, our empirical findings in both a field study of fraud investigation groups and a multitask laboratory study corroborate the simulation results. We thereby contribute to research on power by highlighting a dynamic understanding of power in groups and explaining how current implicit assumptions may lead to opposing findings.

Keywords: power, hierarchy, agent-based simulations, team leader competence

Supplemental materials: http://dx.doi.org/10.1037/apl0000056.supp

Power has long been recognized as an essential element in social and organizational relationships (i.e., Bunderson, 2003; Fiske, 1993; Keltner, Gruenfeld, & Anderson, 2003; Lammers & Galinsky, 2009; Russell, 1938). Despite the consensus on power’s importance, whether the concentration of power in one group member, that is, high power disparity, helps or hampers group performance is often debated. On the one hand, the functionalist theory of power posits that high power disparity in groups enhances structure, clarity, and coordination, and thereby group performance (e.g., Halevy, Chou, & Galinsky, 2011; Lammers & Galinsky, 2009; Magee & Galinsky, 2008). On the other hand, the conflict theory of power suggests that high power disparity undermines group performance through higher competition, conflict, and political behavior (e.g., Bloom, 1999; Edmondson, 2002; Eisenhardt & Bourgeois, 1988; Greer & van Kleef, 2010; Siegel & Hambrick, 2005). Clearly, these conflicting views highlight the need for a more integrative theory that better explains the effect of power disparity on group performance.

In this article, we develop and test a theoretical framework that may provide insights into when and why power disparity helps or hurts group performance. Our theoretical framework highlights three implicit assumptions that differentiate functionalist and conflict theories of power, and shows that prior conflicting findings can be reconciled by making these assumptions explicit and by integrating them into theoretical and empirical tests of power disparity. First, the functionalist theory of power assumes a static view in which power does not change over time and hierarchies are self-sustaining (Magee & Galinsky, 2008), whereas the conflict theory of power implicitly suggests that individual ranks in a team can and do change (Greer & van Kleef, 2010). Second, the functionalist theory of power assumes that the most competent individual, that is, the individual who meets task requirements the most (for a review, see Morgeson, Delaney-Klinger, Mayfield, Ferrara, & Campion, 2004), is the most powerful group member (Anderson & Brown, 2010; Van Vugt, 2006; Weber, 1978). By contrast, the conflict theory of power presumes that competence and power may not always be aligned, and that this misalignment may lead to conflict and power struggles within groups (e.g., Bendersky & Hays, 2012). Third, on the one hand, advocates of the functionality of power disparity assume that equality is not possible or sustainable—that groups will always gravitate toward disparity (Gruenfeld & Tiedens, 2010; Magee & Galinsky, 2008); on the other hand, the conflict theory of power (e.g., Greer & van
Kleeff, 2010) and other related areas of research, including that on organizational leadership (e.g., Carson, Tesluk, & Marrone, 2007; Morgeson, 2005), assume that equality is possible, and at times, even preferable.

To investigate the validity of these implicit assumptions and to create a more integrative theory of power disparity in groups, we implement a two-step hypothesis-building and hypothesis-testing approach. First, using agent-based simulations, we experiment with how group performance changes when (a) an individual’s power dynamically changes from one period to another, (b) the most competent group member is (or is not) allowed to hold power, and (c) power can possibly be distributed equally within a group. Second, we test our theory that emerges from the simulation using a multisource field study of 46 fraud investigation groups and a controlled, longitudinal multiround laboratory experiment with 76 interacting groups.

We make four contributions to power and hierarchy research and the literature on leadership and groups. First, we elucidate the controversy surrounding the nature of the relationship between power disparity and group performance by providing a more fine-grained comparison of functionalist and conflict theories of power, which allows us to develop a more integrative theory of power disparity in groups. Second, we show that, depending on the team or context, power can be both static and dynamic, which extends earlier studies that considered power to be a static concept (e.g., Anderson & Galinsky, 2006; Gruenfeld, Inesi, Magee, & Galinsky, 2008; Halevy, Chou, Galinsky, & Murnighan, 2012) or advocated only a dynamic perspective (e.g., Aime, Humphrey, DeRue, & Paul, 2014; Cronin, Weingart, & Todorova, 2011; Pettit, Sivanathan, Gladstone, & Marr, 2013; Sturm & Antonakis, 2015). Third, we contribute more broadly to research on leadership and groups by demonstrating that the task competence of the leader (i.e., the power holder) interacts with power disparity and plays an active role in determining the relationship between power disparity and group performance. Fourth, in response to recent critiques of the methods used in power research (Sturm & Antonakis, 2015), we propose a simulation technique that can enrich scholars’ arsenal for investigating complex phenomena, such as power and leadership within groups (Vancouver & Weinhardt, 2012).

Divergent Views on Power Disparity

Given its central position in social and organizational relationships, power has been defined in many ways (for a recent review, see Sturm & Antonakis, 2015). Amid this conceptual diversity, researchers most often converge on a definition of power that refers to a position that allows a person to carry out his or her will despite resistance (Sturm & Antonakis, 2015, p. 139; Weber, 1978, p. 53). For example, an individual who ranks higher in the organizational hierarchy can be said to have more power over an individual with a lower rank because the former’s formal position bestows him or her with a legitimate right to influence and the means to reward or sanction. When individuals work in groups, the groups vary in terms of how individual power is distributed among group members. To capture and study this variation, we define power disparity within a group as “the differences in the concentration of power among group members” (Greer & van Kleef, 2010, p. 1032; Harrison & Klein, 2007). According to this definition, a group’s power disparity reaches its highest level when power is concentrated in the hands of one group member, whereas disparity is at a minimum when power is distributed equally (Harrison & Klein, 2007).

Do groups with higher or lower power disparity achieve better performance? This question has divided research on power disparity into two distinct and opposing camps. On the one hand, the functionalist theory of power regards power disparity as a source of structure and order, and suggests a positive relationship between power disparity and group performance. Recent reviews (e.g., Halevy et al., 2011; Magee & Galinsky, 2008) have argued that higher power disparity supports divisions of labor, clarifies roles, enables coordination, and facilitates the distribution of resources. For example, Halevy et al. (2012) recently found that high power disparity improved cooperation and coordination in basketball teams, which led to better performance. Similarly, He and Huang (2011) found that boards of directors with high power disparity performed better than boards with low power disparity.

On the other hand, research on the conflict theory of power argues that higher power disparity within groups impairs group functioning and performance. According to this research, high power disparity is detrimental because differences and inequities in the amount of power held by different group members can encourage political and competitive behaviors, as members attempt to move up the ranks and/or protect their valued position in the group hierarchy. These dynamics can cause conflict and inhibit a group’s ability to foster participation, information sharing, learning, and helping behaviors (Edmondson, 2002; Eisenhardt & Bourgeois, 1988; Greer & van Kleef, 2010; Siegel & Hambrick, 2005). Indeed, Greer and van Kleef (2010) showed that power disparity creates power struggles and harms team dynamics, particularly in politicized contexts. Similarly, in a study of major league baseball teams, Bloom (1999) found power disparity to harm individual and group performance because of the competitive and conflictual dynamics that power differences engender. Reconciling these two divergent propositions on the effects of power disparity on group performance and building a more integrative theory of the effects of power disparity on group performance require a closer examination of the implicit assumptions made by both sides in this debate.

Assumption 1: Power Is Static

According to the functionalist theory of power, a hierarchy provides clarity in terms of roles and coordination (e.g., Halevy et al., 2011). This benefit stems from the premise that hierarchies are generally stable and that individuals’ power does not change. Functionalist power theorists refer to system justification processes, which allow power disparities to become self-reinforcing over time (for a review, see Magee & Galinsky, 2008). By contrast, the conflict theory of power argues that power differences create a space for competition and conflict, which may result in dynamic shifts in individuals’ power rankings within the group (e.g., Greer & van Kleef, 2010). Conflict theorists of power posit that individuals continuously compare their power with that of others, as power is a valued resource in organizations (Greer & van Kleef, 2010), and when these comparisons are unfavorable, individuals deliberately attempt to change their power positions (van Dijke & Poppe, 2003) via, for example, intragroup power struggles (e.g., Greer & van Kleef, 2010) and conflicts (e.g., Bendersky &
Hays, 2012). In fact, that individuals’ power can and does change is central to the conflict theory of power; this possibility of change is why people fight for power.

Despite their theoretical differences on the stability of power, empirical studies in both streams overlook changes in individuals’ power and do not discern who gains power from one time period to another. For example, experimental researchers in social psychology often randomly assign power to an individual using different priming methods that remain constant over time (e.g., Anderson & Galinsky, 2006; Galinsky, Magee, Gruenfeld, Whitson, & Liljenquist, 2008; Galinsky, Magee, Inesi, & Gruenfeld, 2006; Gruenfeld et al., 2008); in addition, empirical studies associate the level of power disparity with performance only at one point in time (e.g., Halevy et al., 2012; Siegel & Hambrick, 2005). However, to understand whether the predictions of the conflict or functionalist theory of power hold, both dynamic and stable conceptualizations of power should be considered.

Assumption 2: Power Is Based on Competence

Power holders’ competence has been another implicit assumption that sets functionalist and conflict theories of power apart. Task competence refers to an individual’s ability to fulfill the task requirements (Morgeson et al., 2004). Namely, the functionalist perspective assumes that, in a steep hierarchy, the most competent group members obtain the most powerful positions (Anderson & Brown, 2010; Van Vugt, Hogan, & Kaiser, 2008; Weber, 1978). Supporting this assumption, extensive research has shown that groups tend to accord higher ranks to members who exhibit superior abilities (Anderson & Brown, 2010; Berger, Cohen, & Zelditch, 1972; Driskell & Mullens, 1990; Van Vugt, 2006; Van Vugt et al., 2008). The specific abilities required to attain high rank may depend on group-specific tasks (Anderson, Spatharo, & Flynn, 2008; Van Vugt et al., 2008); however, in general, individuals are accorded a higher rank if they exhibit expertise related to the group’s technical problems as well as social and leadership skills (Van Vugt, 2006). Van Vugt et al. (2008) argued that expertise is positively associated with a powerful position, and that those whose expertise does not match their position are overthrown. However, the conflict theory of power suggests that mismatches in power and competence occur and can lead to ongoing conflict and competition within the group. For example, Bunderson (2003) showed that groups are not always successful in identifying the most competent member and in furnishing that person with power. Indeed, groups have often been shown to overestimate the competence of the most powerful individuals (see also Anderson & Brion, 2014), and people have been shown to rise to power for reasons other than competence, such as interpersonal dominance (Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013) or a personal sense of power (Anderson & Kilduff, 2009).

To understand how this implicit assumption drives the effects found in studies of functionalist versus conflict theories of power, we argue that highlighting and testing whether power in groups is indeed associated with competence is important. For example, if a group is able to identify the most competent group member, and allows for power to concentrate in that person, the power holder can then guide the group toward solutions that offer higher performance. However, if the group member who seizes power is, in fact, not the most competent member, the power holder may create a culture of destructive conflict within the group (Aime et al., 2014). Therefore, whether the conflict or functionalist theories of power are more likely to hold depends on the power holder’s task competence.

Assumption 3: Power Equality Cannot Exist

The final pivotal assumption that we examine concerns the presumed ubiquity of social hierarchy. Functionalist power theorists posit that both humans and primates are stratified by nature, and that organizational groups are no exception (Anderson & Brion, 2014; Magee & Galinsky, 2008; Van Vugt, 2006). In fact, formal and informal hierarchies inevitably emerge within groups (Anderson, John, Keltner, & Kring, 2001), partly because they provide functional solutions for group coordination (Haley et al., 2011).

However, conflict power theorists suggest that hierarchies can have disadvantages, and that power equality can exist and motivate cooperation within groups (e.g., Greer & van Kleef, 2010). As contemporary organizations are increasingly characterized by peer networks that surpass hierarchical layers, transformational relationships that replace carrot-and-stick coordination mechanisms and incentives, and flatter organizational structures with interconnected capabilities and processes (Siggelkow & Rivkin, 2005), the existence of power equality is becoming increasingly likely. Indeed, research has shown that more egalitarian power structures, such as shared leadership, can and do exist, and they may benefit group outcomes (for a recent review, see Wang, Waldman, & Zhang, 2014). Thus, if groups with perfectly egalitarian power distributions were possible, would these low-disparity groups outperform their higher disparity counterparts? Therefore, whether conflict or functionalist theories of power disparity hold may depend on whether groups allow power to be equally distributed among their members.

Overview of Studies

Our objective is to reconcile the functionalist and conflict theories of power by demonstrating the central role of these underlying assumptions in explaining the potential positive or negative relationships between power disparity and group performance. We believe that scholars can develop a more comprehensive and integrated theory to explain the relationship between power disparity and group outcomes only through a deeper understanding of the prevalent assumptions in power research.

We examine the power disparity phenomenon and the inconsistent findings in the literature with a two-step theory-building and theory-testing approach. In the first step, we use agent-based simulations to investigate cases in which an individual’s power changes or remains static, and power is centralized in the group.
member with the highest competence; in addition, we compare these cases with groups with an equal distribution of power. In the second step, we empirically test the findings emerging from the simulation study in a laboratory and field setting. Study 2 closely complements the simulation study using a laboratory experiment in which power is allowed to shift in a dynamic manner based on competence or other characteristics. This experiment allowed us to test the theory that emerged from our simulation study and the aforementioned assumptions—that power disparity is only functional when it is dynamic and aligned with task competence. In all other situations, power disparity may cause conflict and detract from group performance. Finally, in Study 3, we seek to establish the external validity of our findings and test our emergent theory using field data from 46 fraud investigation groups.

**Study 1: An Agent-Based Simulation Study of Power Disparity**

According to French and Raven (1959, p. 259), “the processes of power are pervasive, complex, and often disguised.” For such complex phenomena, agent-based simulation methods are regularly applied as powerful theory-building tools to study the behavior of groups (Davis, Eisenhardt, & Bingham, 2007). We use agent-based simulations as a computational laboratory to investigate “what might be” (Burton & Obel, 2011) when the aforementioned assumptions either hold or do not hold.

We focus on three different types of groups defined in terms of power disparity and the power holder’s high or low competence. First, we consider a group with high power disparity and low competence, such that all group members have low power and a stable leader has high power. This model focuses on the first assumption, according to which group members’ power remains stable. Additionally, this model relaxes the second assumption that the leader always has high competence. Second, we consider a group with low power disparity, in which each group member has the same amount of power. This model targets the third assumption and allows a perfectly egalitarian distribution of power. A low-power-disparity model does not involve the contingency of competence because finding the most competent group member and assigning power to that person does not make sense when each group member has the same amount of power. Third, we consider a group with high power disparity and high competence. This model dynamically assigns power to the most competent individual in each period. Hence, this model addresses the first assumption, in which individuals’ power changes from one period to another; the second assumption, in which the most competent group member is allowed to wield power; and the third assumption, in which the power holder is allowed to change.

We create a simulation model as follows. We simultaneously simulate three groups characterized by the aforementioned models (i.e., high-power-disparity/low-competence, high-power-disparity/high-competence, and low-disparity). Each group collaboratively searches for solutions in a landscape with a certain level of complexity, and the solutions are associated with a certain performance outcome (i.e., the altitude on the landscape). Individuals have a certain amount of power that affects their influence on and resistance to group members, which thus affects their position and the velocity of their movement in the landscape. That is, our simulation model includes the following aspects: (a) power and power disparity, (b) problem space, (c) search, and (d) competence.2

First, we represent an individual’s power with a single parameter, \( m \), which signifies his or her power relative to that of other group members. In a low-power-disparity group, all group members have the same level of power, whereas one group member has high power and others have low power in a high-disparity group. This conceptualization distinguishes power from its antecedents and effects (Sturm & Antonakis, 2015). That is, we model individuals’ power regardless of the basis of their power. The exercise of power is not included in this definition, but power’s influence is expressed as the change in another individual’s position. As a result of interactions with powerful others, a group member is influenced more by powerful members than by individuals with little power.

Second, we present a task with a landscape in which group members search to find peaks. A landscape maps the space of available solutions onto a performance outcome, and individuals aim to find the best solution. We utilize the Gaussian landscape generator (Gallagher & Yuan, 2006). A landscape is generated using a mixture of multivariate normal distributions (i.e., Gaussian functions). Intuitively, a single Gaussian function can be considered as a hill, and a mixture of Gaussians creates a rugged landscape, in which the performance outcome is simply defined as the maximum value over all Gaussians. This approach allowed us to create landscapes with a similar complexity structure within a continuous task space.

Third, we model group members collaboratively searching in a landscape to find the peaks, and assume that their decisions are steered by the quality of the solutions and power differences, in which group searches can still yield outcomes by chance. This assumption is consistent with the three traditional decision-making paradigms (i.e., bounded rationality, politics and power, and garbage can paradigms). Eisenhardt and Zbaracki (1992, p. 32) summarized these paradigms as follows: “Most scholars believe that people are boundedly rational, that decision making is essentially political, and that chance matters.” A well-known, agent-based simulation method called the particle swarm optimization (PSO) algorithm provides the necessary tools to model group searches and interactions along these lines. In the standard PSO models, individuals’ movement is influenced by the best solution they have found thus far as a result of their personal search and the best solution found by any group member (i.e., the bounded rationality paradigm), while room remains for randomness in the search (i.e., the garbage can paradigm). However, the standard PSO model does not consider the effect of power differences on the search process. Hence, we adapt PSO such that individuals’ interactions related to power differences are embedded in the search process. The group search is now also affected by social influences related to power differences, which are proportional to power (i.e., power paradigm).

Fourth, we model an individual’s competence in terms of the quality of the solutions that the individual has previously found. In the high-disparity/high-competence case, the individual who finds the solution that provides the highest performance value

---

2 Please refer to the online supplement for technical details and parameter selection.
will obtain the highest power in the following period, whereas low performers will lose power. We recognize that individuals may have high power not only for being star performers, but also because they have other competencies, such as leadership and managerial skills. Nevertheless, this assumption is line with status characteristics theory, which argues that power in a decision-making group is distributed according to the performance expectations that individuals hold for themselves and other members (Berger et al., 1972; Bunderson, 2003).

Experimental Setting

Our main experiment compares the performance of different disparity levels with respect to the power holder’s task competence under varying levels of complexity. In intellective and problem-solving tasks, group performance is often measured as the best solution produced individually or collectively by any of the group members (McGrath, 1984). Hence, we define group performance as the best solution found by the group when its search ends. We alternatively measured group performance as the median solution and obtained qualitatively similar results because this measurement does not affect the search process—it only quantifies the final outcome. The model’s timeline and parameters are as follows. A task was created by using the Gaussian landscape generator between \([-2, 2]\) with a moderate complexity ($\gamma = 10$). The global optimum, $\Psi^*$, and the second-best peak, $\rho_2^*$, were set at 1 and 0.75, respectively. Then, a group of 10 individuals, which is a team size often observed in organizational teams (Carton & Cummings, 2012) and sufficiently large to allow variations in group search, was randomly positioned in the landscape, and random initial velocities drawn from the uniform distribution were assigned. In each time period, group members searched the landscape and updated their positions according to Equation 4 in the online supplement. This search continued for $T = 100$ periods.

We ran each power model on the same landscape with the same initial positions and velocities. Hence, any observed difference in performance can solely be attributed to the power models. To smooth random variations, we created $S = 1,000$ tasks for each complexity level. Regarding the values of power, high power (i.e., $m_{\text{high}}$) was defined as a mass of 0.9, whereas low-power individuals had $m_{\text{low}}$ values of 0.1. Additionally, we solely focused on a five-dimensional space. The results were robust for varying levels of dimensionality and numbers of iterations. In addition to these parameters that were specific to our study, we tuned PSO parameters to ensure that groups converged in their search, as suggested by Clerc and Kennedy (2002). They suggested making the coefficients for local, $c_1$, and group search, $c_2$, equal. Similarly, we set the coefficient for the gravitational attraction, $c_3$, equal to the other two acceleration coefficients. The equal acceleration coefficients enable a balance between individual and group search and movement related to power differences.

Analysis and Hypothesis Development

In this section, we present the results of our simulation experiments and develop a hypothesis regarding the interplay between power disparity and competence. We proceed as follows to interpret the simulation results. First, we investigate how the search process occurs in the different power models. This step lays out the mechanisms that drive performance differences between the power models. Second, we compare the different models in terms of their performances. We summarize the latter result in the form of a hypothesis.

Figure 1 depicts how different power models search as the hierarchy becomes steeper, that is, as the power level of the power holder increases. On the $y$-axis of the upper panel, we plot the similarity between individuals’ positions across a two-period time lag. This similarity indicates the average autocorrelation between individuals’ positions over time. The $x$-axis indicates the power holder’s power in high-disparity cases and group members’ power in the low-disparity case, in which all members have equal power. In Figure 1, we observe that the solutions that group members in the high-power-disparity/low-competence and low-disparity models find over time are quite similar. Their solutions are similar because the center of gravity that attracts group search in these models does not change over time. The former group is attracted to the power holder, whereas the latter group moves toward its center. Because the direction of the search remains stable, group members visit correlated positions. By contrast, the center of gravity in the high-disparity/high-competence model changes: A different individual with the best performance gains high power. This change reorients the group search toward the individual with the best solution. Consequently, in Figure 1, we observe that individuals’ positions are less similar in the high-disparity/high-competence model than in the other models. This search behavior implies that high-disparity/high-competence groups reorient their search in each period, whereas low-disparity groups search based on the group’s average opinion, and high-disparity/low-competence groups are confined to the vicinity of the power holder’s position.

To understand how these search behavior differences manifest themselves in performance differences, Figure 2 compares the performances of different power models. It shows that high-power-disparity groups perform better than low-disparity groups only if the central power holder is highly competent at the group task. However, if the power holder is not competent, high-power-disparity groups perform worse than the low-disparity groups. When power is concentrated in the most competent group member who provides the highest quality solutions (i.e., high competence) in a high-power-disparity group, the power holder directs the remainder of the group toward better solutions. Because of the previous success, the power holder gains more of an influence over other group members.

Figure 1. Average autocorrelation between the group members’ positions during the search in a moderately complex problem space.
members. Whereas members of low-power-disparity groups move toward center of the group, which does not necessarily offer better performance, merit-based groups with high disparity win power battles and convince group members to pursue higher quality solutions, thus achieving superior performance. Consequently, we propose that power disparity is positively associated with group performance when the power holder is highly competent.

However, if power is not concentrated in the hands of the most competent group member, the power holder may derail the group from finding quality solutions that could offer higher performance. First, the power holder will influence group members toward the desired solution, and this solution is more likely to be suboptimal. Second, the power holder will eventually restrain the group from moving toward better solutions, as Figure 1 shows that the direction of group search changes little over time. For example, Tost, Gino, and Larrick (2013) documented that centralized power holders suppress others’ voices, which may be particularly problematic when the power holder is incompetent. Although the search direction in groups with low power disparity also changes little, members in these groups have the opportunity to influence group decisions. The wisdom of the group supersedes the solution that the power holder has suggested. For example, Bunderson (2003) showed that expertise drives searches in low-power-disparity groups. Hence, although one person may propose an inadequate solution, the rest of the group can continue searching for alternative solutions and weigh the proposals according to their quality rather than considering who suggested them. Therefore, we argue that power disparity is negatively associated with group performance when the power holder is not competent. These arguments lead us to the following hypothesis:

Hypothesis 1: The power holder’s task competence moderates the relationship between power disparity and group performance. That is, compared with groups with low power disparity, high-disparity groups perform better when the power holder is (dynamically) aligned with task competence, and perform worse when the power holder is not (dynamically) aligned with task competence.

Sensitivity Analyses

We conducted three additional simulation experiments as robustness checks to investigate the effects of problem complexity, group size, and the acceleration coefficient for gravitational attraction. We investigated these effects because they may affect the extent to which group members are influenced by power differences and the difficulty of finding the best solutions.

Figure 3 compares the performance of different power disparity models under varying degrees of problem complexity. As the problem space becomes more complex, finding the optimal solution becomes more difficult because the landscape becomes rugged. Consequently, the performance outcome of a group search decreases, as shown by the downward sloping performance curves of the power disparity models. However, our main result remains stable, as the high-power-disparity groups outperform the low-power-disparity groups only if the best performer is granted power under varying degrees of problem complexity.

We also compared power models while varying the group size. Larger groups can span wider areas in a landscape; consequently, we observed higher average performances in larger groups for all power disparity models. However, the relative performance differences did not change. Furthermore, the acceleration coefficient for power differences, \( c_3 \), in relation to the acceleration coefficients for personal and group search, \( c_1 \) and \( c_2 \), respectively, determines the extent to which an individual is influenced by power differences. We checked the robustness of the findings in our main simulation experiment by varying \( c_3 \). Because \( c_3 \) indicates the degree of influence by power difference, the performance difference between power models faded away as \( c_3 \) approached zero. However, our results remained qualitatively similar as long as \( c_3 \) was sufficiently large.

Study 2: An Empirical Analysis of Power Disparity and Group Performance

The objective of this study is to explore the validity of the three implicit assumptions of power research and to causally test our central hypothesis in a laboratory setting. This laboratory experiment complements the simulation in three important ways. First, it
provides empirical evidence for the predictions of the model. Second, it allows individuals’ power to change from one period to another, thus enhancing our understanding of power as a dynamic phenomenon. Third, the experiment provides explorative insights regarding who gains power in a group and when that power changes hands.

Method

Participants and design. Two hundred ninety-four students (185 women, 109 men) from a large private university in the United States participated in this study. The average age was 22.89 years ($SD = 4.61$). We manipulated the level of power disparity within a group (low vs. high power disparity) in a between-subjects design.

Procedure. After being welcome, participants completed a brief survey to provide demographic information. Participants were then randomly placed in teams and were informed that they would work on a task as a team. Because of the idiosyncratic arrivals, the teams had three to five members ($M = 3.54, SD = 0.66$). In total, 83 teams were formed. To encourage individuals and teams to engage in the task, the lab assistant informed the participants of two lotteries. First, the participant with the best individual score would receive $50 at the end of the study. Second, each member of the team with the best team score would receive an additional $50.

After the participants were assigned to teams, power disparity was manipulated. The participants in the low-disparity condition had no central power position assigned. They were instructed as follows:

You will now work together to survive in this desert. You should discuss how to solve this task together. You have 10 minutes for group discussion. When the time is up, I will come in and collect the final rankings.

In the high-power-disparity condition, teams were asked to select a power holder who would be in charge of managing group discussions, making the final call if there were disagreements, and handing in the final rankings.

The participants worked on the desert survival task (Lafferty & Pond, 1974). The task has participants consider a scenario in which they are left in the desert; to survive, they need to rank 12 items in order of importance for survival in the desert. The participants worked on the desert survival task for two 10-min periods. Between the first and second round, teams in the high-power-disparity condition were informed of their team and individual performances. After this feedback was provided, the teams in the high-power-disparity condition were allowed to reselect a power holder. Then, all teams were given the chance to rerank the items in the second 10-min session. These deliberate design choices allowed us to explore whether the power holder would change from one period to another (i.e., Assumption 1) and whether the teams would select the team member with the highest competence to be the power holder (i.e., Assumption 2). After the second round, the study was complete, and the participants were thoroughly debriefed and thanked for their participation.

Measures. The main dependent variable is team task performance. We assessed a team’s final performance as its ranking of 12 items with respect to the expert’s rankings. We used Spearman’s rank correlation coefficient because of the ordinal nature of the rankings. The performance score increases according to the degree to which the team rankings match the expert’s rankings.

In line with Study 1, our main independent variable includes three categories of power disparity. The low-disparity category refers to the teams in which no formal power holder was assigned (power disparity = 0). In cases in which the teams were asked to select a central power holder, we categorized the team as a high-disparity/low-competence group if the team member with the highest competence was not given the central power position (power disparity = 1), and as a high-disparity/high-competence group otherwise (power disparity = 2).

Similar to Study 1, the experimental task requires team members to find items that can help them survive in the desert. Hence, we define competence as the quality of the solutions provided by an individual and measure it as the participants’ individual performance scores in the first round of the team task. As with team performance, individual performance was measured by comparing individuals’ rankings of survival items to the expert’s ranking.

Additional variables. We included several additional variables to use in our manipulation checks and to explore the drivers of shifts in individuals’ power and their appointment to the central power position in the team. We measured two variables in the manipulation checks. First, at the end of the study, team members were asked to indicate whether they were required to select a team power holder in each period. Second, to ascertain whether teams in the low-power-disparity condition had lower levels of hierarchical stratification, we asked each participant a round-robin question to assess the power that other team members had during team decision making (Anderson et al., 2008; Greer, Caruso, & Jehn, 2011; Venkataramani & Tangirala, 2010). This question was used to measure individuals’ perceived power and the perceived degree of power disparity within the team. We calculated each team member’s perceived power as the average ratings of his or her teammates. To quantify the degree of perceived power disparity within the team, we calculated the coefficient of variation in individuals’ perceived power (Harrison & Klein, 2007).

In addition, we measured participants’ physical cues, such as age, gender, height, and weight, because psychical appearance is often associated with power and competence (Bunderson, 2003; Schubert, Waldzus, & Giessner, 2009). We also measured whether participants characterized their stable personalities using adjectives related to dominance. These adjectives were retrieved from Wiggins’s (1979) Interpersonal Adjective Scale. We used a 7-point scale, ranging from 1 (not at all) to 7 (extremely). We calculated dominance as the average of dominance adjectives and reversed submissive adjectives (Cronbach’s alpha = .85; Tiedens, Unzueta, & Young, 2007). Pearson correlations among these additional variables are given in Table 1.

Results

Manipulation checks. We conducted two manipulation checks. First, when we asked participants whether their team was required to select a formal power holder, seven teams did not agree with their experimental condition. We discarded these teams from our analyses. Second, we checked whether more perceived power disparity existed in high-disparity teams than in low-disparity teams. A $t$ test showed that teams in the high-power-disparity
condition perceived more disparity than teams in the low-power-disparity condition, \( t(55) = 1.97 \), \( p < .05 \), one-tailed.

**Hypothesis testing.** We conducted a one-way ANOVA to compare the performance of teams with low disparity, high-disparity/low-competence or high-disparity/high-competence. We set the low-power-disparity condition as the baseline. Our results showed significant performance differences across the power disparity conditions, \( F(2, 73) = 11.16, p < .01, \omega = 0.46 \) (Bartlett’s test of equal variances = 0.44, ns). Follow-up planned contrasts revealed that teams in the low-power-disparity condition \( (M = 0.43, SD = 0.23) \) performed significantly better than teams in the high-disparity/high-competence condition \( (M = 0.26, SD = 0.26) \), \( r(73) = -2.34, p < .05, r = .26 \). However, as predicted, teams in the high-disparity/low-competence condition \( (M = 0.57, SD = 0.25) \) performed significantly better than teams in the low-power-disparity condition, \( r(73) = 1.98, p = .05, r = .23 \), and teams in the high-disparity/low-competence condition, \( r(73) = 4.72, p < .01, r = .48 \). These results provide empirical support for our hypothesis that high power disparity helps (hurts) teams when the power holder has high (low) competence.

**Exploring Assumption 1.** Although we did not explicitly hypothesize regarding if and why teams changed their power holders, our experimental design afforded us the opportunity to gain initial explorative insights into these hierarchical dynamics. Our results show that 41% of the teams changed their power holders, suggesting that the assumption that hierarchy is stable, as we have purported, is incorrect. By contrast, hierarchies can be dynamic as well.

Next, we conducted a probit regression to predict the likelihood of a power change in relation to team- and member-level drivers of this power shift. The results in the first column of Table 2 show that teams were significantly less likely to change their central power holder when the power holder had high competence. We also documented a slightly significant effect of perceived power disparity and the power holder’s height. These results demonstrate that individuals’ power is not always stable and that a power shift may occur based on member characteristics (e.g., competence and height) and perceived power disparity.

**Exploring Assumption 2.** Our lab experiment allowed us to also explore the assumption that power differences are aligned with differences in competence. We observed that only 55% of the teams initially selected the most competent team member to be the central power holder (i.e., the ratio of teams in high-disparity/high-competence condition). Clearly, this percentage is quite lower than the case in which all power holders in high-disparity conditions have high competence, as Assumption 2 posits.

Next, we explored the other potential reasons that a team member was selected as the central power holder. We conducted a probit regression to estimate whether a participant was attributed high power by the team members and assigned as the power holder. The results are presented in the second column of Table 2. Although competence was the strongest predictor, its effect was small, suggesting that many factors other than competence are likely to determine who the group first chooses as leader. These

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Selected as leader</td>
<td>.13</td>
<td>.04</td>
<td>.11</td>
<td>.02</td>
<td>.05</td>
<td>.01</td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Team leader change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Gender diversity</td>
<td>.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Age diversity</td>
<td>.17</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Perceived power disparity</td>
<td>.27</td>
<td><strong>.15</strong></td>
<td>.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Task competence</td>
<td>-.30</td>
<td><strong>-.36</strong></td>
<td>.02</td>
<td>-.01</td>
<td>.07</td>
<td>.14</td>
<td>.09</td>
<td>-.04</td>
<td>.05</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>7. Age</td>
<td>.04</td>
<td>.01</td>
<td><strong>.28</strong></td>
<td>.11</td>
<td>.03</td>
<td>-.02</td>
<td>.06</td>
<td>-.06</td>
<td>.11</td>
<td>-.02</td>
<td></td>
</tr>
<tr>
<td>8. Height</td>
<td>.12</td>
<td>.07</td>
<td>.09</td>
<td>.06</td>
<td>.19</td>
<td>.04</td>
<td>.69</td>
<td><strong>-.59</strong></td>
<td>.04</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>9. Weight</td>
<td>.05</td>
<td>-.03</td>
<td>.25</td>
<td>.25</td>
<td>.23</td>
<td>.06</td>
<td>.71</td>
<td><strong>-.55</strong></td>
<td>-.01</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>10. Gender</td>
<td>-.00</td>
<td>-.24</td>
<td>-.02</td>
<td>-.09</td>
<td>-.05</td>
<td>.06</td>
<td>-.50</td>
<td><strong>-.50</strong></td>
<td>-.59</td>
<td><strong>-.16</strong></td>
<td>-.09</td>
</tr>
<tr>
<td>11. Dominance</td>
<td>.13</td>
<td>.04</td>
<td>.15</td>
<td>.23</td>
<td>.15</td>
<td>.17</td>
<td>.21</td>
<td>-.25</td>
<td><strong>-.16</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Perceived power</td>
<td>.03</td>
<td>.00</td>
<td>.22</td>
<td>.46</td>
<td><strong>.17</strong></td>
<td>.31</td>
<td><strong>.10</strong></td>
<td>.21</td>
<td>-.09</td>
<td><strong>.32</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** The lower triangle presents the correlations between team-level variables and the leader’s characteristics in Assumption 1 (\( N = 56 \)). The upper triangle presents the correlations between individual-level variables in Assumption 2 (\( N = 193 \)).

- \( p < .10 \)
- \( ** p < .05 \)
- \( *** p < .01 \)

With \( \alpha = .05 \), we applied a Bonferroni correction for multiple comparisons, resulting in a corrected \( \alpha = .01 \).
results do not support Assumption 2, as the most competent team member does not always obtain the high-power position.

**Exploring Assumption 3.** Is an egalitarian distribution of power possible? While we experimentally manipulated whether high or low power disparity exists, our laboratory experiment allows us to check whether an informal power hierarchy emerges in the low-power-disparity condition. The average perceived power disparity in teams with low power disparity was 0.13 (SD = 0.07). We conducted a $t$ test and found that this value is in fact significantly different than its theoretical minimum, that is, zero (Harrison & Klein, 2007), $t(19) = 8.72$, $p < 0.01$, two-tailed. Hence, we conclude that, even in teams with low disparity, hierarchies emerge. In addition, we investigated whether teams with egalitarian power structures generally outperformed teams with high power disparity. However, as mentioned in our hypothesis testing section, teams with low power disparity outperformed high-disparity/low-competence teams, though they did not outperform teams with high power disparity and highly competent power holders. This finding shows that equality can exist, and its effectiveness, compared with that of hierarchy, depends on whether a hierarchy can be legitimated in a given situation.

**Discussion**

Study 2 complements our simulation study in two important ways. First, it empirically explores the validity of three implicit assumptions prevalent in power research in a controlled laboratory setting. The results demonstrate that an individual’s power is not always stable and may change dynamically (Assumption 1); teams do not always successfully identify the team member with the highest competence and assign a powerful position (Assumption 2); and, even in the low-power-disparity condition, hierarchical stratification emerges to a certain extent (Assumption 3). Second, Study 2 provides empirical evidence for our main hypothesis that the performance differences between teams with high and low power disparity are contingent on whether the power holder has high or low competence. Although Studies 1 and 2 complement one another and communicate a unified message, neither study provides evidence of whether their results will hold in organizational settings. To remedy this shortcoming, we now present Study 3, in which we tested our main hypothesis in the field.

**Study 3: Field Evidence**

The objective of this study is to empirically test our central hypothesis in an organizational setting, and thereby to improve the external validity of our findings. Providing empirical evidence on the relationship between power disparity and group performance is important for several reasons. First, this evidence suggests that individuals with high power are not always those who provide creative solutions and direction for the group. Second, this empirical evidence complements our simulation study by integrating the predictions of the functionalist and conflict theories of power.

**Method**

**Sample.** We collected data from a public organization in the Netherlands. The organization is composed of 1,190 employees. The management team and support departments account for 5% of the staff. Ninety-five percent (1,126 employees) of the staff actively work in 49 groups to perform the complex intellectual task of detecting tax evasion and fiscal fraud, which includes excise duty fraud, money laundering, insider trading, and property fraud. These investigation groups constitute our target sample. The average investigation group had 23.98 (SD = 8.07) members. The average age of the employees was 49.62 years, 99% were Dutch, and 26.47% were female.

Prior to our data collection efforts, we conducted interviews with the managing director, the head of the workers’ council, a team leader (in this organization, the power holder is called the **team leader**), and a group member. Note that the group size in this organization is larger than those often observed in other work teams (Carton & Cummings, 2012). Our fieldwork showed that these groups work under a team leader; whenever a new project materializes, small, temporary project teams are formed. During the on-site interviews, we explicitly inquired whether employees view themselves as a part of the regular group or the small project teams. Our interviewees unanimously indicated that employees view themselves as belonging to the regular group. In this way, these groups satisfy Kozlowski and Bell’s (2003, p. 334) definition of work teams: two or more individuals who come together to perform organizationally relevant tasks, share one or more common goals, and interact socially.

Furthermore, these interviews also supported our assertion that this organization provides an excellent setting to test our hypotheses. First, the groups are comparable because they perform similar jobs (i.e., detecting fraud). Second, the jobs that the employees perform are nonroutine and intellectual in nature. In addition, our interviewees indicated that power holders’ competence (in our case, team leaders’ competence) played an important role in the group processes because group members consulted with the team leader when they encountered problems. Third, as a public organization with clear hierarchical boundaries, power and power differences are reinforced in day-to-day operations.

We obtained archival data from company records, which included the employees’ age, gender, nationality, group membership, and hierarchical rank. The remaining data of interest were collected via an online survey. The managing director announced the survey, and the researchers sent out the survey 1 week after the director’s communication. Participation in this study was voluntary. The anonymity and confidentiality of the participants’ responses were emphasized in the invitation e-mail and on the first page of the survey.

After excluding the top management team and support teams, we invited 1,126 employees to complete the survey. Of these, 742 participated; the response rate was 66%. On average, 68% of group members (SD = 0.15) responded. Independent sample $t$ tests and chi-square tests revealed no difference between respondents and nonrespondents with respect to age, $t = −1.53$, $p = ns$, and gender, $χ^2(1) = 2.28$, ns. We were required to drop three groups from our analyses because their leadership positions were vacant, which resulted in a sample of 46 groups.

**Measures.** Our dependent variable was group performance. Because the organization did not maintain any objective measure of group performance, we relied on subjective assessments. Our interviews revealed that the following performance criteria were relevant for these groups: efficiency, quality, cooperation, speed, and overall performance. We used a 5-point scale (1 = much
worse; 3 = average; 5 = much better; Cronbach’s alpha = .85). We asked each team leader and two randomly selected group members to assess his or her group’s performance compared with that of other groups (Van der Vegt & Bunderson, 2005; Van der Vegt, de Jong, Bunderson, & Molleman, 2010). We obtained assessments from multiple raters in 39 groups. A median interrater agreement score (Rwgj) of 0.95 indicated that the respondents agreed on their ratings of the performance items. An intraclass correlation coefficient (ICC1) of 0.23 showed that performance scores varied across groups. However, an ICC2 of 0.39 indicated that group means may not have been very stable, which was because of the low number of raters and the formative nature of the performance construct (Klein & Kozlowski, 2000).

Our primary independent variable was power disparity, which we measured in two steps. First, we measured each group member’s formal power using their hierarchical ranks. Previous research has offered three common measurement practices: measuring power in terms of (a) influence (e.g., through a round-robin design), (b) perception (e.g., asking individuals to think of a scenario in which they occupy a powerful position), and (c) control over resources (e.g., hierarchical rank). The first option confounds “what power does” with “what power is” (Fiske & Bernahl, 2007; Sturm & Antonakis, 2015). The second option does not utilize the power relationships already existing in our field setting, because respondents do not need to think of a power scenario, but they live it. Therefore, this option is often applied in lab settings with students. Although the third option primarily focuses on formal power, hierarchical rank is prevalent and activated in field settings, does not confound the outcome with its measurement, and is employed quite often by researchers (e.g., Becker & Baloff, 1969; Greer & van Kleef, 2010; Patel & Cooper, 2014). Therefore, we opted to measure power using the group members’ hierarchical ranks. These data were retrieved from company records. Second, to calculate the degree of power disparity within a group, we calculated the coefficient of variation in individuals’ power, as recommended by Harrison and Klein (2007).

We hypothesized that the effect of power disparity is moderated by the power holder’s task competence. In this organization, we observed that the highest formal power was held by the team leaders. Therefore, we focused on the team leaders’ task competence. In our simulation study, we relied on the quality of the solutions that an individual previously proposed to define task competence; however, such historical and objective data were not available here. Zhou and George’s (2001) validated creativity scale captures task competence in this setting very well because task competence entails developing new ways and approaches to detect diverse criminal activities. Previously, Hirst, Van Knippenberg, Chen, and Sacramento (2011) used this validated scale in a similar context involving Taiwanese customs officers who were attempting to detect trafficking, fraud, and smuggling activities. Using a 5-point scale (1 = strongly disagree; 3 = neutral; 5 = strongly agree), each group member answered the following question regarding his or her leader: “To what extent do you think the following statement is a characteristic of your team leader?” Sample items included “My team leader consistently seeks new ideas and ways to develop new solutions” and “My team leader suggests new ways to achieve goals or objectives.” The measure displayed a high degree of internal consistency (Cronbach’s alpha = .96). Furthermore, the respondents exhibited a high degree of agreement on their assessment of the leader’s task competency (Rwgj = 0.92), which varied sufficiently across teams (ICC1 = 0.23) and presented high consistency (ICC2 = 0.82).

We controlled for job complexity, diversity in tenure, and education-level diversity. When tasks are complex, finding a solution becomes difficult; more importantly, many possible solutions may be offered by group members. Deciding among these proposed solutions creates battles of choice, and competing solutions can amplify the importance of a competent team leader. We controlled for job complexity using four items from Morgeson and Humphrey’s (2006) Work Design Questionnaire (Cronbach’s alpha = .88, Rwgj = 0.88, ICC1 = 0.05, ICC2 = 0.42). A sample item reads, “The tasks on the job are simple and uncomplicated” (reverse scored). Diversity in terms of tenure and education level determines the knowledge resources available to a group, which, in turn, affect its performance. We obtained each team member’s education level and tenure via a single survey question. Consequently, we controlled for diversity in tenure by taking the standard deviation of the team members’ tenure (Harrison & Klein, 2007). Because of the ordinal nature of the education level, we measured diversity in education level using Blair-Lacy’s (1-2°) index (Blair & Lacy, 2000). The means, standard deviations, and Pearson correlations are presented in Table 3. Survey items are presented in the online supplement.

### Analysis and Endogeneity

Our hypothesis proposes that power disparity’s effect on group performance is contingent on the power holder’s perceived competence. An ideal test of this hypothesis requires the random assignment of leaders with high and low competences to groups with varying degrees of power disparities. Although our simulation analyses provided an ideal experimental design, the field study did not (in Study 2, we also allowed team members select their power holders instead of exogenously assigning them). In fact, perceptions of power holder competence may be endogenously affected by unobserved factors (Bunderson, 2003; cf. Sturm & Antonakis, 2015). To account for endogeneity concerns, we analyzed our data using two-stage least squares with instrumental variables (Antonakis, Bendahan, Jacquart, & Lalive, 2014).

We instrumented task competence using the leaders’ hierarchical rank, age, and leadership style. Powerful individuals are more likely to be perceived as competent (Anderson & Brion, 2014). Furthermore, physical cues, such as age, may create biased per-

### Table 3

Descriptive Statistics and Pearson Correlations among Study 3 Variables

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Job complexity</td>
<td>4.17</td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Education level diversity</td>
<td></td>
<td>.72</td>
<td>.05</td>
<td>.44**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Tenure diversity</td>
<td>8.68</td>
<td>1.97</td>
<td>.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Power disparity</td>
<td>.23</td>
<td>.06</td>
<td>.57***</td>
<td>.58***</td>
<td>.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Competence</td>
<td>3.08</td>
<td>.46</td>
<td>.04</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Group performance</td>
<td>3.37</td>
<td>.42</td>
<td>.08</td>
<td>.02</td>
<td>.03</td>
<td>.10</td>
<td>.25</td>
</tr>
</tbody>
</table>

*Note. N = 46 groups. 
*p < .10. **p < .05. ***p < .01.
ceptions of competence (cf. Bunderson, 2003). Finally, leadership style may also affect how a group perceives its leader’s competence. Our theory conceptualizes group members searching for solutions, and leader task competence is defined as the quality of the solutions offered by the leader. Therefore, we measured leadership style in terms of goal-focused leadership using five items developed by Colbert and Witt (2009). A sample item for a leader’s followers reads, “My team leader provides direction and defines priorities” (Cronbach’s alpha = .89). We aggregated the followers’ answers at the leader level (Rwgj = 0.88, ICC1 = 0.19, ICC2 = 0.77).

We conducted two tests to ascertain whether endogeneity exists and, if so, whether our instruments are strong. First, the Wu-Hausman test results show that the coefficients of the OLS model significantly differ from those of the instrumental variables (IV) model, F(2, 37) = 4.79, p < .05. This result indicates that endogeneity exists, OLS estimates are inconsistent, and the IV model should thus be used. Second, an IV model may yield biased results if weak instruments are used, that is, if the instruments do not explain competence. An F test showed that our instruments were strong, evidenced by the large F statistic of the model, with and without the instrumental variable, which was F(6, 35) = 21.64, p < 0.01 (above the suggested threshold value of 10; Antonakis et al., 2014). Consequently, in the following subsection, we will report the results of the IV model.

Results

To increase the interpretability of our interaction effect, we standardized the main effects and then calculated the interaction term by multiplying power disparity and the team power holder’s competence. The results are summarized in Table 4. Model 1 includes only the control variables. None of the control variables significantly explained group performance. In Model 2, we added power disparity and competence, which did not explain group performance. In contrast to these models, the interaction term in Model 3 had a positive and significant coefficient (b = 0.46, p < .05). Because R² does not have a counterpart in IV regressions (Greene, 2012, p. 266), we compared Models 2 and 3 using a chi-square test, which indicated that Model 3 had a significantly better fit than Model 2, χ² (1) = 6.15, p < .05. This result supports our hypothesis that the relationship between power disparity and group performance is contingent on the power holder’s competence.

To better understand the interplay between power disparity and competence, Figure 4 plots the relationship between power disparity and group performance at high and low competence level (1.5 standard deviations above and below the mean, respectively). Accordingly, we observe that power disparity was positively associated with group performance when the power holder was highly competent. This relationship was reversed when the power holder had low competence. These results are consistent with our hypothesis. Additionally, we conducted simple slope tests that support the finding that power disparity was positively associated with group performance when competence was high (b = 0.74, p < .05, two-tailed). By contrast, power disparity was negatively associated with group performance when competence was low (b = -0.64, p < .10, two-tailed). Overall, these results support our moderation hypothesis that the effect of power disparity is contingent on the power holder’s ability to provide high-quality solutions. Simple slope tests showed that higher (lower) power disparity fosters (hampers) group performance when the power holder cannot (can) offer high-quality solutions.

We tested two additional extensions of our model: a three-way interaction among power disparity, competence, and job complexity, and a curvilinear effect of power. The results presented in the online supplement ruled out these alternative explanations. Note that our analyses might have failed to identify those effects simply because of our small sample size (N = 46) or our particular empirical setting. In fact, unearthing these curvilinear and three-way interaction effects and the underlying mechanisms are very promising areas for future research.

General Discussion

Despite the broad consensus on the importance of power (Bunderson, 2003; Fiske, 1993; Keltner et al., 2003; Lammers & Galinsky, 2009; Russell, 1938), considerable disagreement persists regarding whether groups with high or low power disparity achieve superior performance (for reviews, see Anderson & Brown, 2010; Lammers & Galinsky, 2009). Any attempt aiming to reconcile these contradictory views and to develop an integrative theory of power must address existing implicit assumptions. Hence, using agent-based simulations, we first explicitly investigated the following existing assumptions in power research: (a) group members’ power does not change over time, (b) a steep hierarchy ensures that the most competent group member seizes the highest power position, and (c) an equal distribution of power is not possible. This investigation revealed that the hypothesis positing that high power disparity can help or hurt group performance was contingent on whether the power holder had high task competence. Second, using a laboratory experiment, we empiri-
cally explored the validity of these assumptions and found support for our hypothesis. Third, a field study of fraud investigation teams provided additional evidence for our hypothesis and reinforced the external validity of our results.

When we consider the overall results of our multimethod set of studies, our results provide a means to reconcile the functionalist and conflict theories of power. In contrast to the functionalist theory, which favors high power disparity over an egalitarian distribution of power, we found that low power disparity yields higher performance if the power holder is not replaced over time with the most competent group member. However, we found support for the functionalist argument that high power disparity leads to high performance when power is dynamically centralized in each time period and awarded to the most competent member. We contribute to this debate by proposing that the relative level of power disparity does not determine whether power disparity helps or hurts groups; instead, this effect depends on the competence level of the person at the top of the power hierarchy. Our findings thus suggest that the legitimacy of power disparity is critical in determining whether functionalist or conflict theories of power hold.

Our explorative analyses in Study 2 elucidate the validity of the three implicit assumptions in power research. First, we found that both stable and dynamic hierarchies exist. On the one hand, in contemporary organizations, informal forms of power are pervasive, and an individual’s power may change from one period to another as a function of his or her competence. Such a dynamic understanding of power has only recently begun drawing attention (cf. Cronin et al., 2011; Pettit et al., 2013). On the other hand, we showed that individuals’ power may remain stable, contingent upon the perceived power disparity and the power holder’s characteristics. As such, we propose that the literature on power should not be confined to either a static or dynamic conceptualization of power; it should instead consider and integrate the fact that some individuals’ power changes dynamically and others’ remain unchanged at the same time. Second, we report that although competence is a slightly important factor in predicting an individual’s appropriation of a powerful position, not all hierarchies allocate power to the team member with the most competence. This result calls for a richer understanding of when and how power and competence align. Third, we showed that informal power disparity emerges even in egalitarian groups, albeit to a lesser degree than in hierarchical groups, and benefits these groups more than high-disparity groups with less competent leaders. How such a low degree of power disparity emerges in egalitarian groups and how it helps group performance are fruitful avenues for future research.

Furthermore, our article also offers interesting insights for the leadership literature. First, our findings speak, for example, to the debate in the leadership literature about whether traditional models of leadership, in which leadership is centralized in a single individual (i.e., high power disparity), are more or less functional than leadership models in which leadership is shared across team members (i.e., lower power disparity; for a recent review, see Wang et al., 2014). We find that centralized leadership, when aligned with member competence, may be the most beneficial structure for teams. However, shared leadership can also enhance performance and certainly produces better performances than centralized leadership when the member in charge lacks competence. We should note that these findings do not advocate simply rotating leadership across members (e.g., Davis & Eisenhardt, 2011); instead, they suggest that power should be allocated to the most competent member. Second, our findings speak to the literature on leader–member exchange (LMX; for recent meta-analyses, see Dulebohn, Bommer, Liden, Brouer, & Ferris, 2012; Rockstuhl, Dulebohn, Ang, & Shore, 2012). Our results suggest that hierarchical relationships between a leader and followers are most rewarding when the leader has high task competence. By contrast, when the leader does not have high task competence, minimizing the power differences between leaders and followers is advisable.

Finally, this article contributes to formal approaches in power research by proposing a novel agent-based simulation technique, namely, PSO. Developing such tools increases the alternatives readily available to scholars. For instance, our simulation method can be readily applied to study complex nonlinear relationships and feedback loops between a leader and his or her followers. LMX scholars can study performance effects attributable to various leader and follower characteristics (e.g., power differences between the leader and follower, their noisy perceptions of solution quality, and performance feedback), interpersonal relationship characteristics (e.g., the communication network structure and friction in top-down and bottom-up communication), and contextual variables (e.g., task complexity and landscape change) in controlled combinatorial laboratories of agent-based simulations (Burton & Obel, 2011).

Managerial Implications

Our findings provide important implications regarding how managers and leaders should structure power within teams and organizations. Namely, we show that organizations should only have power disparity within teams when the teams are able to identify the most competent member and when all members are willing to shift power when task requirements (and, thereby, relative competence levels) change. This finding suggests several actionable steps for managers. First, if managers have teams with high coordination demands and wish to use power disparity (or hierarchy) within the teams to coordinate and order interactions, making sure that the members with the most competence are given the most power is critical. Hence, our findings point to the vital importance of data-driven hiring and promotion practices that are based on objective assessments of actual task competence (e.g., matching qualifications with task requirements), such that powerful positions in the organization are obtained based on competence rather than political skill or impression management. Performance reviews and quantitative indicators of task performance (earnings, sales, etc.) can also be useful indicators to draw on when ensuring that power is aligned with task competence.

Second, managers should help team members understand the unique skills that all members bring to the team (via, e.g., cross-training or job-crafting interventions). By helping members identify and understand the different competences that members bring to the team, shifting power (such as decision control) within the team when task requirements change will be easier. Taken together, we highlight the role of data-driven promotion practices in assessing who should move up the corporate ladder; in addition, we emphasize the importance of hiring individuals who are innately willing and able to share leadership with others when others have more competence, and help members know the different
competences that they bring to the team to facilitate dynamic power shifts when team tasks evolve and change.

Additionally, our findings also suggest that managers should only choose to concentrate power in a team if the team is capable of successfully detecting the “stars” and implementing a policy that flexibly assigns power to those stars. If no such capability exists, our results suggest that low power disparity may be more functional, giving credence to the rising trend of flatter organizational structures, such as holacracy. When teams are unable to achieve the fluidity to shift power whenever task demands change, conflicts may emerge, as ill-suited leaders may breed resentment among followers and cause hierarchies to be sources of contention rather than coordination. In such situations, organizations may be better served by having a more egalitarian power structure as the basis for social interactions.

Limitations

We attempted to decrease the methodological limitations of our cross-sectional empirical data, laboratory experiments, and agent-based simulations by using the methods in a complementary manner. Nevertheless, this article is not free of limitations. For example, we did not consider individuals’ acquiescence, that is, when a group member complies with a solution as a result of the influence of other members but does not identify with or internalize this choice. Moreover, individuals may differ in their willingness to accept the inequalities of power disparity. Hofstede (1986) raised this issue by arguing that power is a cultural phenomenon. Similarly, leaders can vary in their attitudes toward individuals with less power (Halevy et al., 2011). To sustain the parsimony of our theoretical model, we suppressed these aspects. Furthermore, in this article, we focused on intellecitive and problem-solving tasks (McGrath, 1984). Hence, future research can extend our findings to different settings, tasks, and actors, such as factory floor production teams.

Conclusion

This study enriches our understanding of power disparity and clarifies when functionalist versus conflict theories of power disparity are more likely to apply. We show that the nature of the effect of power disparity on group performance depends on the ability of groups and organizations to assign power according to competence in a dynamic manner. By identifying the implicit assumptions in power disparity research, we provide an important integration of different theoretical perspectives on power disparity in groups, and provide important managerial implications for how to best structure power in groups and organizations.

References


Anderson, C., & Kilduff, G. J. (2009). Why do dominant personalities tend to different settings, tasks, and actors, such as factory floor production teams.


Received December 15, 2014
Revision received August 26, 2015
Accepted September 1, 2015