

Journal of Personality and Social Psychology

Power Dynamics and the Reciprocation of Trust and Distrust

Marlon Mooijman

Online First Publication, April 10, 2023. <https://dx.doi.org/10.1037/pspi0000424>

CITATION

Mooijman, M. (2023, April 10). Power Dynamics and the Reciprocation of Trust and Distrust. *Journal of Personality and Social Psychology*. Advance online publication. <https://dx.doi.org/10.1037/pspi0000424>

Power Dynamics and the Reciprocation of Trust and Distrust

Marlon Mooijman

Management Department (Organizational Behavior), Jones Graduate School of Business, Rice University

Because trust is essential in relationships, scholars have sought to determine what causes people to trust each other. A burgeoning area of research on trust has focused on power dynamics. Yet, although successful trust development in relationships is a function of one individual initiating trust and another individual reciprocating this trust, research has focused exclusively on the impact of power on trust initiation and left unaddressed the impact of power on trust reciprocation. In the current research, I examine the impact of power dynamics on trust reciprocation—people trusting someone who first trusts them. Across five preregistered experiments, I demonstrate that people are more likely to trust high-power individuals than low-power individuals when these individuals first trust them. I also demonstrate that people are more likely to distrust high-power individuals than low-power individuals when these individuals first distrust them. Power dynamics amplify trust and distrust reciprocation because people believe that having power means making decisions based on dispositions rather than the situation. This belief makes people think that—compared to low-power individuals—high-power individuals view them as more trustworthy when trusting them, but more untrustworthy when distrusting them, which motivates reciprocation. Taken together, these findings show that power helps and hurts trust development and highlight when and why this occurs. This provides some clarity to a literature dominated by inconclusive findings on the relationship between power and trust, provides a blueprint for understanding the impact of power on longer term reciprocal trust development, and provides important practical implications for power holders.

Keywords: power, trust and distrust reciprocity, dispositional attribution, metaperceptions of trustworthiness

Supplemental materials: <https://doi.org/10.1037/pspi0000424.supp>

Trust, which is the willingness to be vulnerable to others (Kramer, 1999), is essential in well-functioning relationships. When people trust each other, they view each other positively (Ferrin et al., 2008), report high levels of relationship commitment and satisfaction (Wieselquist et al., 1999), and coordinate and cooperate effectively (Balliet & Van Lange, 2013). This makes trust crucial for the functioning of teams, organization, and societies. Teams perform better (Dirks, 2000), organizations are more profitable (Burke et al., 2007), and societies are more politically stable (Nunn et al., 2018) when people trust each other. It is no surprise, then, that a myriad of scholars in psychology (Weiss et al., 2021), management (Dirks & De Jong, 2022), sociology (Schilke et al., 2021), and economics (Fehr, 2009) have examined what causes people to trust each other.

A recent, and burgeoning, area of research on trust has focused on the role of power dynamics (Brion et al., 2019; du Plessis et al., 2023; Feenstra et al., 2020; Inesi et al., 2012; Mooijman et al., 2015, 2019;

Schilke et al., 2015; Weber et al., 2005). Although this area of research has demonstrated that power dynamics impact trust, it has focused exclusively on the impact of power on trust initiation and left unaddressed the impact of power on trust reciprocation. This is a significant omission. Successful trust development between two individuals involves one individual initiating trust and another individual reciprocating trust. When an individual initiates trust that is not reciprocated, trust development breaks down and relationships are more likely to develop distrust and conflict (Parco et al., 2002).

In the current research, I address this lacuna in the literature and examine the impact of power dynamics on trust reciprocation—people trusting someone who first trusts them. I propose that people are more likely to trust high- compared to low-power individuals when these individuals first trust them. I also examine the impact of power dynamics on the reciprocation of distrust, which is the unwillingness to be vulnerable to others (Kramer, 1999). I propose that people are also more likely to distrust high- compared to low-power individuals when these individuals first distrust them. I argue that these effects occur because people believe that having power means making decisions based on dispositions rather than the situation. People, consequently, view high- compared to low-power individuals' decision to trust or distrust as more indicative of their perception that people are trustworthy or not, which amplifies reciprocation. I find evidence for these predictions in five preregistered experiments.

The present work advances our understanding of the relationship between power dynamics and trust in several important ways. First, while prior work on power and trust has taken a unidirectional theoretical lens—focusing on one individual's trust in another

Marlon Mooijman  <https://orcid.org/0000-0002-3388-9120>

The author extends special thanks to Richard Schwartz for statistical help.

Marlon Mooijman played lead role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, visualization, writing—original draft and writing—review and editing.

The Open Science Framework page can be accessed at https://osf.io/5u7zj/?view_only=abe06395e0c742daa29599e6f9d1465e.

Correspondence concerning this article should be addressed to Marlon Mooijman, Management Department (Organizational Behavior), Jones Graduate School of Business, Rice University, 1900 Rice Boulevard, Houston, TX 77005, United States. Email: marlon.mooijman@rice.edu

individual—I take a *bidirectional* perspective, which involves an individual responding to being trusted or distrusted by another individual (Korsgaard et al., 2015). Doing so provides a more accurate, dynamic, and complete picture of the relationship between power and trust, as people do not merely initiate trust or distrust but also respond to it. In fact, the present research provides a blueprint for understanding the impact of power on longer term bidirectional trust development. High-power individuals' decision to initiate trust or distrust is more likely to set the tone for the rest of the exchange, leading high-power individuals to elicit reciprocal trust or distrust more easily from low-power individuals rather than vice versa. This trust and distrust dynamic is likely to spiral over longer periods of time (Ferrin et al., 2008).

Second, the present findings show *when* people are more likely to trust or distrust high- compared to low-power individuals—that is, when people are first trusted or distrusted by these individuals—and the psychological process by which this occurs. This provides a crucial context, and some theoretical clarity, to a literature that has shown that people are more likely to trust *and* distrust high-compared to low-power individuals (Ross et al., 2001; Schaerer et al., 2021; Schilke et al., 2015; Van der Toorn et al., 2011). Indeed, the current literature on power and trust is dominated by inconclusive and opposing findings, with scholars calling for work to “improve our understanding of the complex relationship between power and trust” (Schilke et al., 2016).

Finally, in demonstrating that power dynamics amplify trust and distrust reciprocation, the present work provides practical implications for high-power individuals interested in promoting trust. Instead of trying to promote trust and prevent distrust by flattening power hierarchies and reducing resource inequalities (Wilkinson & Pickett, 2007, 2009), the current research shows that power hierarchies and high levels of trust in power holders can coexist, as long as individuals with power trust others. For high-power individuals who want to be trusted, trusting others is particularly effective, whereas distrusting others is particularly destructive.

Trust, Distrust, Metaperceptions of Trustworthiness, and Reciprocity

Trust is typically defined as the act of accepting vulnerability to another person whereas distrust is typically defined as the act of not accepting vulnerability to another person (Kramer, 1999; see also Hardin, 2001; Zand, 1997). An individual trusts when giving someone an important letter to mail, money to invest, or a personal secret to keep based on the expectation that this person delivers the letter, invests the money, or keeps the secret. An individual distrusts when refusing to share mail; money; or secrets in fear of someone throwing the mail away, pocketing the money, or spilling the secret. Trust and distrust are therefore related to, but distinct from, perceived trustworthiness and untrustworthiness, which refers to the perception that another person will, or will not, take your interests into account (Mayer et al., 1995). For instance, in the Trust Game (TG), an individual trusts by giving money to another person based on the expectation that this person gives this money back at the end of the game. An individual distrusts in this setting by refusing to give money to another person based on the expectation that this person will not give money back at the end of the game (Berg et al., 1995).

In the current article, I examine both trust and distrust because they are two sides of the same coin—accepting or not accepting

vulnerability—and focus on the extent to which people trust or distrust in response to being trusted or distrusted. This can mean people sharing or refusing to share personal secrets with an individual who has shared or refused to share personal secrets with them. This can also mean people giving or refusing to give money to an individual who has given or refused to give money to them. Thus, when I refer to trust and distrust reciprocity, I refer to the reciprocation of trust with trust and distrust with distrust.

I distinguish between two stages of the trust and distrust reciprocity process. In the first stage, people are trusted or distrusted by someone else, which determines whether people think this person views them as trustworthy or untrustworthy. Trustworthiness and untrustworthiness, in this context, refer to *metaperceptions*—the perception that a person expects you to take, or not take, their interests into account (Mayer et al., 1995). For instance, in the Trust Game, people think that a person who trusts them, and gives them money, expects them to be trustworthy and return this money at the end of the game, whereas people think that a person who distrusts them, and does not give them money, expects them to be untrustworthy and not return any money at the end of the game (Pillutla et al., 2003). Being trusted or distrusted is therefore an antecedent of metaperceptions of (un)trustworthiness. People think that someone who trusts them sees them as trustworthy but that someone who distrusts them sees them as untrustworthy.

In the second stage, people base their decision to reciprocate trust or distrust on whether they think a person views them as trustworthy or not. People see themselves as trustworthy, want others to view them as trustworthy, and react negatively when treated with suspicion (Mooijman et al., 2017; Pillutla et al., 2003; Sherman & Cohen, 2006; Tyler et al., 2015). In fact, when no prior breach of rules has been displayed, people consider it fair to be seen as trustworthy but unfair to be seen as untrustworthy (Dunning et al., 2019) and base their trust in, and distrust toward, others on whether they feel fairly treated by these others (e.g., interpersonal justice perceptions; Tyler & Lind, 1992). Indeed, when I think that a person views me as trustworthy, I think they treat me appropriately and I am inclined to trust them. When I think that a person views me as untrustworthy, I think they treat me inappropriately and I am inclined to distrust them. Metaperceptions of trustworthiness and untrustworthiness are therefore antecedents of trust and distrust reciprocity. People reciprocate trust in part because of metaperceptions of trustworthiness, and people reciprocate distrust in part because of metaperceptions of untrustworthiness.

How do power dynamics shape this two-stage reciprocation process? I propose that people are more likely to reciprocate the trust and distrust from high- compared to low-power individuals because people think that—compared to low-power individuals—high-power individuals view them as more trustworthy when trusting them, but more untrustworthy when distrusting them. Thus, power dynamics may shape reciprocation through metaperceptions.

Power, Dispositions, Metaperceptions of Trustworthiness, and Reciprocity

I define power as asymmetric control over valuable resources (Magee & Galinsky, 2008) such that a high-power individual controls the resources that a low-power individual desires to obtain (Emerson, 1962). These resources can be monetary (e.g., salaries and bonuses), physical (e.g., desirable office spaces), or social (e.g., inclusion in

high-status teams; Fiske, 2010), and grant the high-power individual control over their own outcomes (independence; Lammers et al., 2016) and the outcomes of the low-power individual (social control; Fiske, 1993). Because of this, power liberates individuals from situational constraints. Instead of having to consider the perspectives (Galinsky et al., 2006), emotions (van Kleef et al., 2008), and attitudes (Galinsky et al., 2008) of others, power allows individuals to rely on their own preferences and desires. When making decisions, for instance, high-power individuals are more likely to rely on their own implicit attitudes (Guinote et al., 2012), authentic feelings (Hecht & LaFrance, 1998), value orientations (Galinsky et al., 2008), and personality traits (Overbeck & Droutman, 2013) than low-power individuals. In fact, high-power individuals are less likely to be affected by “the press” of the situation and less likely to conform to prevailing norms and conventions than low-power individuals (Keltner et al., 2003; Uskul et al., 2016). Thus, power allows individuals to behave more according to their inherent inclinations, preferences, and desires—their dispositions—rather than situational pressures.

Importantly, this well-documented link between power and dispositional decision-making is reflected in people’s beliefs about what it means to have power. People recognize that, and overestimate the extent to which, power allows individuals to make decisions based on dispositions. For instance, high-power individuals are seen as able to make decisions (e.g., to work overtime, have certain political opinions) because of their personality and idiosyncratic preference, whereas low-power individuals are seen as making these decisions because they must follow prevailing norms and appease others (Overbeck et al., 2006). These beliefs persist even when there is diagnostic information available about the actual situational pressures that decision-makers face. This suggests that the belief that power liberates individuals from situational pressures reflects not only reality but also a biased stereotype about how situationally unconstrained high-power individuals really are (Overbeck et al., 2006). Moreover, consistent with the notion that high-power individuals’ decisions are more likely to be perceived as reflecting their dispositions, Fragale et al. (2009) showed that individuals with well-connected, politically prominent families are seen as more intentional about engaging in alleged financial misconduct (e.g., improper tax return) than individuals without well-connected, politically prominent families (see also Bauman et al., 2016; Reeder et al., 2004). Taken together, people hold the belief that having power means making decisions based on inherent inclinations, preferences, and desires—dispositions—rather than situational pressures and constraints.

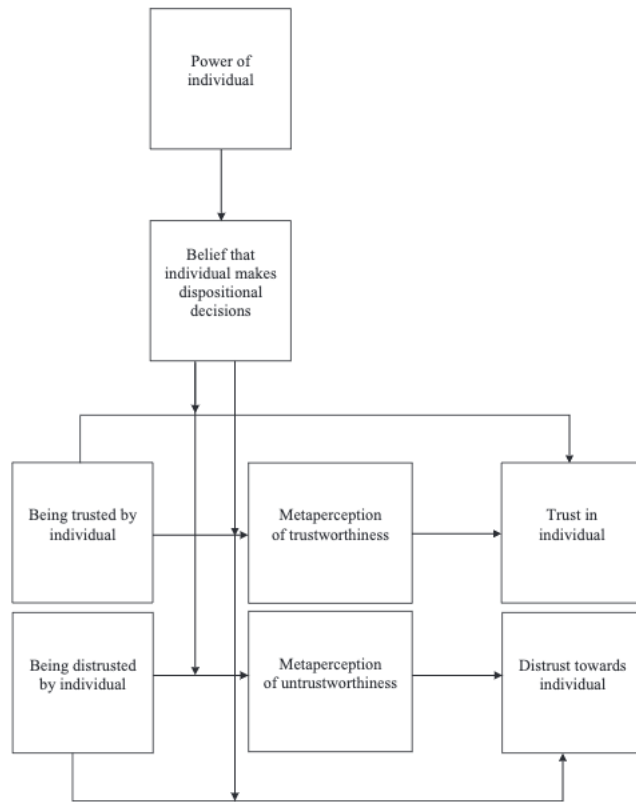
How does this belief shape the trust and distrust reciprocation process? I suggest that it makes people view high- compared to low-power individuals’ decision to trust or distrust as more strongly caused by the expectations they have about others’ trustworthiness. Indeed, someone who is seen as making dispositional decisions (e.g., based on their personality and idiosyncratic personal preference) is likely to be seen as basing the decision to trust or distrust on the internal factor that is most relevant to this decision: the expectation that others will be trustworthy or untrustworthy (Kelley, 1973; Weiner, 1992). In contrast, an individual who is seen as trusting or distrusting because of the situation—because they think accepting or refusing vulnerability is normative or because they are conforming to the demands of others—is less likely to be seen as basing this decision on their expectations about others’ trustworthiness.

Situational factors can conflict with an individual’s internal state, making it unclear whether an individual’s decision was driven by their expectations about others or not. Perhaps an individual does not think you are trustworthy but thinks that accepting vulnerability to your actions is situationally normative (Dunning et al., 2019); or an individual thinks you are trustworthy but is instructed by more powerful others to not accept vulnerability to your actions (Keltner et al., 2003). People may therefore think that—compared to low-power individuals—high-power individuals view them as more trustworthy when trusting them, but more untrustworthy when distrusting them. An individual’s power, in other words, makes being trusted or distrusted by them a stronger antecedent of people’s metaperceptions of trustworthiness.

Consequently, following the notion that people reciprocate trust and distrust because of their perception that someone views them as trustworthy or untrustworthy (e.g., Hanna et al., 2019; Mooijman et al., 2017; Pillutla et al., 2003; Tyler et al., 2015), people may be more likely to reciprocate the trust and distrust of high-power individuals compared to low-power individuals. Indeed, this power-reciprocation effect may be caused by people’s belief that high-power individuals’ decisions are caused by their dispositions. This leads people to think that high- compared to low-power individuals view them as more trustworthy when trusting them, but more untrustworthy when distrusting them. Taken together, I hypothesize that because people believe that power equals dispositional decision-making, an individual’s power makes being trusted or distrusted by this individual a stronger antecedent of people’s metaperceptions of trustworthiness. This, in turn, makes people more likely to reciprocate high- compared to low-power individuals’ trust and distrust. See Figure 1 for a schematic depiction of these hypotheses.

These hypotheses can be contrasted with predictions derived from other theoretical perspectives. Resource-dependence theories predict that people are motivated to think positively about high-power individuals because they are dependent on them for obtaining valuable resources. This perspective suggests that people trust high-power individuals *more* than low-power individuals regardless of whether these individuals trust or distrust them (Schilke et al., 2015; Van der Toorn et al., 2011; Weber et al., 2005). In contrast, encapsulated-interest theories (Farrell, 2004; Hardin, 2001) predict that people assume that high-power individuals have no incentive to encapsulate others’ interests into their behavior, making people *distrust* high-power individuals more than low-power individuals regardless of whether these individuals trust or distrust them. Last, stereotype-content models predict that people expect to be distrusted by high-power individuals (because they are seen as cold and selfish; Fiske et al., 2002; see also Overbeck et al., 2006), which should make people react more positively to being trusted but *not* more negatively to being distrusted by them. Indeed, a positive surprise should elicit a more positive response than the confirmation of a positive expectation, and a negative surprise should elicit a more negative response than the confirmation of a negative expectation (Burgoon, 1993). Following this logic, this perspective would predict that, in contrast to my prediction, people react *more* negatively to being distrusted by low-power individuals rather than high-power individuals because they expect to be trusted by low-power individuals but distrusted by high-power individuals. In sum, the hypotheses of the current article can be distinguished from predictions derived from other theoretical perspectives.

Figure 1
Schematic Depiction of Hypotheses



Note. All lines indicate positive causal effects.

Overview of Research, Transparency and Openness, and Statistical Power

I tested my hypotheses in five preregistered experiments. In Experiment 1, I tested whether employees were more likely to reciprocate the trust and distrust from their high- compared to low-power coworker because they thought that—compared to low-power coworkers—high-power coworkers viewed them as more trustworthy when trusting, but more untrustworthy when distrusting. In Experiment 2, I used a Trust Game, replicated the core findings of Experiment 1, and measured the extent to which participants believed that other, high-power TG players trusted or distrusted them based on dispositions or situational pressures. I also ruled out several alternative explanations (e.g., expectations about whether the other party will trust or distrust, risk perceptions, and desire to conform to other party's wishes). In Experiment 3, I manipulated rather than measured the degree to which participants believed that high-power individuals made trusting or distrusting decisions based on their dispositions. Similarly, in Experiment 4, I manipulated rather than measured metaperceptions of trustworthiness and untrustworthiness in a Trust Game setting. This allowed me to replicate the findings of Experiments 1 and 2 while providing causal rather than mediational evidence for the proposed psychological mechanisms. Finally, in Experiment 5, I tested the full theoretical model in a sample of coworker dyads and temporally separated the predictor variable from subsequent variables with a

time-lag design. This allowed me to test my hypotheses in a field setting and reduce common method bias (Podsakoff et al., 2003).

For each experiment, I report how I determined sample sizes, all manipulations, and all measures. Data, syntax, and research materials for all experiments are available on Open Science Framework (OSF). All experiments were preregistered and participants could only take part in the experiment after they passed an attention check and a captcha test. Institutional review board approval was obtained from the author's respective institution. All items in the studies could be answered on 7-point scales (1 = *completely disagree*; 7 = *completely agree*), unless stated otherwise, and all participants provided informed consent and were debriefed. No participants were excluded from any experiment for any post hoc reason. I followed journal article reporting standards (Kazak, 2018) and used G*Power3 to conduct a priori power analyses and determine the sample sizes needed for Experiments 1–4. In Experiment 5, I based the number of participants on the availability of coworker dyads and conducted a sensitivity analysis. I used relatively conservative estimates for effect size estimates (e.g., $\eta_p^2 = .01$) and ensured that each experiment had at least 90% statistical power. Where appropriate, I conducted principal component analyses (PCAs) on items and considered item-component loadings above 0.40 as sufficient. In total, I recruited 4,864 participants who were working professionals, coworker dyads, or online participants, and who were confronted with role-playing paradigms; autobiographical priming tasks; Trust Game settings; or real-life instances of trust, distrust, and power dynamics.

Please note that I report the main and interaction effects for the main dependent variable—participants' trust in, or distrust toward, an individual—in the result section of each respective experiment, but I report the main and interaction effects for the mediator variables dispositional attribution and metaperceptions of trustworthiness in tables, figures, or the supplementary online materials (referred to as online Supplemental Materials throughout the remainder of the article). I did this to reduce the size of the results section and increase its readability. I indicate in the results sections where the additional results can be found in the online Supplemental Materials. Results for the mediator variable of dispositional attribution closely followed my hypothesis: participants thought that high-power individuals made more dispositional trust and distrust decisions than low-power individuals. The results for the mediator variable of metaperceptions of trustworthiness also closely followed my hypothesis: participants thought that—compared to low-power individuals—high-power individuals viewed them as more trustworthy when trusting them, but more untrustworthy when distrusting them.

Last, I also report additional experiments in the online Supplemental Materials that are relevant for the current article but that did not include control conditions, used inferior measurements, did not manipulate power, or focused on mechanisms that do not fit with the framing of the current article. I ran these experiments before I ran the five preregistered experiments reported in the current article. For the sake of transparency, I briefly describe the results of these experiments in the online Supplemental Materials. Where relevant, these experiments replicate the findings reported below or hint at other promising directions for future research.

Experiment 1

In Experiment 1, I tested my core hypotheses on the relationships between power, metaperceptions, and reciprocity in a real-world

context. I recruited a sample of working professionals who had a coworker at the time of study and measured participants' perceptions of their coworker's power. I used an autobiographical recall method to manipulate whether it was salient to participants that their coworker trusted or distrusted them, measured participants' meta-perceptions of trustworthiness regarding this coworker, and measured their trust and distrust by measuring their willingness or unwillingness to be vulnerable to the coworker (e.g., admit mistakes they made in the workplace). I included a measure of liking to rule out that any observed effects simply reflected participants' like or dislike of their coworker. I also controlled for demographic characteristics such as gender, age, income, Social Economic Status, and education. I preregistered at https://aspredicted.org/blind.php?x=VY6_NDC.

Method

Participants and Design

I recruited through the community subject pool of a U.S. university 645 working professionals who had a coworker at the time of the study (410 males; $M = 28.45$, $SD = 7.87$) and randomly assigned them to three conditions: being trusted, control, and being distrusted. Participants held on average a college degree and were most likely to make between \$25k–\$50k a year. Participants rated their social class 5.70 ($SD = 1.37$) out of 10 on the MacArthur scale of SES. All participants indicated having a job and a direct coworker, with professions ranging from IT technicians, security guards, data analysts, to accountants, clerks, and waitresses. I conducted an a priori power analysis to determine the minimum sample size needed to find a stronger impact of the trust versus control and distrust versus control contrast on the dependent variable. I used G*Power3 with a δ between slopes of .30, error probability of .05, and power of .90. This analysis suggested a sample size of 215 per condition. I therefore decided to run 645 participants.

Procedure

Power

I measured the power of the coworker by presenting participants with six items taken from the perceived power scale developed by Yu et al. (2019). Sample items included, "My coworker has a great deal of power at work" and "My coworker's designated role allows them a lot of control over resources" ($\alpha = .70$; $M = 4.76$, $SD = 0.82$).

Being Trusted or Distrusted Manipulation

I adapted the autobiographical recall task from the literature on power and used it to prime whether it was salient to working professionals that their coworker trusted or distrusted them (Galinsky et al., 2003). The idea behind this manipulation is to make cognitively salient an event in participants' lives where their coworker accepted versus did not accept vulnerability to participants' future actions. This, in turn, should influence how participants perceive their coworker. Participants in the being trusted (distrusted) condition were instructed to:

Please recall and write down an event in your relationship with your coworker where they made [did not make] themselves vulnerable to your actions in the workplace. That is, please describe an event from

your own life where your coworker accepted [did not accept] personal vulnerability to your actions. This can relate to things your coworker said to you or did not say to you, the things your coworker shared or did not share with you, or the behaviors your coworker displayed towards you. Please write at least 3–5 sentences.

In the control condition, participants were instructed to "recall and write about the last time that you worked together with your co-worker. Please write at least 3–5 sentences."

Metaperceptions of Trustworthiness

I measured the degree to which participants perceived their coworker as having relatively positive or negative expectations about participants' trustworthiness with 17 items taken from Mayer and Davis (1999). These items distinguished between the benevolence, integrity, and ability dimensions of trustworthiness that prior research has demonstrated are relevant (Colquitt et al., 2007). Five items referred to the benevolence dimension of trustworthiness (e.g., "My coworker thinks that I take into account their needs and desires"), six items referred to the integrity dimension of trustworthiness (e.g., "My coworker thinks I have a strong sense of justice"), and five items referred to the ability dimension of trustworthiness (e.g., "My coworker thinks that I am capable of performing my job"). I combined all 17 items into one scale ($\alpha = .98$; $M = 4.39$, $SD = 1.69$). A PCA showed that all items loaded on one (the first) component (cf. Baer et al., 2021).

(Dis)Trust in Coworker

I measured the degree to which participants (dis)trusted their coworker using an adapted version of Mayer and Gavin's (2005) trust scale. Five items measured trust (e.g., "I would tell my coworker about any mistakes I made on the job, even if they could damage my reputation") and five items measured distrust (e.g., "I would be uncomfortable giving my coworker a task or problem that was critical to me without the ability to monitor their actions"). A PCA showed that all items loaded on one (the first) component with the distrust items having negative loadings. I reverse-coded the distrust items and averaged all 10 items to create a (dis)trust scale ($\alpha = .92$; $M = 4.03$, $SD = 1.11$). I interpret scores above the midpoint of the scale as relative trust and scores below the midpoint of the scale as distrust.

Liking of Coworker

I measured liking of coworker with three items (e.g., "My coworker is a likeable person"; $\alpha = .90$; $M = 5.17$, $SD = 1.35$).

Being Trusted and Distrusted Manipulation Check

I included the following item: "My coworker tends to accept vulnerability to my actions in the workplace"; $M = 4.16$, $SD = 1.43$).

Results

Being Trusted and Distrusted Manipulation Check

First, one independent coder who was blind to the hypotheses verified that all participants wrote a story related to the prompt. Second, confirming the manipulation, the being trusted or distrusted manipulation impacted participants' responses to the manipulation

check item, $\beta = .28$, $t(644) = 7.24$, $p < .001$, 95% CI [.35, .61], $rsp = .28$. I found no main effect of power nor an interaction effect, $ps > .29$. Participants in the being trusted condition indicated that their coworker tends to accept more vulnerability to them ($M = 4.68$, $SD = 1.45$) than participants in the control condition ($M = 4.09$, $SD = 1.36$), $t(429) = 4.33$, $p < .001$, Cohens' $d = 0.42$, and being distrusted condition ($M = 3.71$, $SD = 1.32$), $t(426) = 7.18$, $p < .001$, $d = 0.69$. Participants in the being distrusted condition indicated that their coworker tends to accept less vulnerability to them than participants in the control condition, $t(429) = -2.92$, $p = .002$, $d = 0.28$.

(Dis)Trust in Coworker

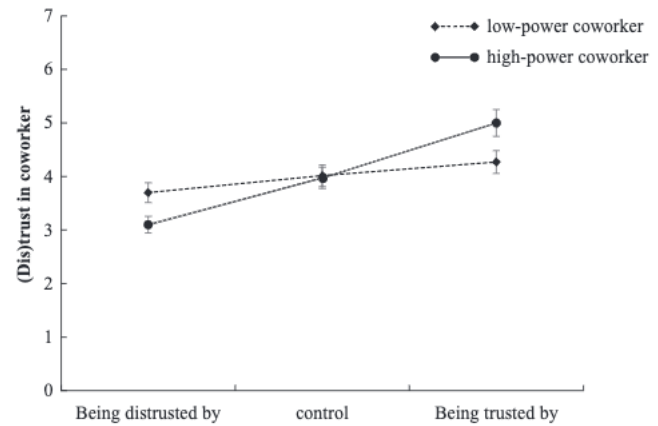
I conducted a multiple regression analysis that included the being trusted or distrusted manipulation (1 = *being trusted*, 0 = *control*, -1 = *being distrusted*), power (mean centered; Step 1), and their interaction (Step 2) as predictors. The (dis)trust scale was the dependent variable. This analysis yielded the predicted Being Trusted or Distrusted \times Power interaction effect, $\beta = .18$, $t(644) = 5.09$, $p < .001$, 95% CI [.17, .39], $rsp = .18$ (see Supplemental Table S2, for the main effects). To decompose this interaction effect, I ran two similar multiple regression analyses that either included as predictors the being trusted contrast (1 = *being trusted*, -1 = *control*) or the being distrusted contrast (1 = *being distrusted*, -1 = *control*) alongside power and their interaction term.

For the being trusted contrast, I found main effects of being trusted, $\beta = .34$, $t(430) = 7.59$, $p < .001$, 95% CI [.24, .41], $rsp = .23$, and power, $\beta = .12$, $t(430) = 2.66$, $p = .008$, 95% CI [.03, .20], $rsp = .12$, as well as the predicted interaction effect between being trusted and power, $\beta = .20$, $t(430) = 3.98$, $p < .001$, 95% CI [.10, .29], $rsp = .18$. Consistent with my hypothesis that power motivates trust reciprocity, being trusted predicted participants' trust in their coworker more strongly when this coworker had high power, +1 SD ; $\beta = .54$, $t(430) = 8.09$, $p < .001$, 95% CI [.39, .64], $rsp = .36$, compared to low power, -1 SD ; $\beta = .13$, $t(430) = 1.95$, $p = .052$, 95% CI [-.01, .26], $rsp = .09$. In fact, as is visible in Figure 2, the slopes for high- versus low-power coworkers were significantly different from each other, $t(426) = 4.23$, $p < .001$, and participants had more trust in high- compared to low-power coworkers when participants were first trusted by these coworkers, $\beta = .37$, $t(213) = 5.80$, $p < .001$, 95% CI [.27, .54], $rsp = .37$.

For the being distrusted contrast, I found a main effect of being distrusted, $\beta = -.25$, $t(430) = -5.29$, $p < .001$, 95% CI [-.38, -.18], $rsp = .25$, and the predicted interaction effect between being distrusted and power, $\beta = -.10$, $t(430) = -2.17$, $p = .030$, 95% CI [-.20, -.01], $rsp = .10$. I observed no main effect of power, $\beta = -.07$, $t(430) = -1.54$, $p = .13$, 95% CI [-.17, .02], $rsp = .07$. Consistent with my hypothesis that power motivates distrust reciprocity, being distrusted predicted participants' distrust toward their coworker more strongly when this coworker had high power, +1 SD ; $\beta = -.34$, $t(430) = -5.38$, $p < .001$, 95% CI [-.52, -.24], $rsp = .25$, compared to low power, -1 SD ; $\beta = -.16$, $t(430) = -2.52$, $p = .012$, 95% CI [-.31, -.04], $rsp = .12$. As is visible in Figure 2, the slopes for high- versus low-power coworkers were significantly different from each other, $t(426) = 2.02$, $p = .04$, and participants distrusted high-power coworkers more strongly than low-power coworkers when participants were first distrusted by them, $\beta = -.17$, $t(213) = 2.48$, $p = .014$, 95% CI [-.34, -.04], $rsp = .17$.

Figure 2

The Degree to Which Participants Trusted or Distrusted Their Coworker as a Function of Being Trusted or Distrusted and Coworker Power



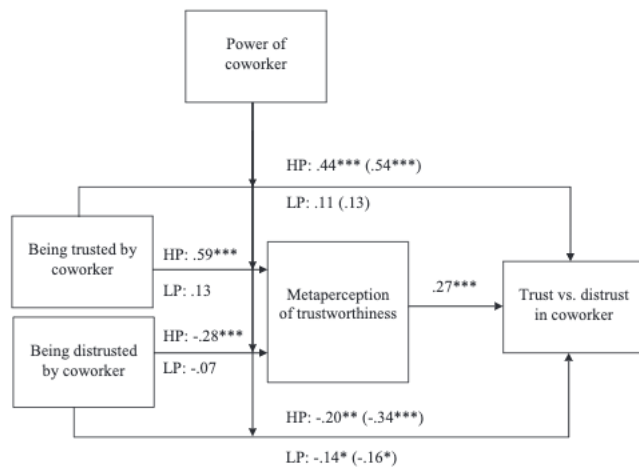
Note. Error bars represent SE. Experiment 1. SE = standard error.

Please note that all these effects remained the same in terms of direction and significance when adding age, gender, education, income, Social Economic Status, and liking of the coworker as covariates to the model (see Supplemental Table S3). See also online Supplemental Material: Experiment 1 for the regression results for liking.

Mediation Analysis

As part of my theoretical model, I hypothesized that changes in metaperceptions of trustworthiness would explain why people are more likely to reciprocate the trust and distrust of high- compared to low-power coworkers. To test this hypothesis, I ran two bootstrapping analyses (10,000 resamples) with either the being trusted contrast or the being distrusted contrast as independent variable. I added metaperceptions of trustworthiness as mediator, participants' (dis)trust in their coworker as dependent variable, and coworker power as moderator to both analyses. As expected, these analyses yielded two significant moderated mediation effects, $b = .06$, $SE = .02$, 95% [.03, .10]; $b = -.02$, $SE = .01$, 95% [-.05, -.01], respectively. I report the multiple regression results for metaperceptions in Figure 3 and Supplemental Table S2. As is visible in Figure 3, the effect of being trusted on participants' greater trust in high- compared to low-power coworkers was mediated by the perception that coworkers viewed them as more trustworthy when this coworker had high power ($b = .13$, $SE = .03$, 95% [.06, .20]) compared to low power ($b = .03$, $SE = .01$, 95% [.01, .07]). The effect of being distrusted on participants' greater distrust toward high- compared to low-power coworkers was mediated by the perception that coworkers viewed them as more untrustworthy when this coworker had high power ($b = -.04$, $SE = .02$, 95% [-.09, -.01]) compared to low power ($b = -.01$, $SE = .01$, 95% [-.04, .02]). These effects held after adding age, gender, education, income, Social Economic Status, and liking as covariates to the model (see Supplemental Table S4). The correlation table is in Supplemental Table S1.

Figure 3
Mediation Model for Experiment 1



Note. Coefficients within brackets are direct effects. HP = high power; LP = low power.

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

Discussion

Experiment 1 demonstrated that power dynamics shape trust and distrust reciprocation in a real-world context. Participants trusted their high-power coworkers more than their low-power coworkers when they thought that these coworkers trusted them, whereas participants distrusted their high-power coworkers more than their low-power coworkers when they thought that these coworkers distrusted them. These effects remained significant after controlling for demographics and liking and were explained by participants' perception that—compared to low-power coworkers—high-power coworkers viewed them as more trustworthy when trusting them, but more untrustworthy when distrusting them. In sum, Experiment 1 provides evidence for my core hypotheses on the relationships between power, metaperceptions, and reciprocity. Of course, some alternative explanations for these observed relationships remain (e.g., power shaping fear or risk taking; Anderson & Galinsky, 2006; Keltner et al., 2003) and Experiment 1 did not directly test the role of dispositional attribution. I address these issues in Experiment 2.

Experiment 2

In Experiment 2, I examined the hypothesized relationships between power, dispositional attribution, metaperceptions, and reciprocity in a Trust Game setting. This meant I used a different measure of trust and distrust—money received or not received, and money given or not given. I also controlled for more alternative explanations. Based on the notion that the observed effects could be accounted for by changes in risk perceptions (Anderson & Galinsky, 2006), fear of punishment (Keltner et al., 2003), felt pressure to conform to others (Case et al., 2015), or expectations about whether one will be trusted (Fiske et al., 2002), I statistically controlled for whether participants: (a) considered it risky to trust and give money, or distrust and withhold money; (b) were afraid of being punished in response to reciprocating, or not reciprocating, trust or distrust; (c) behaved the way they did because they were conforming to the

perceived wishes of the other party; and (d) expected to be trusted and given money, or distrusted and not given money. I contrasted a power condition with a control condition to demonstrate that the observed effects are not due to the target individual's low power (cf. Schaerer et al., 2018) and measured dispositional attribution. I preregistered at https://aspredicted.org/RP5_G8V.

Method

Participants and Design

I recruited 1,256 participants from the Cloud Research platform (812 males; $M = 34.00$, $SD = 11.13$) and randomly assigned them to a 3 (being trusted vs. control vs. being distrusted) \times 2 (power vs. control) between-participants design. I conducted an a priori power analysis to determine the sample size needed to find the hypothesized interaction effect. I used G*Power3 with an η_p^2 of .01, error probability of .05, power of .90, numerator df of 2, and number of groups of 6. This analysis suggested a sample size of 1,256.

Procedure

Power Manipulation

Following the power manipulation developed by Khademi et al. (2021), participants were informed that they are part of a six-person group. Participants in the power condition were told that one person in this group, who is not the participant, was randomly chosen and given five lottery tickets to distribute among the other five persons. This person has full control over how to distribute these tickets but cannot keep the tickets for themselves. This person makes their allocation at the end of the experiment. Participants in the control condition were not given this information. Please note that the lottery stood at \$135,000,000 at the time of the study.

Trust Game

Participants were then informed that they would play an exchange task with one person of the six-person group. Participants in the power condition were informed that the person they will play the task with controls the lottery tickets and decides on its distribution at the end of the experiment. Participants were told that

in this exchange task, the other player will be endowed with an additional \$1 and given the choice to give this \$1 to you or not. When the other player gives you \$1, it will triple in value. This means you will receive \$3. You can then decide to give any amount of this \$3 back to the other player or keep it yourself. When the other player gives you \$0, you will receive nothing, and the other player will keep the \$1.

Participants answered two verification questions to ensure they understood these instructions. I adapted these instructions from prior work (e.g., Berg et al., 1995; Lount & Pettit, 2012).

Dispositional Attribution

Following Overbeck et al. (2006), I measured dispositional attribution with six items adapted to the Trust Game context. Three items referred to dispositions (e.g., "The other player's behavior in this task will reflect who they are"), and three items referred to the situation (e.g., "The other player's behavior in this task will reflect

their feeling that the situation they are in dictates this behavior to them"). Because a PCA showed that all items loaded on one (the first) component, I reverse-coded the situational items and combined them with the dispositional items to construct the full scale ($\alpha = .87$; $M = 4.03$, $SD = 1.14$).

Being Trusted or Distrusted Manipulation

Participants in the being trusted condition were instructed that the other player had decided to give them \$1—their full endowment. This demonstrated to participants that the other player accepted vulnerability to their future actions. Participants in the being distrusted condition were instructed that the other player had decided to give them \$0. This demonstrated to participants that the other player did not accept vulnerability to their future actions. Participants in the control condition were not told whether the other player had decided to give them \$1 or \$0 and continued with the task.

Metaperceptions of Trustworthiness

Participant indicated on two items whether they thought the other player had positive or negative expectations about their trustworthiness ("The other player thinks that I would never keep money for myself"; $r = .67$; $M = 4.01$, $SD = 1.91$).

(Dis)Trust in the Other Player

Participants were then told that any remaining responses to this version of the exchange task were not needed and that they would play the same exchange task again with the same person.¹ This time, however, they would be the player endowed with \$1 who is able to decide whether to give this \$1 to the other player. The other player, this time, could give money back to the participant. Thus, participants switched roles between the first iteration of the game where the other player initiated trust or distrust and the second iteration of the game where participants initiated trust or distrust. Participants across conditions therefore faced the same choice—to trust or distrust the other player after being trusted, distrusted, or not knowing whether they were trusted or distrusted. All participants answered two verification questions to ensure they understood the instructions for this version of the game and indicated whether they would give their \$1 to the other player (0 = *I will give \$0 to the other player*; 1 = *I will give \$1 to the other player*).

I used this design rather than having participants return any money given to them by the other player in the first version of the task (as is customary in Trust Game designs; Berg et al., 1995) because participants in the being distrusted condition and control condition did not receive any money from the other player, making it impossible to compare across conditions how much money participants returned to the other player. I also chose this design because trust and distrust reciprocation is different from reciprocation as a form of actual trustworthiness. Although Trust Game scholars label participants' return of others' money as reciprocity, this behavior has nothing to do with participants' trust or distrust in these others. It is possible to give someone's money back to them out of a sense of obligation, but not give your own money to them out of a fear they will not give it back (Pillutla et al., 2003). The Trust Game designs I use, then, measure the reciprocation of trust with trust and distrust with distrust—the focal construct.

Control Variables

I measured participants' expectation that they would be trusted,² fear, conformity, and risk. That is, before the trust/distrust manipulation, participants were asked to indicate whether they expected to receive \$1 or \$0 (0 = *The other player will not give their \$1 to me*, 1 = *The other player will give their \$1 to me*). After the manipulation, participants indicated whether they were "afraid of negative repercussions from the other player if I did not act the way I did in this task" ($M = 2.54$, $SD = 1.27$) and whether they thought "the other player was afraid of receiving negative repercussions from me" ($M = 2.20$, $SD = 1.24$). Participants also indicated whether "I behaved the way I did in this task because that's how the other player wanted me to act" ($M = 3.91$, $SD = 1.62$), "giving money in this task is risky" ($M = 5.54$, $SD = 1.57$), and "not giving money in this task is risky" ($M = 2.58$, $SD = 1.46$).

Power Manipulation Check

Participants indicated how powerful the other player was ("The other player has power over me"; $M = 4.27$, $SD = 1.47$).

Being Trusted and Distrusted Manipulation Check

Participants were asked "how much money did the other player give you in the exchange task?" After this, participants were told the study had come to an end. They were paid the maximum amount possible, given a lottery ticket, and were debriefed.

Results

Power Manipulation Check

A univariate analysis of variance (ANOVA) with being trusted or distrusted and power as independent variables, and the perceived power of the other individual as dependent variable, yielded a main effect of power, $F(1, 1250) = 126.43$, $p < .001$, $\eta_p^2 = .09$. Confirming the manipulation, participants in the power condition rated their partner as more powerful ($M = 4.71$, $SD = 1.37$) than participants in the control condition ($M = 3.82$, $SD = 1.44$). I observed no main effect of being trusted or distrusted and no interaction effect, $p > .49$.

¹ Participants in the being trusted condition were told that at the very end of the experiment, they could decide whether to give the \$3 they have in their possession as the result of the other player's behavior (i.e., after the \$1 being given to them by the other player in the first round tripled in value) back to this other player. Please also note that Experiments 2 and 4 involved some level of deception, as participants were not "live playing" the TG. I made sure participants were paid the maximum possible and were fully debriefed about the need for deception.

² I measured whether participants expected to be trusted or distrusted by the other player before they knew whether they were actually trusted or distrusted by this player because participants may have viewed high-power players as less likely to trust, which may have shaped participants' subsequent response to being trusted. Thus, this measurement connects to one of the alternative explanations I describe in the theory section where the observed pattern of findings is due to an interaction between expectations about being trusted and actually being trusted. As noted in the online Supplemental Materials, I find little evidence in favor of this idea.

Being Trusted and Distrusted Manipulation Check

All participants correctly identified the amount of money (if any) they received in the first interaction of the exchange task.

(Dis)Trust in Player

A logistic regression analysis with being trusted or distrusted (1 = *being trusted*, 0 = *control*, -1 = *being distrusted*), power, and their interaction as predictors, and participants' decision to trust or distrust as dependent variable, yielded the predicted interaction effect between being trusted or distrusted and power, $B = .35$, $SE = 0.08$, $Wald = 19.86$, $p < .001$, $Exp(B) = 1.42$ (see Supplemental Table S3, for the main effects). To decompose this three-by-two interaction effect, I ran two similar regression analyses that either included as predictors the being trusted contrast (1 = *being trusted*, -1 = *control*) or the being distrusted contrast (1 = *being distrusted*, -1 = *control*) alongside power and their interaction.

For the being trusted contrast, I found main effects of being trusted, $B = .35$, $SE = 0.08$, $Wald = 19.48$, $p < .001$, $Exp(B) = 1.42$, and power, $B = .19$, $SE = 0.08$, $Wald = 6.11$, $p = .013$, $Exp(B) = 1.22$. I also found the predicted being Trusted \times Power interaction effect, $B = .19$, $SE = 0.08$, $Wald = 5.69$, $p = .017$, $Exp(B) = 1.21$. Consistent with my hypothesis that power motivates trust reciprocity, being trusted predicted participants' trust in the other player more strongly when this player had power, $B = .54$, $SE = 0.12$, $Wald = 20.44$, $p < .001$, $Exp(B) = 1.72$, compared to when they did not, $B = .16$, $SE = 0.11$, $Wald = 2.37$, $p = .12$, $Exp(B) = 1.18$. In fact, when participants first received \$1, more participants gave \$1 to the players who had power (84.7%) compared to players who did not have power (71.8%), $B = .39$, $SE = 0.12$, $Wald = 9.99$, $p = .002$, $Exp(B) = 1.48$ (see Table 1).

For the being distrusted contrast, I found main effects of being distrusted, $B = -.47$, $SE = 0.07$, $Wald = 43.63$, $p < .001$, $Exp(B) = 0.62$ and power, $B = -.15$, $SE = 0.07$, $Wald = 4.54$, $p = .033$, $Exp(B) = 0.86$. I also observed the predicted interaction effect between being distrusted and power, $B = -.16$, $SE = 0.07$, $Wald = 4.95$, $p = .026$, $Exp(B) = 0.85$. Consistent with my hypothesis that power motivates distrust reciprocity, being distrusted predicted participants' distrust toward the other player more strongly when this player had power, $B = -.63$, $SE = 0.10$, $Wald = 37.92$, $p < .001$, $Exp(B) = 0.53$, compared to when they did not, $B = -.31$, $SE = 0.10$, $Wald = 9.87$, $p = .002$, $Exp(B) = 0.73$. In fact, when participants first received \$0, more participants gave \$0 to players who had power (65.6%) compared to players who did not have power (50.5%), $B = -.31$, $SE = 0.10$, $Wald = 9.69$, $p = .002$, $Exp(B) = 0.73$.

Table 1

Participants (Pp) Trusting or Distrusting the Other Player as a Function of Being Trusted, Control, or Being Distrusted by This Other Player, and the Power of This Other Player

Trust and distrust reciprocity	Being trusted by		Control being		Distrusted by	
	PP	CP	PP	CP	PP	CP
Pp trust player	84.7%	71.8%	65.1%	64.8%	34.4%	49.5%
Pp distrust player	15.3%	28.2%	34.9%	35.2%	65.6%	50.5%

Note. PP = power player; CP = control player. Experiment 2: $N = 1,256$.

Please note that all the reported effects remained the same in terms of direction and significance when controlling for fear, conformity, expectations, and risk perceptions (see Supplemental Table S8, for the output). For the main and interaction effect of being trusted or distrusted and power on dispositional attribution and metaperceptions, see Supplemental Tables S6 and S7.

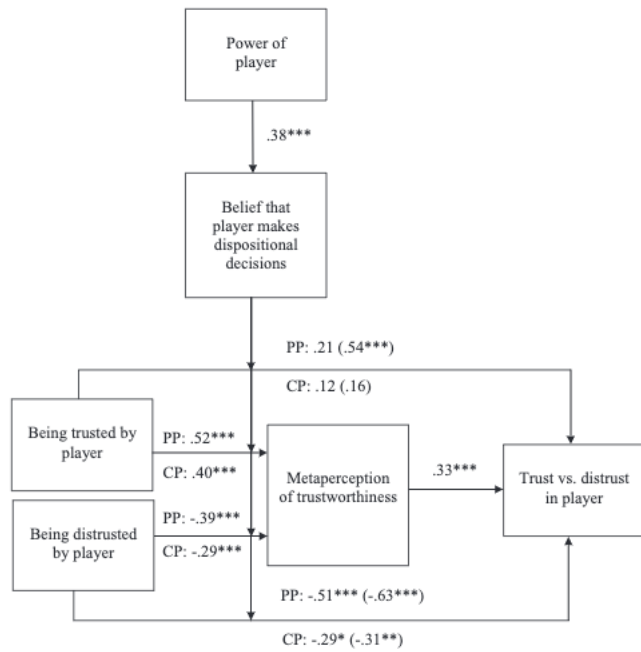
Mediation Analysis

The belief that power players are more likely to trust and distrust because of dispositions than control players should lead people to think that—compared to control players—power players view them as more trustworthy when trusting them, but more untrustworthy when distrusting them. This, in turn, should explain why people are more likely to reciprocate the trust and distrust of power players compared to control players. To test this, I ran a bootstrapping analysis (10,000 resamples) with the being trusted (coded as 1) versus control (coded as -1) contrast as independent variable, metaperceptions of trustworthiness as mediator, participants' trust in (giving \$1) or distrust toward (giving \$0) the other player as dependent variable, and the other player's power as moderating the relationship between being trusted and metaperceptions of trustworthiness. I also added an effect to this model where this moderating role of power occurred through dispositional attribution, per the theoretical model depicted at Figure 1. In addition, I ran the same analysis but now with the being distrusted (coded as 1) versus control (coded as -1) contrast as independent variable.

These two analyses yielded two significant moderated mediation effects for power ($b = .04$, $SE = .02$, 95% [.01, .09] for being trusted contrast; $b = -.03$, $SE = .02$, 95% [-.07, -.01] for being distrusted contrast). I report the multiple regression results for metaperceptions in Figure 4 and Supplemental Table S7. As is visible in Figure 4, metaperceptions predicted trust in (vs. distrust toward) the other player. The effect of being trusted on participants' greater trust in power players compared to control players was mediated by participants' perception that trusting players viewed them as more trustworthy when in power ($b = .13$, $SE = .03$, 95% [.06, .20]) compared to when not in power ($b = .03$, $SE = .01$, 95% [.01, .07]). Similarly, the effect of being distrusted on participants' greater distrust toward power players compared to control players was mediated by participants' perception that distrusting players viewed them as more untrustworthy when in power ($b = -.08$, $SE = .04$, 95% [-.16, -.01]) compared to when not in power ($b = -.05$, $SE = .02$, 95% [-.11, -.01]). This pattern of mediation findings replicates Experiment 1 and provides additional support for my hypotheses on the relationships between being trusted or distrusted, power, metaperceptions, and reciprocity.

Moreover, consistent with my predictions, power strengthened the indirect effect of metaperceptions of trustworthiness through the belief that power players' decision to trust or distrust reflected their dispositions ($b = .07$, $SE = .02$, 95% [.04, .12] for the being trusted contrast; $b = -.03$, $SE = .02$, 90% [-.07, -.01] for the being distrusted contrast). I report the multiple regression results for dispositional attribution in Supplemental Table S6. Indeed, as is visible in Figure 4, power predicted dispositional attribution. For the being trusted contrast, the indirect effect of metaperceptions of trustworthiness was stronger when this player was seen as making relatively dispositional decisions (+1 SD ; $b = .27$, $SE = .06$, 95% [.14, .40]) compared to making relatively nondispositional decisions

Figure 4
Mediation Model for Experiment 2



Note. Coefficients within brackets are direct effects. PP = power player; CP = control player.

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

(-1 SD; $b = .09$, $SE = .03$, 95% [.04, .15]). For the being distrusted contrast, the indirect effect of metaperceptions of untrustworthiness was stronger when this player was seen as making relatively dispositional decisions ($+1$ SD; $b = -.10$, $SE = .04$, 95% [-.19, -.02]) compared to making relatively nondispositional decisions (-1 SD; $b = -.02$, $SE = .01$, 95% [-.05, -.01]). This pattern of mediation findings provides support for the hypotheses that dispositional attribution and metaperceptions of trustworthiness explain why people are more likely to reciprocate the trust and distrust of high- versus low-power individuals.

Please note that all the reported indirect effects remained the same in terms of direction and significance when controlling for fear, conformity, expectations, and risk perceptions (see Supplemental Table S9, for the associated output). For the correlation table, see Supplemental Table S5, and for the main effects and interaction effects on all control variables, see online Supplemental Materials: Experiment S2.

Discussion

Experiment 2 demonstrated the hypothesized relationships between power, dispositional attribution, metaperceptions, and reciprocity in a Trust Game setting. A Trust Game player's power made it more likely that participants reciprocated this player's decision to trust or distrust. This effect was explained by participants' belief that power players were more likely to make decisions because of dispositions, which made participants think that—compared to a control player—a power player viewed them as more trustworthy when trusting them,

but more untrustworthy when distrusting them. These effects held when controlling for fear, the desire to conform, expectations about being trusted or distrusted, and risk perceptions. Taken together, Experiments 1 and 2 provide converging evidence for my hypotheses across two different samples, research designs, and measurements. Yet, in demonstrating the psychological processes underlying trust and distrust reciprocity in power contexts, I so far relied solely on mediational evidence. Although mediation analyses have their merits, they are unable to demonstrate the causal nature of hypothesized relationships (Fiedler et al., 2018). In Experiment 3, I therefore aimed to establish the causal role of dispositional attribution using a role-playing paradigm from the power literature (e.g., Blader & Chen, 2012).

Experiment 3

In Experiment 3, I manipulated power, being trusted or distrusted, and dispositional attribution. I used a role-playing paradigm (Mooijman et al., 2019, 2020) and manipulated whether I described a target individual as trusting or distrusting participants because of the target individual's dispositions or not. I expected that describing high-power and low-power individuals as trusting or distrusting because of their dispositions would attenuate the extent to which people reciprocated the trust and distrust of high-power individuals more strongly than the trust and distrust of low-power individuals. Indeed, when high-power and low-power individuals are seen as equally as motivated by their dispositions, then their decision to trust or distrust should be perceived as equally as diagnostic of their trustworthiness perceptions. This means that high-power individuals and low-power individuals should also be trusted or distrusted equally. I preregistered at https://aspredicted.org/TDC_KFR.

Method

Participants and Design

I recruited 1,256 participants from the Cloud Research platform (647 males; $M = 41.99$, $SD = 12.83$) and randomly assigned them to a 3 (being trusted vs. control vs. being distrusted) \times 2 (high power vs. low power) \times 2 (dispositional attribution vs. control) between-participants design. I conducted an a priori power analysis to determine the sample size needed to find the hypothesized interaction effect. I used G*Power3 with an η_p^2 of .01, error probability of .05, power of .90, numerator df of 2, and number of groups of 12. This analysis suggested a sample size of 1,256.

Procedure

Power Manipulation

Following prior power manipulations (e.g., Blader & Chen, 2012; Mooijman et al., 2019, 2020), participants in the high-power (low-power) condition were told to imagine

an individual in a firm who is one of the most [least] powerful individuals within the company. This individual has control over many [few] organizational resources, frequently evaluates others [is frequently evaluated by others] in the organization, and has others directly report to them [directly reports to others] in the firm.

Being Trusted or Distrusted Manipulation

All participants were then told that “this individual works on a project that is important for their career and that involves various stakeholders throughout the company. You also work on this project.” Following prior work (Mayer & Gavin, 2005), participants in the being trusted (being distrusted) condition were then told that “This individual is comfortable [uncomfortable] with you having an influence on this project and rarely monitors you or checks your work [always tries to monitor you and check your work].” This demonstrated to participants that this individual accepted or did not accept vulnerability to participants’ actions, consistent with the current operationalization of being trusted or being distrusted.³ Participants in the control condition were not given this additional information.

Dispositional Attribution Manipulation

Participants in the dispositional attribution condition were then instructed to

imagine that this individual’s behavior in the workplace is primarily driven by their dispositions. This means that their behavior is primarily due to their personality, because they want to behave that way, and/or because they have a personal preference to act that way.

Participants in the control condition were not given this information.

Dispositional Attributions Manipulation Check

I then measured the degree to which participants attributed dispositions to the individual with six items (e.g., “This individual’s behavior is caused by their personal preference to act that way”). A PCA showed that all items loaded on one (the first) component. I therefore used the full scale with the three situational items reverse-coded ($\alpha = .77$; $M = 4.82$, $SD = 1.11$).

Metaperceptions of Trustworthiness

I measured participants’ metaperceptions using the same items as in Experiment 1, this time referring to the “individual” rather than “coworker.” A PCA showed that all items loaded on one (the first) component. As such, I combined all seventeen items into one scale ($\alpha = .98$; $M = 4.40$, $SD = 1.55$).

(Dis)Trust in Individual

I measured the degree to which participants trusted or distrusted the individual using the same items as in Experiment 1, this time referring to the “individual” rather than “coworker.” A PCA showed that all items loaded on one (the first) component with the five distrust items having negative loadings. As such, I reverse-coded the distrust items and averaged all ten items to create the (dis)trust scale ($\alpha = .89$; $M = 3.89$, $SD = 1.38$).

Power Manipulation Check

Participants indicated how powerful the individual described in the organizational scenario was on one item (“This individual has a lot of power in the organization”; $M = 4.40$, $SD = 2.21$).

Being Trusted and Distrusted Manipulation Check

The being trusted or distrusted manipulation check also involved one question (“this individual accepts vulnerability to my actions in the workplace”; $M = 4.28$, $SD = 2.23$).

Results

Power Manipulation Check

A univariate ANOVA with power, being trusted or distrusted (1 = *being trusted*, 0 = *control*, -1 = *being distrusted*), and dispositional attribution as independent variables, and the perceived power of the individual as dependent variable, yielded only a main effect of power, $F(1, 1244) = 737.89$, $p < .001$, $\eta_p^2 = .37$. Confirming the power manipulation, participants in the high-power condition rated the individual as more powerful ($M = 5.75$, $SD = 1.37$) than participants in the low-power condition ($M = 3.05$, $SD = 2.07$). I observed no other main or interaction effects, $ps > .11$.

Being Trusted and Distrusted Manipulation Check

A univariate ANOVA with power, being trusted or distrusted (1 = *being trusted*, 0 = *control*, -1 = *being distrusted*), and dispositional attribution as independent variables, and the being trusted or distrusted manipulation check item as dependent variable, yielded only a main effect of the being trusted or distrusted manipulation, $F(1, 1244) = 643.82$, $p < .001$, $\eta_p^2 = .51$. Confirming the validity of the manipulation, participants in the being trusted condition scored higher ($M = 6.19$, $SD = 1.00$) than participants in the control condition, $M = 4.31$, $SD = 1.49$; $t(834) = 21.45$, $p < .001$, $d = 1.49$, and being distrusted condition ($M = 2.31$, $SD = 2.04$), $t(842) = 35.13$, $p < .001$, $d = 2.42$. Participants in the being distrusted condition scored lower than participants in the control condition, $t(830) = -16.13$, $p < .001$, $d = 1.12$. I observed no other main or interaction effects, $ps > .12$.

Dispositional Attribution Manipulation Check and Replication of Experiment 2

A univariate ANOVA with power, being trusted or distrusted (1 = *being trusted*, 0 = *control*, -1 = *being distrusted*), and the dispositional attribution manipulation as independent variables, and the dispositional attribution scale as dependent variable, yielded a main effect of power, $F(1, 1244) = 43.06$, $p < .001$, $\eta_p^2 = .03$, and a main effect of the dispositional attribution manipulation, $F(1, 1244) = 295.72$, $p < .001$, $\eta_p^2 = .19$. Participants, on average, attributed more dispositions to the behavior of high-power individuals ($M = 4.99$, $SD = 1.07$) than low-power individuals ($M = 4.64$, $SD = 1.12$). Participants also, on average, attributed more dispositions to the behavior of individuals after the dispositional attribution manipulation

³ To ensure the validity of this manipulation, I followed the procedures outlined by Hinkin and Tracey (1999) and asked 50 additional Mturk participants to rate the correspondence of these manipulations to the definition of trust as the decision to accept vulnerability to the actions of another person and the definition of distrust as not accepting vulnerability to the actions of another person (1 = *item is an extremely bad match to the definition*, 7 = *item is an extremely good match to the definition*). Results showed high levels of definitional correspondence for the being trusted manipulation ($M = 6.11$, $SD = 1.01$) and being distrusted manipulation ($M = 6.21$, $SD = 1.21$).

($M = 5.29$, $SD = 0.89$) compared to control attribution condition ($M = 4.35$, $SD = 1.12$).

In addition, the analysis also yielded an interaction effect between power and the dispositional attribution manipulation, $F(1, 1244) = 18.88$, $p < .001$, $\eta_p^2 = .02$. Replicating Experiment 2, participants in the control condition (where nothing was mentioned about dispositions) attributed more dispositions to the behavior of high-power individuals ($M = 4.65$, $SD = 1.15$) than low-power individuals ($M = 4.04$, $SD = 1.01$), $t(627) = 6.99$, $p < .001$, $d = 0.56$. Participants in the dispositional attribution condition, however, showed no significant difference in attributing dispositions to the behavior of high-power individuals ($M = 5.35$, $SD = 0.84$) versus low-power individuals ($M = 5.24$, $SD = 0.89$), $t(625) = 1.59$, $p = .11$, $d = 0.13$. This demonstrates that the dispositional attribution manipulation worked as intended. When nothing was mentioned about dispositions, participants thought that high- compared to low-power individuals' decision to trust or distrust was more motivated by their dispositions. When high-power and low-power individuals' decision to trust or distrust was described as equally motivated by their dispositions, this effect disappeared.

(Dis)Trust in Individual

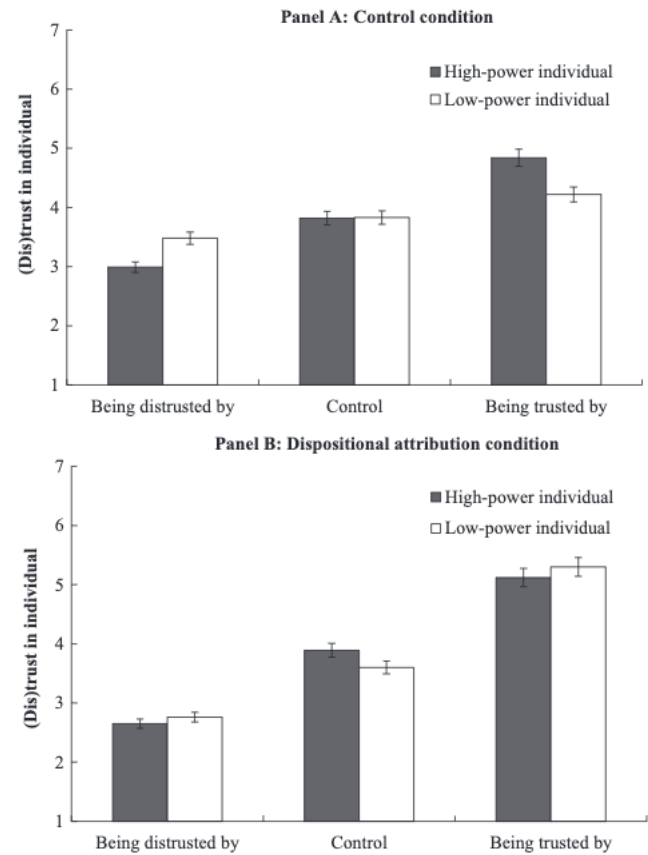
A univariate ANOVA with being trusted or distrusted (1 = *being trusted*, 0 = *control*, -1 = *being distrusted*), power, and the dispositional attribution manipulation as independent variables, and (dis)trust in the individual as dependent variable, yielded the predicted three-way interaction, $F(2, 1244) = 10.55$, $p < .001$, $\eta_p^2 = .02$ (for the main and other interaction effects, see Supplemental Table S11). Decomposing this three-way interaction effect for the dispositional attribution condition and control condition separately, I observed a two-way interaction effect between being trusted or distrusted (1 = *being trusted*, 0 = *control*, -1 = *being distrusted*) and power in the control condition, $F(2, 623) = 13.74$, $p < .001$, $\eta_p^2 = .04$, but not in the dispositional attribution condition, $F(2, 621) = 1.65$, $p = .19$, $\eta_p^2 = .00$.

Control Condition. For the control condition, I decomposed the significant three-by-two interaction effect by running two ANOVA analyses that either included the being trusted contrast (1 = *being trusted*, -1 = *control*) or the being distrusted contrast (1 = *being distrusted*, -1 = *control*), power, and their interaction as predictors. For the being trusted contrast, I found main effects of power, $F(1, 412) = 7.92$, $p = .005$, $\eta_p^2 = .02$, and being trusted, $F(1, 412) = 54.35$, $p < .001$, $\eta_p^2 = .12$. As is visible in panel A of Figure 5, the being trusted manipulation interacted with power, $F(1, 412) = 11.49$, $p < .001$, $\eta_p^2 = .03$. Consistent with my hypothesis that power motivates trust reciprocity, being trusted had a stronger impact on participants' trust in high-power individuals, $t(209) = 7.54$, $p < .001$, $d = 1.04$, compared to low-power individuals, $t(203) = 2.85$, $p = .005$, $d = 0.40$. Compared to low-power individuals ($M = 4.22$, $SD = 0.99$), participants trusted high-power individuals more when these individuals first trusted them ($M = 4.84$, $SD = 0.99$), $t(207) = 4.58$, $p < .001$, $d = 0.63$.

For the being distrusted contrast, I found main effects of power, $F(1, 416) = 5.93$, $p = .015$, $\eta_p^2 = .01$, and being distrusted, $F(1, 416) = 23.97$, $p < .001$, $\eta_p^2 = .05$. As is visible in panel A of Figure 5, the being distrusted contrast also interacted with power, $F(1, 416) = 3.69$, $p = .055$, $\eta_p^2 = .01$. Consistent with my hypothesis that power motivates distrust reciprocity, being distrusted had a stronger impact

Figure 5

The Degree to Which Participants Trusted the Individual as a Function of Dispositional Attribution, Power, and Being Trusted or Distrusted



Note. Error bars represent SE. Experiment 3. SE = standard error.

on participants' distrust toward high-power individuals, $t(208) = -5.38$, $p < .001$, $d = 0.75$, compared to low-power individuals, $t(208) = -1.92$, $p = .056$, $d = 0.27$. Compared to low-power individuals ($M = 3.49$, $SD = 1.45$), participants distrusted high-power individuals more when these individuals first distrusted them ($M = 2.99$, $SD = 1.45$), $t(211) = -2.91$, $p = .004$, $d = 0.40$. These findings replicate Experiments 1 and 2.

Dispositional Attribution Condition. Participants in the dispositional attribution condition did not differ in their trust in, versus distrust toward, high-power individuals and low-power individuals that first trusted ($M = 5.12$, $SD = 0.99$ vs. $M = 5.30$, $SD = 0.95$; $p = .23$) or distrusted them ($M = 2.77$, $SD = 1.06$ vs. $M = 2.66$, $SD = 1.32$; $p = .49$). See panel B of Figure 5.

In sum, when nothing was mentioned about whether individuals made dispositional decisions, power motivated trust and distrust reciprocity (cf. Experiments 1 and 2). However, when participants thought that both high-power and low-power individuals made dispositional decisions, power no longer motivated trust and distrust reciprocity. In fact, participants reciprocated low-power individuals' decision to trust or distrust just as strongly as high-power individuals' decision to trust or distrust when this decision was seen as motivated equally by dispositions. This suggests that dispositional attribution

plays a causal role in explaining why participants reciprocate the trust and distrust of high- rather than low-power individuals.

Mediation Analyses

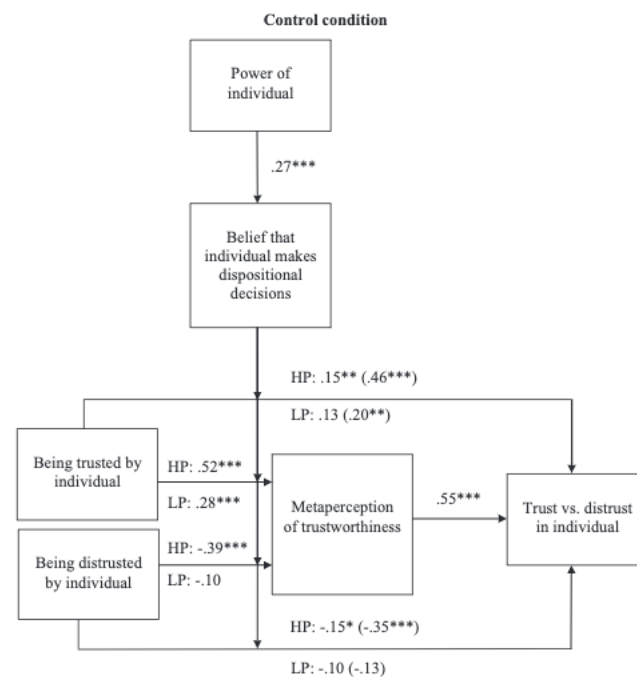
Control Condition. In the control condition, the belief that high-power individuals are more likely to make dispositional decisions should make people think that—compared to low-power individuals—high-power individuals view them as more trustworthy when trusting them, but more untrustworthy when distrusting them. This, in turn, should explain why people are more likely to reciprocate the trust and distrust of high- rather than low-power individuals. To test this, I ran a bootstrapping analysis (10,000 resamples) with the being trusted (coded as 1) versus control (coded as -1) contrast as independent variable, metaperceptions as mediator, participants' (dis)trust in the individual as dependent variable, and the individual's power as moderating the relationship between being trusted and metaperceptions of trustworthiness. I added an effect to this model where the effect of power occurred through dispositional attribution. I also ran the same analysis but now with the being distrusted (coded as 1) versus control (coded as -1) contrast as independent variable.

These two analyses yielded two significant moderated mediation effects for power ($b = .09$, $SE = .05$, 95% [.01, .19] for being trusted; $b = -.15$, $SE = .06$, 95% [-.27, -.05] for being distrusted). For the main effects and interaction effects on metaperceptions, see Supplemental Table S12. As is visible in Figure 6, metaperceptions predicted trust in (vs. distrust toward) the individual. The effect of being trusted on participants' greater trust in high- compared to low-power individuals was mediated by participants' perception that high-power individuals that trusted them viewed them as more trustworthy ($b = .13$, $SE = .03$, 95% [.06, .20]) than low-power individuals that trusted them ($b = .03$, $SE = .01$, 95% [.01, .07]). Similarly, the effect of being distrusted on participants' greater distrust toward high- compared to low-power individuals was mediated by participants' perception that high-power individuals that distrusted them viewed them as more untrustworthy ($b = -.12$, $SE = .03$, 95% [-.16, .06]) than low-power individuals that distrusted them ($b = .01$, $SE = .01$, 95% [-.01, .03]).

Moreover, power strengthened the indirect effect of metaperceptions of trustworthiness through participants' belief that high-power individuals' decision to trust or distrust reflected their dispositions rather than situational pressures ($b = .10$, $SE = .03$, 95% [.01, .21] for the being trusted contrast; $b = -.05$, $SE = .02$, 95% [-.11, -.01] for the being distrusted contrast). As is visible in Figure 6, for the being trusted contrast, the indirect effect of metaperceptions was stronger when the individual was seen as making relatively dispositional decisions (+1 SD ; $b = .29$, $SE = .05$, 95% [.19, .40]) compared to relatively nondispositional decisions (-1 SD ; $b = .07$, $SE = .03$, 95% [.02, .14]). For the being distrusted contrast, the indirect effect of metaperceptions was stronger when this individual was seen as making relatively dispositional decisions (+1 SD ; $b = -.22$, $SE = .05$, 95% [-.32, -.22]) compared to relatively nondispositional decisions (-1 SD ; $b = -.02$, $SE = .03$, 95% [-.09, .04]). This pattern of mediation directly replicates Experiment 2 (for the correlation table, see Supplemental Table S10).

Dispositional Attribution Condition. I then ran the same two mediation analyses as in Experiment 2 but now in the dispositional attribution condition. These two analyses did not yield moderated

Figure 6
Mediation Model for Experiment 3, Control Condition



Note. Coefficients within brackets are direct effects. HP = high power; LP = low power.

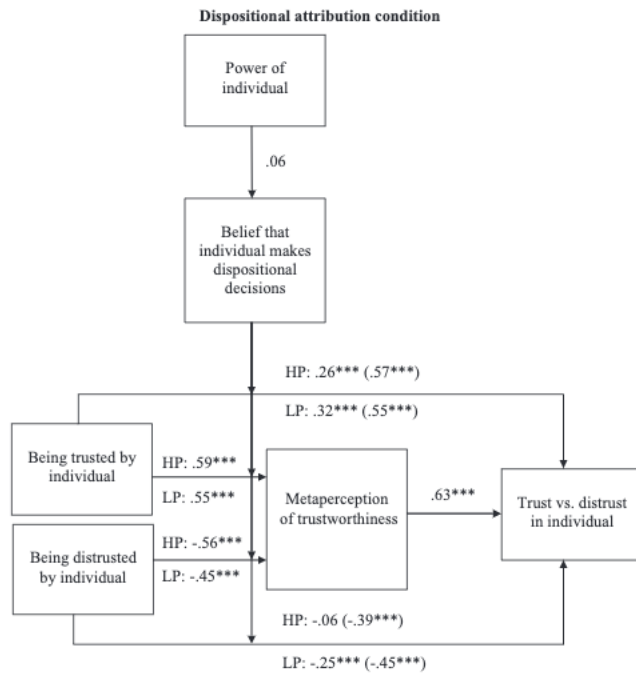
* $p < .05$. ** $p < .01$. *** $p < .001$.

mediation effects for power and no effect of power through dispositional attribution (all $b < .02$, $SE < .02$, 95% [-.05, .09]). As is visible in Figure 7, after the dispositional attribution manipulation, being trusted or distrusted by the individual predicted participants' metaperceptions of trustworthiness to a similar degree when it came to high-power individuals and low-power individuals. Thus, the dispositional attribution manipulation prevented participants from thinking that a high-power individual viewed them as more trustworthy when trusting them, but more untrustworthy when distrusting them. This, in turn, prevented the individual's power from shaping trust and distrust reciprocity.

Discussion

Experiment 3 replicated Experiment 2 in a role-playing paradigm and provided causal evidence for the role of dispositional attribution. Describing high- and low-power individuals as making trust and distrust decisions because of their dispositions prevented participants from thinking that—compared to a low-power individual—a high-power individual viewed them as more trustworthy when trusting them, but more untrustworthy when distrusting them. This, in turn, prevented power from shaping trust and distrust reciprocity. This provides direct causal evidence for the role of dispositional attribution in shaping metaperceptions of trustworthiness and, subsequently, reciprocity. Of course, Experiment 3's limitation was its hypothetical scenario method, which is why I again took a more "in person" approach in Experiment 4. Specifically, in Experiment 4, I continued manipulating rather than measuring the psychological process. This time, I manipulated metaperceptions in a Trust Game setting.

Figure 7
Mediation Model for Experiment 3, Dispositional Attribution Condition



Note. Coefficients within brackets are direct effects. HP = high power; LP = low power.

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

Experiment 4

So far, the hypothesized causal relationship between metaperceptions and trust in, or distrust toward, another individual has only been measured rather than manipulated. Experiment 4 addressed this shortcoming by manipulating metaperceptions of trustworthiness in a Trust Game setting. I used a similar Trust Game design as in Experiment 2 but added a condition where being trusted was coupled with a message that the other player considered the participant trustworthy. I also added a condition where being distrusted was coupled with a message that the other player considered the participant untrustworthy. I expected this manipulation to attenuate the extent to which people reciprocated trust and distrust more strongly for high- rather than low-power individuals. As such, I did not focus on, or measure, dispositional attribution. Rather, I aimed to demonstrate that power no longer shapes reciprocity when participants think that high- and low-power individuals view them as equally as (un)trustworthy. This highlights the causal nature of the relationship between metaperceptions and participants' trust in, or distrust toward, the other Trust Game player. I preregistered at https://aspredicted.org/4Q8_38H.

Method

Participants and Design

I recruited 1,407 Prolific participants (709 females; $M = 41.04$, $SD = 13.68$) and randomly assigned them to a 4 (being trusted vs.

being distrusted vs. being trusted + metaperception of trustworthiness vs. being distrusted + metaperception of untrustworthiness) \times 2 (power vs. control) between-participants design. I based the number of participants on a G*Power3 analysis with an η_p^2 of .01,⁴ error probability of .05, power of .90, numerator df of 3, and number of groups of 8. This analysis suggested to run 1,407 participants.

Procedure

Power Manipulation

I used the same power manipulation as in Experiment 2.

Trust Game

I used the same instructions as in Experiment 2.

Being Trusted or Distrusted Manipulation

Participants in the being trusted condition were told that the other player had sent them a message. In the being trusted condition, they received a message from the other player through the Chatplat platform that stated: "hey! I decided to give you \$1." Participants in the being distrusted condition also received a message from the other player that stated: "hey! I decided to give you \$0."

Metaperceptions of Trustworthiness Manipulation

Participants in the metaperceptions conditions were also told that the other player had sent them a message. In the being trusted + metaperceptions of trustworthiness condition, they received a message from the other player through the Chatplat platform that stated: "hey! I decided to give you \$1 because I think you will be trustworthy." Participants in the being distrusted + metaperceptions of untrustworthiness condition received a message from the other player that stated: "hey! I decided to give you \$0 because I think you will be untrustworthy."

(Dis)Trust in the Other Player

I used the same design as in Experiment 2.

Power Manipulation Check

I used the same item as in Experiment 2 ($M = 4.27$, $SD = 1.45$).

Metaperceptions of Trustworthiness Manipulation Check

I used the same items as in Experiment 2 ($r = .60$; $M = 4.09$, $SD = 1.79$).

Being Trusted Versus Distrusted Manipulation Check

Participants were also asked "how much money did your partner give you in the money exchange task?" After this, participants were debriefed about the nature of the study, paid the maximum possible, and given a lottery ticket.

⁴ The preregistration states that the effect size I used was .04 but it was .01 consistent with the prior experiments (hence the 1,407 participants).

Results

Power Manipulation Check

A univariate ANOVA with power, the four being trusted or distrusted conditions as independent variables, and perceived power as dependent variable yielded only a main effect of power, $F(1, 1396) = 354.61, p < .001, \eta_p^2 = .20$. Participants rated the power player as more powerful ($M = 4.92, SD = 1.22$) than the control player ($M = 3.62, SD = 1.36$). I observed no other main effect or interaction effect, $p > .24$. This confirms the validity of the manipulation.

Being Trusted Versus Distrusted Manipulation Check

All participants correctly identified the amount of money they had received.

Metaperceptions of Trustworthiness Manipulation Check and Replication of Experiments 1–3

A univariate ANOVA with the being trusted versus distrusted conditions and power as independent variables, and metaperceptions as dependent variable, yielded an interaction effect between the being trusted or distrusted conditions and power, $F(3, 1399) = 15.34, p < .001, \eta_p^2 = .03$ (for the main effects, see Supplemental Table S13). I decomposed this four-by-two interaction effect by examining the interaction effect between being trusted versus distrusted when the metaperceptions of trustworthiness manipulation was present or absent.

Absence of Metaperceptions Manipulation. I observed a two-way interaction effect between being trusted versus distrusted and power when the metaperceptions manipulation was absent, $F(1, 700) = 42.08, p < .001, \eta_p^2 = .06$. I also found a main effect of being trusted versus distrusted, $F(1, 700) = 1001.63, p < .001, \eta_p^2 = .06$. I found no main effect of power, $p = .99$. When participants did not receive a message from the other player about their trustworthiness, being trusted versus distrusted had a larger impact on participants' perception that the other player viewed them as trustworthy when this player had power, $t(350) = 28.86, p < .001, d = 3.08$, compared to when they did not, $t(350) = 16.76, p < .001, d = 1.78$. Compared to control players who trusted them ($M = 5.16, SD = 1.11$), participants thought that power players who trusted saw them as more trustworthy ($M = 5.71, SD = 0.88$), $t(351) = 5.06, p < .001, d = 0.54$. Participants thought that power players who distrusted saw them as more untrustworthy ($M = 2.52, SD = 1.17$) than control players who distrusted them ($M = 3.06, SD = 1.23$), $t(349) = -4.22, p < .001, d = 0.45$. Thus, when nothing was mentioned about whether individuals saw participants as trustworthy or not, the player's power made it more likely that participants thought that this player viewed them as trustworthy when trusting them, but untrustworthy when distrusting them. This replicates Experiments 1–3.

Presence of Metaperceptions Manipulation. I did not observe a two-way interaction effect between being trusted versus distrusted and power when the metaperceptions manipulation was present, $F(1, 699) = 1.77, p = .18, \eta_p^2 = .00$. I also found a main effect of being trusted versus distrusted, $F(1, 700) = 1455.44, p < .001, \eta_p^2 = .67$. I found no main effect of power, $p = .75$. When participants received a message about their perceived trustworthiness or untrustworthiness,

participants perceived no significant differences between power players and control players who trusted ($M = 5.65, SD = 0.91$; $M = 5.51, SD = 0.98$, respectively) or distrusted them ($M = 2.52, SD = 1.44$; $M = 2.60, SD = 1.14$, respectively), $ps > .18$. This demonstrates that the metaperceptions of trustworthiness manipulation was successful in inducing a perception among participants that the other player considered them trustworthy or untrustworthy in the Trust Game, regardless of player power.

(Dis)Trust in the Other Player

A logistic regression analysis with power and the being trusted versus distrusted conditions, and their interaction, as independent variables, and participants' decision to trust or distrust as dependent variable, yielded an interaction effect between the being trusted versus distrusted conditions and power, $B = .12, SE = 0.05, Wald = 6.14, p = .013, Exp(B) = 1.13$ (for the main effects, see Supplemental Table S14). I decomposed this four-by-two interaction effect by examining the interaction effect between being trusted versus distrusted when the metaperceptions of trustworthiness manipulation was absent or present.

Absence of Metaperceptions Manipulation. I observed a two-way interaction effect between being trusted versus distrusted and power when the manipulation was absent, $B = 0.94, SE = 0.18, Wald = 29.03, p < .001, Exp(B) = 2.56$, in addition to a main effect of being trusted versus distrusted, $B = 1.74, SE = 0.17, Wald = 106.41, p < .001, Exp(B) = 5.69$. I found no main effect of power, $p = .08$. Consistent with my hypothesis that power motivates trust reciprocity, when participants did not receive a message about their trustworthiness, being trusted versus distrusted predicted participants' decision to trust or distrust the other player more strongly when this player had power, $B = 2.75, SE = 0.27, Wald = 105.05, p < .001, Exp(B) = 15.67$, compared to when they did not, $B = .87, SE = 0.22, Wald = 15.09, p < .001, Exp(B) = 2.39$. In fact, when participants first received \$1, more participants gave \$1 to the players who had power (83.6%) compared to players who did not have power (70.5%), $B = .38, SE = 0.13, Wald = 8.45, p = .004, Exp(B) = 1.46$. When players first received \$0, more participants gave \$0 to players with power (75.4%) than players without power (50%), $B = -.56, SE = 0.12, Wald = 23.49, p < .001, Exp(B) = 0.57$. This pattern of findings closely replicates Experiments 1–3.

Presence of the Metaperceptions Manipulation. I did not observe a two-way interaction effect between being trusted versus distrusted and power in the presence of the metaperceptions manipulation, $B = -.07, SE = 0.19, Wald = 0.12, p = .73, Exp(B) = 0.94$. I did find a main effect of being trusted versus distrusted, $B = -2.59, SE = 0.19, Wald = 182.43, p < .001, Exp(B) = .08$, but no main effect of power, $p = .78$. When participants received a message about their expected trustworthiness or untrustworthiness, I observed no differences in participants reciprocating trust or distrust for power players versus control players. Participants were as likely to give \$1 when the other player first gave them \$1 and had power (86.4%) or did not have power (85.7%), $p = .86$. Participants were also as likely to give \$0 when the other player first gave them \$0 and had power (69.3%) or did not have power (67.6%), $p = .73$.

These findings demonstrate that when nothing was mentioned about whether a player saw participants as trustworthy or untrustworthy, this player's power motivated participants' trust and distrust reciprocity (cf. Experiments 1–3). It also demonstrates that the

metaperceptions of trustworthiness manipulation reduced the extent to which power motivated participants' trust and distrust reciprocity (Table 2).⁵

Mediation Analysis

For the conditions where the metaperceptions manipulation was absent, I ran a bootstrapping mediation analysis (10,000 resamples) with being trusted versus distrusted as the independent variable, metaperceptions of trustworthiness as mediator, and trust (\$1) or distrust (\$0) in the player as dependent variable. This analysis yielded a moderated mediation effect, $b = .25$, $SE = .09$, 95% [.09, .44]. As is visible in Figure 8, participants' perception that the other player viewed them as trustworthy correlated with participants' trust in, versus distrust toward, the other player. The effect of being trusted versus distrusted on participants' greater trust in the power player compared to the control player was mediated by the perception that the power player saw them as more trustworthy when trusting versus distrusting them ($b = .72$, $SE = .23$, 95% [.27, 1.18]) compared to the control player ($b = .48$, $SE = .15$, 95% [.18, .79]). I then ran the same analysis but now for the conditions where the metaperceptions manipulation was present. This analysis yielded no moderated mediation effect ($b = -.06$, $SE = .05$, 95% [-.17, .03]) showing that the manipulation prevented the indirect effect of metaperceptions of trustworthiness from materializing.

Discussion

Experiment 4 replicated Experiments 1–3 and provided causal evidence for the relationship between participants' metaperceptions of trustworthiness and their trust in, versus distrust toward, other players in the Trust Game. Taken together, Experiments 1–4 provide converging correlational and causal evidence for my hypotheses across different samples, research designs, and measurements. In Experiment 5, the final experiment, I tested the full theoretical model one last time in a field study.

Experiment 5

In Experiment 5, I collected data in the field using a sample of coworker dyads. I instructed one randomly chosen coworker to either accept vulnerability or not accept vulnerability to their coworker (the focal participant). I measured coworkers' power over participants. One week after these instructions, I measured participants' attribution of dispositions to the behaviors of their coworker, their metaperceptions of trustworthiness regarding their coworker, and their trust in or distrust toward their coworker. Doing so allowed me to test my hypotheses in a field study, temporally separate the predictor variable from subsequent variables, and test the full theoretical model one final time.

Procedure

Participants and Design

I recruited 300 coworker dyads (total $N = 600$; 310 males; $M = 40.12$, $SD = 13.16$) through the community subject pool of a U.S. university and randomly assigned them to three conditions: being trusted, control, and being distrusted. I based the number of participants on the availability of coworker dyads rather than an

a priori power analysis and preregistered the number of dyads at 300. A sensitivity analysis with 90% power and 100 participants per group demonstrated that this sample size allowed me to detect a 0.23 slope difference between two different groups. Because slope differences in Experiment 1—that used a similar design as Experiment 5—ranged from 0.18 to 0.41, Experiment 5's sample size provided enough (albeit somewhat low) power to detect an effect. Within each condition, I randomly decided which coworker would be instructed to accept vulnerability to their coworker, not accept vulnerability to their coworker, or were not given any explicit instructions. I label the coworker that is instructed to accept, not accept vulnerability, or not given these instructions as the coworker and the other coworker as the participant. I did this because the participant answered most questions and was therefore the more focal point of attention in the description of results below. I preregistered at https://aspredicted.org/ZJ3_KRD.

Procedure

Power

I asked participants to indicate how much power their coworker had over them using the same items as in Experiment 1 ($\alpha = .94$; $M = 4.54$, $SD = 1.66$).

Being Trusted and Distrusted Manipulation

At the same time as I took the first measurements, I instructed the coworker in the being trusted (being distrusted) condition to trust or distrust the participant:

in the next week, try to engage in behavior that shows to your coworker that you accept vulnerability [do not accept vulnerability] to their future actions. For instance, as their coworker, you can involve them in important workplace assignment or share organizational resources and secrets [for instance, as their coworker, you can refuse to involve them in important workplace assignment or not share organizational resources and secrets]. Please engage in a behavior that works for you in your context—the key is not to simply engage in a behavior that we suggest but to try to accept vulnerability [to try to not accept vulnerability] to the future actions of your coworker through behavior that works for you and your coworker.

Coworkers in the control condition were not given these instructions.

Dispositional Attribution

One week after the being trusted or distrusted manipulation instructions, I asked participants to what extent their coworker's behaviors in the workplace reflected their dispositions using similar items as in Experiment 2 (e.g., "My coworker's behavior in the workplace reflects their personal preference to act that way"; $\alpha = .78$;

⁵ Some might wonder what it means for practitioners that the metaperceptions of trustworthiness manipulation in Experiment 4 strongly attenuated the impact of power. Consistent with my theory, the manipulation seemed to reduce ambiguity about whether power players and control players saw participants as trustworthy. Consequently, it reduced the impact of power on the likelihood that an individual is trusted. Of course, the "real world" is messier than a controlled laboratory experiment, but these findings suggest that explicit affirmations of perceived trustworthiness prevent power from affecting the trust development process too strongly—at least in the short-term.

Table 2

Participants (Pp) Trusting or Distrusting as a Function of Being Trusted or Distrusted, the Power of the Target Individual, and the Metaperceptions of Trustworthiness or Untrustworthiness Manipulation

Trust and distrust reciprocity	Being trusted by		Being distrusted by	
	Power player	Control player	Power player	Control player
	Metaperception manipulation: absent (present)	Metaperception manipulation: absent (present)	Metaperception manipulation: absent (present)	Metaperception manipulation: absent (present)
Pp trust	83.6% (86.4%)	70.5% (85.7%)	24.6% (29.5%)	50.0% (31.2%)
Pp distrust	16.4% (13.6%)	29.5% (14.3%)	75.4% (69.3%)	50.0% (67.6%)

Note. Percentages within brackets represent the percentages when the metaperceptions of trustworthiness or untrustworthiness manipulation is present. Experiment 3: $N = 1,407$.

$M = 4.56$, $SD = 1.11$). A PCA showed that all items loaded on one (the first) component. I therefore combined all items into one scale.

Metaperceptions of Trustworthiness

At the same time as the dispositional attribution items, participants filled out the same metaperception items as in Experiment 1 ($\alpha = .99$; $M = 4.46$, $SD = 1.81$). A PCA showed that all items loaded on one (the first) component. I therefore combined all items into one scale.⁶

(Dis)Trust in Coworker

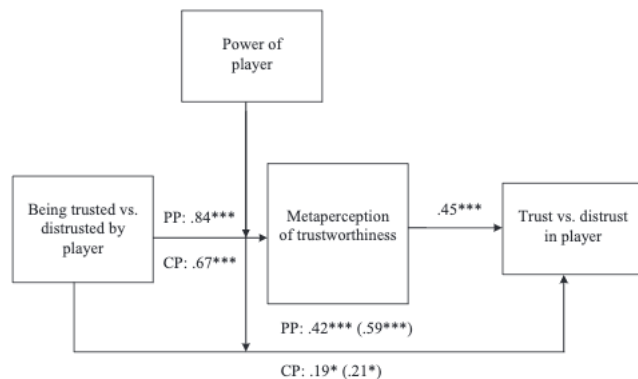
At the same time as the dispositional attribution items, participants filled out the same (dis)trust items as in Experiment 1 ($\alpha = .90$; $M = 4.02$, $SD = 1.39$). A PCA showed that all items loaded on one (the first) component, and I therefore combined all items into one scale.⁶

Being Trusted and Distrusted Manipulation Check

At the same time as the dispositional attribution items, the coworkers indicated whether and how they had accepted or not accepted vulnerability to their coworker.

Figure 8

Mediation Model for Experiment 4



Note. Coefficients within brackets are direct effects. PP = power player; CP = control player.

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

Results

Being Trusted and Distrusted Manipulation Check

All coworkers indicated that they engaged in behavior that meant accepting or not accepting vulnerability to participants (e.g., telling them personal secrets, sharing organizational tips and resources, involving them in workplace projects; withholding resources etc.).⁷

(Dis)Trust in Coworker

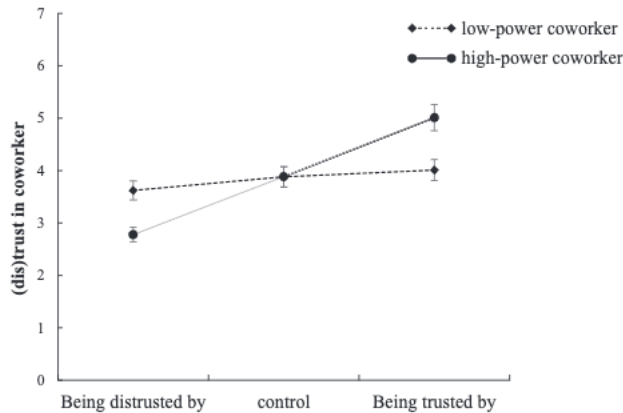
I conducted a multiple regression analysis that included the being trusted or distrusted manipulation (1 = *being trusted*, 0 = *control*, -1 = *being distrusted*), power, and their interaction as predictors. The (dis)trust scale was the dependent variable. This analysis yielded the predicted being trusted or Distrusted \times Power interaction, $\beta = .26$, $t(299) = 4.96$, $p < .001$, 95% CI [.27, .63], $rsp = .26$ (see Supplemental Table S16, for the main effects). To decompose this interaction effect, I ran two multiple regression analyses that either included as predictors the being trusted contrast (1 = *being trusted*, -1 = *control*) or the being distrusted contrast (1 = *being distrusted*, -1 = *control*) alongside power and their interaction term.

For the being trusted contrast, I found a main effect of being trusted, $\beta = .19$, $t(199) = 2.75$, $p = .007$, 95% CI [.07, .42], $rsp = .19$, as well as the predicted interaction effect between being trusted and power, $\beta = .20$, $t(199) = 2.87$, $p = .005$, 95% CI [.08, .41], $rsp = .20$. I found no main effect of power, $\beta = .12$, $t(199) = 1.68$, $p = .094$, 95% CI [-.03, .31], $rsp = .12$. Consistent with my hypothesis that power motivates trust reciprocity, being trusted predicted participants' trust in their coworker more strongly when this coworker had high-power, +1 SD ; $\beta = .38$, $t(199) = 4.01$, $p < .001$, 95% CI [.25, .72], $rsp = .27$, compared to low-power, -1 SD ; $\beta = .00$, $t(199) = .01$, $p = .99$, 95% CI [-.24, .24], $rsp = .00$. As is visible in Figure 9, the slopes for high- versus low-power coworkers were significantly different, $t(196) = 2.88$, $p = .004$, and participants trusted high-power coworkers more than low-power coworkers when participants were

⁶ For all experiments, the metaperceptions of trustworthiness items and (dis)trust items loaded on two different factors with no, or few, cross-loadings when conducting principal component analyses.

⁷ Due to privacy reasons, the exact behaviors participants wrote down are not included in the publicly available data set on OSF. Similarly, due to privacy reasons, the examples given for the stories written in Experiment 1 are also not included in the publicly available data set on OSF. This was done to minimize the risk that participants could find their own coworker in the data set and see how much they were trusted or distrusted.

Figure 9
The Degree to Which Participants Trusted or Distrusted Their Coworker as a Function of Power and Being Trusted or Distrusted



Note. Error bars represent *SE*. Experiment 5. *SE* = standard error.

trusted by these coworkers, $\beta = .31$, $t(99) = 3.25$, $p = .002$, 95% CI [.15, .61], $rsp = .31$.

For the being distrusted contrast, I found main effects of being distrusted, $\beta = -.24$, $t(199) = 3.52$, $p < .001$, 95% CI [-.51, -.14], $rsp = .24$, and power, $\beta = -.21$, $t(199) = 3.14$, $p = .002$, 95% CI [-.49, -.11], $rsp = -.21$, as well as the predicted interaction effect between being distrusted and power, $\beta = -.18$, $t(199) = -2.76$, $p = .006$, 95% CI [-.45, -.07], $rsp = .18$. Consistent with my hypothesis that power motivates distrust reciprocity, being distrusted predicted participants' distrust toward their coworker more strongly when this coworker had high-power, +1 *SD*; $\beta = -.43$, $t(199) = -4.46$, $p < .001$, 95% CI [-.85, -.33], $rsp = .29$, compared to low-power, -1 *SD*; $\beta = -.05$, $t(199) = -.57$, $p = .57$, 95% CI [-.33, .18], $rsp = .04$. As is visible in Figure 9, the slopes for high- versus low-power coworkers were significantly different from each other, $t(196) = 2.83$, $p = .005$, and participants distrusted high-power coworkers more than low-power coworkers when participants were first distrusted by these coworkers, $\beta = -.34$, $t(99) = 4.01$, $p < .001$, 95% CI [-.89, -.30], $rsp = .37$.

Mediation Analyses

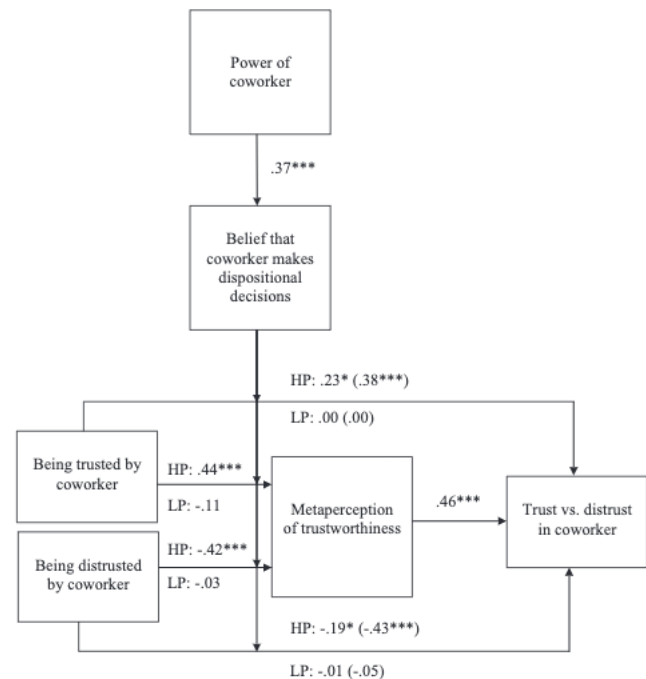
Following the mediation approach of Experiment 3 (control condition), I ran a bootstrapping analysis (10,000 resamples) with the being trusted (coded as 1) versus control (coded as -1) contrast as independent variable, metaperceptions as mediator, participants' (dis)trust in their coworker as dependent variable, and this coworker's power as moderating the relationship between being trusted and metaperceptions of trustworthiness. I added an effect to this model where the effect of power occurred through dispositional attribution. I also ran the same analysis but now with the being distrusted (coded as 1) versus control (coded as -1) contrast as independent variable. See Supplemental Tables S17 and S18 for main effects and interaction effects on dispositional attribution and metaperceptions.

These two analyses yielded two significant moderated mediation effects for power ($b = .08$, $SE = .03$, 95% [.03, .14] for being trusted; $b = -.07$, $SE = .03$, 95% [-.14, -.02] for being distrusted). As is

visible in Figure 10, metaperceptions predicted trust in (vs. distrust toward) the coworker. The effect of being trusted on participants' greater trust in high- compared to low-power coworkers was mediated by participants' perception that high-power coworkers that trusted them viewed them as more trustworthy ($b = .22$, $SE = .07$, 95% [.10, .37]) than low-power coworkers that trusted them ($b = -.06$, $SE = .06$, 95% [-.18, .05]). Similarly, the effect of being distrusted on participants' greater distrust toward high- compared to low-power coworkers was mediated by participants' perception that high-power coworkers that distrusted them viewed them as more untrustworthy ($b = -.37$, $SE = .09$, 95% [-.55, -.20]) than low-power coworkers that distrusted them ($b = -.14$, $SE = .06$, 95% [-.27, -.02]).

Importantly, power strengthened the indirect effect of metaperceptions of trustworthiness through participants' belief that high-power coworkers' decision to trust or distrust reflected their dispositions ($b = .06$, $SE = .01$, 95% [.03, .09] for the being trusted contrast; $b = -.04$, $SE = .01$, 95% [-.10, -.01] for the being distrusted contrast). As is visible in Figure 10, power predicted dispositional attribution. For the being trusted contrast, the indirect effect of metaperceptions was stronger when the individual seen as making relatively dispositional decisions (+1 *SD*; $b = .14$, $SE = .06$, 95% [.04, .27]) compared to relatively nondispositional decisions (-1 *SD*; $b = .01$, $SE = .06$, 95% [-.12, .11]). For the being distrusted contrast, the indirect effect of metaperceptions was stronger when this individual was seen as making relatively dispositional decisions (+1 *SD*; $b = -.43$, $SE = .10$, 95% [-.66, -.24]) compared to

Figure 10
Mediation Model for Experiment 5



Note. Coefficients within brackets are direct effects. HP = high power; LP = low power.
 * $p < .05$. ** $p < .01$. *** $p < .001$.

relatively nondispositional decisions (-1 *SD*; $b = -.07$, $SE = .05$, 95% [$-.18$, $.04$]).

Discussion

Experiment 5's findings provide additional and converging support for my hypotheses in a field study with coworker dyads.

General Discussion

Across five preregistered experiments, I demonstrated that power amplifies trust and distrust reciprocation. People were more likely to trust (distrust) high-power individuals than low-power individuals when these individuals first trusted (distrusted) them. This occurred because people thought that—compared to low-power individuals—high-power individuals trusted and distrusted because of dispositions rather than the situation. This led people to believe that high-compared to low-power individuals viewed them as more trustworthy when trusting them, but more untrustworthy when distrusting them. Taken together, Experiments 1–5 provide consistent evidence for my hypotheses across different samples and research designs.

Theoretical Implications

The present set of findings have important theoretical implications. First, while prior work on power and trust has taken a unidirectional theoretical lens—focusing on one individual initiating trust in another individual—the present set of findings show that there is value in taking a *bidirectional* perspective, which involves an individual responding to being trusted or distrusted by another individual (Korsgaard et al., 2015). In taking this bidirectional perspective, the current article moves research from the impact of power on trust initiation (e.g., Feenstra et al., 2020; Mooijman et al., 2015, 2019; Schilke et al., 2015) to the impact of power on trust reciprocation. Doing so provides a more accurate and complete picture of the relationship between power and trust, as people do not merely initiate trust or distrust but also respond to it. Indeed, successful trust development between two individuals involves one individual initiating trust and another individual reciprocating this trust. When an individual initiates trust that is not reciprocated, mutual trust is unlikely to develop and relationships may develop distrust and conflict (Korsgaard et al., 2015; Parco et al., 2002). Without an understanding of how power impacts the initiation *and* reciprocation of trust, research is at risk of painting a picture on power and trust that is too simplistic. The present findings suggest, for instance, that a high-power individual's initial trust or distrust is more likely to set the tone for the rest of the exchange, leading high-power individuals to elicit mutual trust or distrust more easily from the low-power individual rather than vice versa (e.g., see Ferrin et al., 2008, for how perceived trustworthiness can spiral over time). The present research therefore provides a blueprint for understanding the impact of power on long-term bidirectional trust and distrust development.

Second, the current article shows *when* people are more likely to trust or distrust high- compared to low-power individuals—that is, when they are first trusted or distrusted by these individuals—and the psychological process by which this occurs. This is important because the current literature is dominated by inconclusive and opposing findings on the relationship between power and trust.

Research has shown that people are more likely to trust (Schilke et al., 2015; Smith & Overbeck, 2014; Van der Toorn et al., 2011; Weber et al., 2005) or distrust (Farrell, 2004; Hardin, 2001; Ross et al., 2001; Schaerer et al., 2021) high- compared to low-power individuals. Until the current findings, there was no clear understanding of *when* these opposing effects occurred. In fact, because the current article focused on the impact of power on trust reciprocity, it helps us place prior work on power and trust initiation in a more complete context. Prior research has shown, for instance, that power can make individuals more likely to initiate distrust (e.g., Feenstra et al., 2020; Inesi et al., 2012; Mooijman et al., 2015, 2019; Schilke et al., 2015), suggesting that power is a trust liability. High-power individuals may distrust others *and* they may be more likely to have this distrust reciprocated. At the same time, the current article also suggests that power can be a trust asset. When high-power individuals trusted others, their power helped them become *more* trusted. Combining this insight with research suggesting that power might make individuals more likely to initiate trust (Anderson & Galinsky, 2006; Brion & Anderson, 2013), this challenges and nuances the emerging scholarly conclusion that power differences foster distrust in relationships (du Plessis et al., 2023; Mooijman & Graham, 2018). The present research, then, shows when and why power is a trust development asset *or* liability, and provides a more theoretically integrated understanding of the relationship between power and trust.

Third, the current article suggests that people view high-power individuals' trust and distrust as *more* authentic than low-power individuals' trust and distrust. Indeed, their behavior was perceived as more consistent with their internal psychological state (e.g., expectations about others' trustworthiness and untrustworthiness; Mayer et al., 1995). This finding informs the literature on people's lay beliefs about power. Kim et al. (2017) showed, in contrast to the current findings, that people viewed high-power individuals as *less* authentic than low-power individuals when these individuals expressed sadness and regret about violating a norm. This effect was explained by people's perception that high-power individuals have more control over their emotions and use these emotions more strategically than low-power individuals (e.g., to avoid punishment after a norm violation). How can these findings be reconciled with the current set of findings? It is possible that the perceived authenticity of high-power individuals' actions varies with whether their behavior occurs in the aftermath of a norm violation or not—dispositional attribution might make them seem *less* authentic when these dispositions are seen as acting in the service of manipulating others to gain favors and escape punishment. Whereas dispositional attribution might make them seem *more* authentic when these dispositions are not seen as acting in the service of manipulating others to gain favors and escape punishment. These implications are also relevant for the literature on being trusted. Research commonly uses items such as “this individual gives me influence over issues that are important to him/her” to measure the degree to which people think another individual trusts them (e.g., Baer et al., 2015; Lau et al., 2014; Lau & Lam, 2008; Skiba & Wildman, 2019; Wang & Huang, 2019). Without measuring relational power dynamics, participants may vary in whether they perceive these behaviors as authentic signals of an individual's expectation that they are trustworthy or untrustworthy. The failure to consider the role of power could therefore mean that researchers rely

on items that are interpreted differently by participants depending on hidden relational power dynamics.

Last, the current article engaged in competing theory testing, showing that the present findings are best explained by a causal attribution account (Weiner, 1992) rather than competing theoretical accounts (e.g., encapsulated interest theories, stereotype-content models, resource dependence theories; Farrell, 2004; Fiske et al., 2002; Van der Toorn et al., 2011; Weber et al., 2005). Indeed, consistent with the notion that people attribute more dispositions to the decisions of high- rather than low-power individuals, people saw high-power individuals' decision to trust or distrust as a stronger reflection of their trustworthiness expectations, which motivated both trust and distrust reciprocity. These findings suggest that a causal attribution perspective on the relationship between power and trust deserves more research attention. For instance, although prior work has documented that people believe power holders to be dispositionally motivated (Overbeck et al., 2006), few studies have attempted to understand the consequences of this belief. The present article highlights a novel consequence: it makes being trusted or distrusted a stronger antecedent of people's metaperceptions of trustworthiness. In fact, because general metaperceptions are currently understudied in power research—even though people are motivated to think about how power holders view them (Fiske, 1993; Lammers et al., 2008)—and metaperceptions of trustworthiness had to date not been studied at all in the power literature, the current article breaks new theoretical ground.

Practical Implications

The current article also provides practical implications for high-power individuals in teams, organization, and institutions. Instead of trying to promote trust and prevent distrust by flattening power hierarchies and reducing resource inequalities (as research on hierarchy suggests doing; Wilkinson & Pickett, 2007, 2009), the findings reported in the current article suggest that power hierarchies and high levels of trust in power holders can coexist, as long as individuals with power trust others. This is an important implication, as being trusted benefits power holders. When people trust high-power others, they view them more positively (Ferrin et al., 2008), report higher levels of relationship satisfaction (Wieselquist et al., 1999), and cooperate more effectively with them (Balliet & Van Lange, 2013). The advice to trust others is therefore more likely to be followed by power holders than the advice to flatten power hierarchies and create more egalitarian social structures. Indeed, although greater resource equality is associated with greater levels of trust (e.g., Wilkinson & Pickett, 2007), the advice to flatten power hierarchies means that power holders must give up some of their power. Given the fact that power holders want to maintain their power and associated privileges—even at others' expense (Case & Maner, 2014; Fehr et al., 2013)—this latter advice is unlikely to be followed by power holders. In contrast, the advice to trust, and thus invite reciprocal trust, aligns with power holders' self-interest. When powerful authorities are trusted by others, they gain deference and cooperation from these others (Kirchler et al., 2014; Mooijman et al., 2017; Tyler et al., 2015).

It is worth noting that although the advice that high-power individuals should trust more may seem intuitive, it would have also been possible for people to view high-power individuals' decision to trust them with suspicion. People could have assumed

that the powerful are corrupt (Bendahan et al., 2015), strategic about hiding their genuine intentions (Kim et al., 2017), and uninterested in encapsulating others' interests in their behavior (Galinsky et al., 2006). This would have made it less likely that people trusted power holders (Schaerer et al., 2021). In that case, advising high-power individuals to trust exposes them to the risk of exploitation without the benefits of possible mutual trust development. The current article demonstrates, however, that the benefit of trusting is more pronounced for high- rather than low-power individuals (see also De Cremer et al., 2018). This is perhaps also a surprising conclusion given prior research that has shown how people hold the lay belief that the powerful manipulate those around them for their own goals (Belmi & Laurin, 2016; see also Gruenfeld et al., 2008).

Moreover, the present article suggests that distrusting others is more detrimental to high-power than low-power individuals. When high-power individuals distrust, which prior works suggests they may be inclined to do (du Plessis et al., 2023; Feenstra et al., 2020; Inesi et al., 2012; Mooijman et al., 2015, 2019; Schilke et al., 2015; Weiss et al., 2021), power amplifies the negative effects of distrusting others. Thus, not only is trusting others more beneficial for high-power individuals, but distrusting others is also particularly harmful for high-power individuals. Interestingly, this means that low-power individuals may be able to distrust others with fewer negative consequences. Thus, having power acts like a double-edged sword when it comes to trust development. Trusting others is particularly effective for promoting reciprocal trust, whereas distrusting others is particularly destructive for promoting reciprocal trust.

Limitations and Future Research Directions

Although I presented five experiments that provided evidence for my hypotheses across different samples, research designs, and measurements, there are still limitations and future research directions to be noted. For instance, because I focused on trust and distrust reciprocity, I did not include in my research designs the possibility for people to validate or violate another party's trust in them (e.g., giving back, or keeping, money given to you in a Trust Game). To fully understand the impact of power on longer term trust dynamics, however, researchers should include this possibility. Following research on power and selfish behavior (Dubois et al., 2015), it is possible that high-power individuals are more likely to violate the trust placed in them compared to low-power individuals. How will this affect the longer term course of trust development? There are several possibilities. It could reduce the extent to which high-power individuals are able to sustain relationships with high levels of reciprocal trust, as people may respond particularly negatively to a trust violation from a high-power individual who is perceived as making dispositional decisions. It is also possible that people respond less negatively to high-power individuals violating their trust because they expect these individuals to be selfish and want to maintain a functioning relationship with an individual they depend on. Future research could examine which possibility has more empirical support and examine how power shapes longer term, reciprocal trust development where trust violations are possible.

In addition, the act of trusting or distrusting—when related to giving or withholding resources—is in itself an exercise of power. This means that the current article, like all research on power and trust to date, cannot speak to the impact of complete powerlessness

on trust and distrust (Handgraaf et al., 2008). Instead, it shows the impact of relatively high or low levels of power. This is not a problem from a theoretical or practical perspective. Power is a relational construct, which means that individuals have high or low power relative to each other within their relationship (Magee & Galinsky, 2008). In addition, the experiments in the current article showed that being trusted or distrusted did not impact participants' perception that the other party had power over them. That said, future research on trust could separate low power from powerlessness. Although it is unlikely that behavioral measures of trust can be used in such future research—since these measures grant resources and power to the decision-maker—future research could focus on differentiating the impact of low power and powerlessness on trust expectations (e.g., benevolence, integrity, ability). It is possible that having low, but some, power has a different impact on trust and distrust expectations than having no power at all. Perhaps having no resources at all makes people more motivated to view others positively whereas having a little bit of power may highlight that the other party can still attempt to exploit you and take your resources away (du Plessis et al., 2023; Schilke et al., 2015).⁸

Last, prior research has shown that power and status can have opposing effects on trust initiation (Lount & Pettit, 2012; Mooijman et al., 2015). Although speculative, this raises the possibility that the current set of findings on power do not generalize to status. Status, as opposed to power, can be defined as the respect and admiration that an individual has in the eyes of others (Blader & Chen, 2012). This means that high-status individuals do not necessarily control resources and, as such, may also not be seen as making dispositional decisions. This implies that power dynamics, but not necessarily status dynamics, shape metaperceptions of trustworthiness and motivate trust and distrust reciprocity. Future research could examine this possibility. Similarly, I treated trust and distrust as opposites, manipulating them in a dichotomous manner (you are either trusted or distrusted). Research on ambivalence suggests, however, that people can simultaneously hold positive and negative attitudes toward someone (Schneider & Schwarz, 2017; see also Lewicki et al., 1998). Although the prevalence of simultaneous trust and distrust may be low, future research could attempt to tease the constructs apart. Future research could use Trust Games where people decide to give a certain amount of money (anywhere between 0 and 5 dollars) instead of the dichotomous Trust Games used in the current article (giving or not giving money). Although dichotomous Trust Games are common (Berg et al., 1995) and consistent with the notion that trust involves accepting vulnerability whereas distrust does not (Kramer, 1999), they do not allow for people to draw ambiguous, ambivalent inferences about whether someone else trusts or distrusts them. Perhaps the impact of power on trust and distrust reciprocation is less pronounced when research designs allow for behaviors that signal more ambivalent trust attitudes (e.g., giving half of a \$1 endowment in a TG).

Conclusion

In the current article, I presented five preregistered experiments that examined the impact of power on trust and distrust reciprocation. I show that people are more likely to trust (distrust) high-power individuals than low-power individuals when these individuals first trusted (distrusted) them. This effect is caused by people's belief that

high-power individuals are more likely to trust and distrust because of dispositions, which led people to think that high-power individuals view them as more trustworthy when trusting them, but more untrustworthy when distrusting them. These findings broaden our knowledge about the complex relationships between power, trust, and distrust, and provide practical implications for high-power individuals on how to have power *and* be trusted.

⁸ Of course, high-power compared to low-power individuals have more resources to give (lottery tickets + \$1) and, as such, may also be viewed as more obligated to give at least of these resources to others in trust settings (Dunning et al., 2019). Although I did not find, as reported in the online Supplemental Materials, that participants expected high-power individuals to be more likely to give resources to them, it is possible that the perception that high-power individuals are more obligated to give resources contributed to people reacting more negatively to high-power individuals choosing to withhold rather than give resources—that is, violate the obligation to give. Perhaps it also contributed to people reacting more positively to high-power individuals choosing to give rather than withhold resources—that is, fulfill the obligation to give. Since I did not measure perceived obligation explicitly in the current set of experiments, future research could examine this possibility in more detail.

References

- Anderson, C., & Galinsky, A. D. (2006). Power, optimism, and risk-taking. *European Journal of Social Psychology, 36*(4), 511–536. <https://doi.org/10.1002/ejsp.324>
- Baer, M. D., Dhensa-Kahlon, R. K., Colquitt, J. A., Rodell, J. B., Outlaw, R., & Long, D. M. (2015). Uneasy lies the head that bears the trust: The effects of feeling trusted on emotional exhaustion. *Academy of Management Journal, 58*(6), 1637–1657. <https://doi.org/10.5465/amj.2014.0246>
- Baer, M. D., Frank, E. L., Matta, F. K., Luciano, M. M., & Wellman, N. (2021). Undertrusted, overtrusted, or just right? The fairness of (in)congruence between trust wanted and trust received. *Academy of Management Journal, 64*(1), 180–206. <https://doi.org/10.5465/amj.2018.0334>
- Balliet, D., & Van Lange, P. A. (2013). Trust, conflict, and cooperation: A meta-analysis. *Psychological Bulletin, 139*(5), 1090–1112. <https://doi.org/10.1037/a0030939>
- Bauman, C. W., Tost, L. P., & Ong, M. (2016). Blame the shepherd not the sheep: Imitating higher-ranking transgressors mitigates punishment for unethical behavior. *Organizational Behavior and Human Decision Processes, 137*, 123–141. <https://doi.org/10.1016/j.obhdp.2016.08.006>
- Belmi, P., & Laurin, K. (2016). Who wants to get to the top? Class and lay theories about power. *Journal of Personality and Social Psychology, 111*(4), 505–529. <https://doi.org/10.1037/pspi0000060>
- Bendahan, S., Zehnder, C., Pralong, F. P., & Antonakis, J. (2015). Leader corruption depends on power and testosterone. *The Leadership Quarterly, 26*(2), 101–122. <https://doi.org/10.1016/j.leaqua.2014.07.010>
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122–142. <https://doi.org/10.1006/game.1995.1027>
- Blader, S. L., & Chen, Y. R. (2012). Differentiating the effects of status and power: A justice perspective. *Journal of Personality and Social Psychology, 102*(5), 994–1014. <https://doi.org/10.1037/a0026651>
- Brion, S., & Anderson, C. (2013). The loss of power: How illusions of alliance contribute to powerholders' downfall. *Organizational Behavior and Human Decision Processes, 121*(1), 129–139. <https://doi.org/10.1016/j.obhdp.2013.01.005>
- Brion, S., Mo, R., & Lount, R. B., Jr. (2019). Dynamic influences of power on trust: Changes in power affects trust in others. *Journal of Trust Research, 9*(1), 6–27. <https://doi.org/10.1080/21515581.2018.1552591>

- Burgoon, J. K. (1993). Interpersonal expectations, expectancy violations, and emotional communication. *Journal of Language and Social Psychology, 12*(1–2), 30–48. <https://doi.org/10.1177/0261927X93121003>
- Burke, S. C., Sims, D. E., Lazzara, E. H., & Salas, E. (2007). Trust in leadership: A multilevel review and integration. *The Leadership Quarterly, 6*(6), 606–632. <https://doi.org/10.1016/j.leaqua.2007.09.006>
- Case, C. R., Conlon, K., & Maner, J. K. (2015). Affiliation-seeking among the powerless: Lacking power increases social affiliative motivation. *European Journal of Social Psychology, 45*(3), 378–385. <https://doi.org/10.1002/ejsp.2089>
- Case, C. R., & Maner, J. K. (2014). Divide and conquer: When and why leaders undermine the cohesive fabric of their group. *Journal of Personality and Social Psychology, 107*(6), 1033–1050. <https://doi.org/10.1037/a0038201>
- Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology, 92*(4), 909–927. <https://doi.org/10.1037/0021-9010.92.4.909>
- De Cremer, D., van Dijke, M., Schminke, M., De Schutter, L., & Stouten, J. (2018). The trickle-down effects of perceived trustworthiness on subordinate performance. *Journal of Applied Psychology, 103*(12), 1335–1357. <https://doi.org/10.1037/apl0000339>
- Dirks, K. T. (2000). Trust in leadership and team performance: Evidence from NCAA basketball. *Journal of Applied Psychology, 85*(6), 1004–1012. <https://doi.org/10.1037/0021-9010.85.6.1004>
- Dirks, K. T., & De Jong, B. (2022). Trust within the workplace: A review of two waves of research and a glimpse of the third. *Annual Review of Organizational Psychology and Organizational Behavior, 9*(1), 247–276. <https://doi.org/10.1146/annurev-orgpsych-012420-083025>
- du Plessis, C., Nguyen, M. H. B., Foulk, T. A., & Schaerer, M. (2023). Relative power and interpersonal trust. *Journal of Personality and Social Psychology, 124*(3), 567–592. <https://doi.org/10.1037/pspi0000401>
- 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. <https://doi.org/10.1037/pspi0000008>
- Dunning, D., Fetschenhauer, D., & Schlosser, T. (2019). Why people trust: Solved puzzles and open mysteries. *Current Directions in Psychological Science, 4*(4), 366–371. <https://doi.org/10.1177/0963721419838255>
- Emerson, R. M. (1962). Power-dependence relations. *American Sociological Review, 27*(1), 31–41. <https://doi.org/10.2307/2089716>
- Farrell, H. (2004). Trust, distrust, and power. In R. Hardin (Ed.), *Distrust* (pp. 85–105). Sage Publications.
- Feenstra, S., Jordan, J., Walter, F., & Stoker, J. I. (2020). Antecedents of leaders' power sharing: The roles of power instability and distrust. *Organizational Behavior and Human Decision Processes, 157*, 115–128. <https://doi.org/10.1016/j.obhdp.2020.01.005>
- Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association, 2*(2–3), 235–266. <https://doi.org/10.1162/JEEA.2009.7.2-3.235>
- Fehr, E., Herz, H., & Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. *The American Economic Review, 103*(4), 1325–1359. <https://doi.org/10.1257/aer.103.4.1325>
- Ferrin, D. L., Bligh, M. C., & Kohles, J. C. (2008). It takes two to tango: An interdependence analysis of the spiraling of perceived trustworthiness and cooperation in interpersonal and intergroup relationships. *Organizational Behavior and Human Decision Processes, 107*(2), 161–178. <https://doi.org/10.1016/j.obhdp.2008.02.012>
- Fiedler, K., Harris, C. A., & Schott, M. (2018). Unwarranted inferences from statistical mediation tests. *Journal of Experimental Social Psychology, 75*, 95–102. <https://doi.org/10.1016/j.jesp.2017.11.008>
- Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist, 48*(6), 621–628. <https://doi.org/10.1037/0003-066X.48.6.621>
- 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). Wiley. <https://doi.org/10.1002/9780470561119.socpsy002026>
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902. <https://doi.org/10.1037/0022-3514.82.6.878>
- Fragale, A. R., Rosen, B., Xu, C., & Merideth, I. (2009). The higher they are, the harder they fall: The effects of wrongdoer status on observer punishment recommendations and intentionality. *Organizational Behavior and Human Decision Processes, 108*(1), 53–65. <https://doi.org/10.1016/j.obhdp.2008.05.002>
- Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology, 85*(3), 453–466. <https://doi.org/10.1037/0022-3514.85.3.453>
- 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. <https://doi.org/10.1037/a0012633>
- Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science, 17*(12), 1068–1074. <https://doi.org/10.1111/j.1467-9280.2006.01824.x>
- 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. <https://doi.org/10.1037/0022-3514.95.1.111>
- Guinote, A., Weick, M., & Cai, A. (2012). Does power magnify the expression of dispositions? *Psychological Science, 23*(5), 475–482. <https://doi.org/10.1177/0956797611428472>
- Handgraaf, M. J. J., Van Dijk, E., Vermunt, R. C., Wilke, H. A. M., & De Dreu, C. K. W. (2008). Less power or powerless? Egocentric empathy gaps and the irony of having little versus no power in social decision making. *Journal of Personality and Social Psychology, 95*(5), 1136–1149. <https://doi.org/10.1037/0022-3514.95.5.1136>
- Hanna, J. T., Elms, A. K., Gill, H., Stanley, D. J., & Powell, D. M. (2019). The effect of leader risk-taking on subordinate felt trust. *Leadership and Organization Development Journal, 40*(2), 163–176. <https://doi.org/10.1108/LODJ-11-2018-0386>
- Hardin, R. (2001). Conceptions and explanations of trust. In K. S. Cook (Ed.), *Trust in society* (pp. 3–39). Sage Publications.
- Hecht, M. A., & LaFrance, M. (1998). License or obligation to smile: The effect of power and sex on amount and type of smiling. *Personality and Social Psychology Bulletin, 24*(12), 1332–1342. <https://doi.org/10.1177/01461672982412007>
- Hinkin, T. R., & Tracey, J. B. (1999). An analysis of variance approach to content validation. *Organizational Research Methods, 2*(2), 175–186. <https://doi.org/10.1177/109442819922004>
- Inesi, E. M., Gruenfeld, D. H., & Galinsky, A. D. (2012). How power corrupts relationships: Cynical attributions for others' generous acts. *Journal of Experimental Social Psychology, 48*(4), 795–803. <https://doi.org/10.1016/j.jesp.2012.01.008>
- Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist, 73*(1), 1–2. <https://doi.org/10.1037/amp0000263>
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist, 28*(2), 107–128. <https://doi.org/10.1037/h0034225>
- Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review, 110*(2), 265–284. <https://doi.org/10.1037/0033-295X.110.2.265>
- Khademi, M., Mast, M. S., Zehnder, C., & De Saint Priest, O. (2021). The problem of demand effects in power studies: Moving beyond power priming. *The Leadership Quarterly, 32*(4), Article 101496. <https://doi.org/10.1016/j.leaqua.2021.101496>

- Kim, P. H., Mislin, A., Tuncel, E., Fehr, R., Cheshin, A., & van Kleef, G. A. (2017). Power as an emotional liability: Implications for perceived authenticity and trust after a transgression. *Journal of Experimental Psychology: General*, *146*(10), 1379–1401. <https://doi.org/10.1037/xge0000292>
- Kirchler, E., Kogler, C., & Muelbacher, S. (2014). Cooperative tax compliance: From deterrence to deference. *Current Directions in Psychological Science*, *23*(2), 87–92. <https://www.jstor.org/stable/44318728>
- Korsgaard, M. A., Brower, H. H., & Lester, S. W. (2015). It isn't always mutual: A critical review of dyadic trust. *Journal of Management*, *41*(1), 47–70. <https://doi.org/10.1177/0149206314547521>
- Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology*, *50*(1), 569–598. <https://doi.org/10.1146/annurev.psych.50.1.569>
- Lammers, J., Gordijn, E. H., & Otten, S. (2008). Looking through the eyes of the powerful. *Journal of Experimental Social Psychology*, *44*(5), 1229–1238. <https://doi.org/10.1016/j.jesp.2008.03.015>
- Lammers, J., Stoker, J. I., Rink, F., & Galinsky, A. D. (2016). To have control over or to be free from others? The desire for power reflects a need for autonomy. *Personality and Social Psychology Bulletin*, *42*(4), 498–512. <https://doi.org/10.1177/0146167216634064>
- Lau, D. C., & Lam, L. W. (2008). Effects of trust and being trusted on team citizenship behavior in chain stores. *Asian Journal of Social Psychology*, *11*(2), 141–149. <https://doi.org/10.1111/j.1467-839X.2008.00251.x>
- Lau, D. C., Lam, L. W., & Wen, S. S. (2014). Examining the effects of feeling trusted by supervisors in the workplace: A self-evaluative perspective. *Journal of Organizational Behavior*, *35*(1), 112–127. <https://doi.org/10.1002/job.1861>
- Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, *23*(3), 438–458. <https://doi.org/10.2307/259288>
- Lount, R. B., Jr., & Pettit, N. C. (2012). The social context of trust: The role of status. *Organizational Behavior and Human Decision Processes*, *117*(1), 15–23. <https://doi.org/10.1016/j.obhdp.2011.07.005>
- Magee, J. C., & Galinsky, A. D. (2008). Social hierarchy: The self-reinforcing nature of power and status. *The Academy of Management Annals*, *2*(1), 351–398. <https://doi.org/10.5465/19416520802211628>
- Mayer, R. C., & Davis, J. H. (1999). The effect of the performance appraisal system on trust for management: A field quasi-experiment. *Journal of Applied Psychology*, *84*(1), 123–136. <https://doi.org/10.1037/0021-9010.84.1.123>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734. <https://doi.org/10.2307/258792>
- Mayer, R. C., & Gavin, M. B. (2005). Trust in management and performance: Who minds the shop while the employees watch the boss? *Academy of Management Journal*, *48*(5), 874–888. <https://doi.org/10.5465/amj.2005.18803928>
- Mooijman, M., & Graham, J. (2018). Unjust punishments in organizations. *Research in Organizational Behavior*, *38*, 95–106. <https://doi.org/10.1016/j.riob.2018.10.001>
- Mooijman, M., Kouchaki, M., Beall, E., & Graham, J. (2020). Power decreases the moral condemnation of disgust-inducing transgressions. *Organizational Behavior and Human Decision Processes*, *161*, 79–92. <https://doi.org/10.1016/j.obhdp.2020.04.005>
- Mooijman, M., van Dijk, W. W., Ellemers, N., & van Dijk, E. (2015). Why leaders punish: A power perspective. *Journal of Personality and Social Psychology*, *109*(1), 75–89. <https://doi.org/10.1037/pspi0000021>
- Mooijman, M., van Dijk, W. W., van Dijk, E., & Ellemers, N. (2017). On sanction-goal justifications: How and why deterrence justifications undermine rule compliance. *Journal of Personality and Social Psychology*, *112*(4), 577–588. <https://doi.org/10.1037/pspi0000084>
- Mooijman, M., Van Dijk, W. W., Van Dijk, E., & Ellemers, N. (2019). Leader power, power stability, and interpersonal trust. *Organizational Behavior and Human Decision Processes*, *152*, 1–10. <https://doi.org/10.1016/j.obhdp.2019.03.009>
- Nunn, N., Qian, N., & Wen, J. (2018). *Distrust and political turnover during economic crises* [Working Paper 24187]. <https://doi.org/10.3386/w24187>
- Overbeck, J. R., & Drouman, V. (2013). One for all: Social power increases self-anchoring of traits, attitudes, and emotions. *Psychological Science*, *24*(8), 1466–1476. <https://doi.org/10.1177/0956797612474671>
- 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. <https://doi.org/10.1002/ejsp.353>
- Parco, J. E., Rapoport, A., & Stein, W. E. (2002). Effects of financial incentives on the breakdown of mutual trust. *Psychological Science*, *13*(3), 292–297. <https://doi.org/10.1111/1467-9280.00454>
- Pillutla, M., Malhotra, D., & Murnighan, J. K. (2003). Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology*, *39*(5), 448–455. [https://doi.org/10.1016/S0022-1031\(03\)00015-5](https://doi.org/10.1016/S0022-1031(03)00015-5)
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Reeder, G. D., Vonk, R., Ronk, M. J., Ham, J., & Lawrence, M. (2004). Dispositional attribution: Multiple inferences about motive-related traits. *Journal of Personality and Social Psychology*, *86*(4), 530–544. <https://doi.org/10.1037/0022-3514.86.4.530>
- Ross, C. E., Mirowsky, J., & Pribesh, S. (2001). Powerlessness and the amplification of threat: Neighborhoods of disadvantage, disorder, and mistrust. *American Sociological Review*, *66*(4), 568–591. <https://doi.org/10.2307/3088923>
- 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. <https://doi.org/10.1016/j.obhdp.2018.08.004>
- Schaerer, M., Foulk, T., du Plessis, C., Tu, M.-H., & Krishnan, S. (2021). Just because you're powerless doesn't mean they aren't out to get you: Low power, paranoia, and aggression. *Organizational Behavior and Human Decision Processes*, *165*, 1–20. <https://doi.org/10.1016/j.obhdp.2021.03.005>
- Schilke, O., Reimann, M., & Cook, C. S. (2021). Trust in social relations. *Annual Review of Sociology*, *47*(1), 239–259. <https://doi.org/10.1146/annurev-soc-082120-082850>
- Schilke, O., Reimann, M., & Cook, K. S. (2015). Power decreases trust in social exchange. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(42), 12950–12955. <https://doi.org/10.1073/pnas.1517057112>
- Schilke, O., Reimann, M., & Cook, K. S. (2016). Reply to Wu and Wilkes: Power, whether situational or durable, decreases both relational and generalized trust. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(11), Article E1418. <https://doi.org/10.1073/pnas.1524382113>
- Schneider, I. K., & Schwarz, N. (2017). Mixed feelings: The case of ambivalence. *Current Opinion in Behavioral Sciences*, *15*, 39–45. <https://doi.org/10.1016/j.cobeha.2017.05.012>
- Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 38, pp. 183–242). Academic Press.
- Skiba, T., & Wildman, J. L. (2019). Uncertainty reducer, exchange deepener, or self-determination enhancer? Feeling trust versus feeling trusted in supervisor-subordinate relationships. *Journal of Business and Psychology*, *34*(2), 219–235. <https://doi.org/10.1007/s10869-018-9537-x>
- Smith, P. K., & Overbeck, J. R. (2014). The leaders' rosy halo: Why do we give power holders the benefit of the doubt? In J.-W. van Prooijen & P. A. M. Lange (Eds.), *Power, politics, and paranoia: Why people are*

- suspicious about their leaders* (pp. 53–72). Cambridge University Press. <https://doi.org/10.1017/CBO9781139565417.006>
- Tyler, T. R., Jackson, J., & Mentovich, A. (2015). The consequences of being an object of suspicion: Potential pitfalls of proactive policing. *Journal of Empirical Legal Studies*, 12(4), 603–636. <https://doi.org/10.1111/jels.12086>
- Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 25, pp. 115–191). Oxford University Press. [https://doi.org/10.1016/S0065-2601\(08\)60283-X](https://doi.org/10.1016/S0065-2601(08)60283-X)
- Uskul, A. K., Paulmann, S., & Weick, M. (2016). Social power and recognition of emotional prosody: High power is associated with lower recognition accuracy than low power. *Emotion*, 16(1), 11–15. <https://doi.org/10.1037/emo0000110>
- Van der Toorn, J., Tyler, T. R., & Jost, J. T. (2011). More than fair: Outcome dependence, system justification, and the perceived legitimacy of authority figures. *Journal of Experimental Social Psychology*, 47(1), 127–138. <https://doi.org/10.1016/j.jesp.2010.09.003>
- 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. <https://doi.org/10.1111/j.1467-9280.2008.02241.x>
- Wang, H., & Huang, Q. (2019). Feeling trusted and employee outcomes: The double-edged sword of political behaviour. *Personnel Review*, 48(7), 1653–1668. <https://doi.org/10.1108/PR-11-2017-0368>
- Weber, J. M., Malhotra, D., & Murnighan, J. K. (2005). Normal acts of irrational trust: Motivated attributions and the trust development process. *Research in Organizational Behavior*, 26, 75–101. [https://doi.org/10.1016/S0191-3085\(04\)26003-8](https://doi.org/10.1016/S0191-3085(04)26003-8)
- Weiner, B. (1992). *Human motivation: Metaphors, theories, and research*. Sage Publications.
- Weiss, A., Michels, C., Burgmer, P., Mussweiler, T., Ockenfels, A., & Hofmann, W. (2021). Trust in everyday life. *Journal of Personality and Social Psychology*, 121, 95–114. <https://doi.org/10.1037/pspi0000334>
- Wieselquist, J., Rusbult, C. E., Foster, C. A., & Agnew, C. R. (1999). Commitment, pro-relationship behavior, and trust in close relationships. *Journal of Personality and Social Psychology*, 77(5), 942–966. <https://doi.org/10.1037/0022-3514.77.5.942>
- Wilkinson, R. G., & Pickett, K. E. (2007). The problems of relative deprivation: Why some societies do better than others. *Social Science & Medicine*, 65(9), 1965–1978. <https://doi.org/10.1016/j.socscimed.2007.05.041>
- Wilkinson, R. G., & Pickett, K. E. (2009). *The spirit level: Why equality is better for everyone*. Penguin Books.
- Yu, A., Hays, N. A., & Zhao, E. Y. (2019). Development of a bipartite measure of social hierarchy: The perceived power and perceived status scales. *Organizational Behavior and Human Decision Processes*, 152, 84–104. <https://doi.org/10.1016/j.obhdp.2019.03.011>
- Zand, D. (1997). *The leadership triad: Knowledge, trust, and power*. Oxford University Press.

Received October 16, 2021

Revision received February 10, 2023

Accepted February 21, 2023 ■

