Does power reduce temporal discounting? Commentary on Joshi and Fast (2013)

Min Zhang

Pamela K. Smith

University of California, San Diego

in press, *Psychological Science*

Devaluing future outcomes, known as temporal discounting (Frederick, Loewenstein, & O'Donoghue, 2002), hinders one's ability to act in line with long-term over short-term interests. It is associated with maladaptive behaviors such as smoking (Kirby, Petry, & Bickel, 1999), drug use (Harrison, Lau, & Rutström, 2010), and not saving for retirement (Gubler & Pierce, 2014). Because temporal discounting may impact a variety of behaviors ranging from exercising to energy consumption (Urminsky & Zauberman, 2015), understanding the factors that influence temporal discounting is critical for researchers in psychology, economics, business, and public policy.

Joshi and Fast (2013) provided evidence that increased social power (control over valued resources; Magee & Galinsky, 2008) reduces temporal discounting. This finding has important applied, theoretical, and empirical implications. First, it implies the potential to mitigate temporal discounting with power. Psychological and health science researchers have thus recently advocated for empowerment as an intervention to improve balancing of long-term and short-term interests in decision-making (Gubler & Pierce, 2014; Patton et al., 2016; Urminsky & Zauberman, 2015).

Second, it informs the theoretical debate on how power affects self-control. Temporal discounting may be thought of as a self-control conflict between having a smaller reward sooner versus a larger reward later (Frederick, Loewenstein, & O'Donoghue, 2002; Mischel, Rodriguez, and Shoda, 1989). The approach-inhibition theory of power (Keltner, Gruenfeld, & Anderson, 2003) posits that high power activates the behavioral approach system, which increases impulsivity and sensitivity to rewards. Therefore, high-power individuals should be more likely to prefer earlier to delayed rewards, and thus show more temporal discounting, than low-power individuals. In contrast, the social distance theory of power (Magee & Smith, 2013) predicts the reverse, that high-power individuals should show less temporal discounting than low-power individuals. According to the social distance theory, because of their greater independence, high-power individuals feel more psychologically distant from low-power individuals than vice versa. Increased psychological distance has been shown to decrease temporal discounting (e.g., Pronin, Olivola, & Kennedy, 2007), in part by leading individuals to construe situations more abstractly (Trope & Liberman, 2010), thus highlighting the value of the delayed reward (e.g., Fujita, Trope, Liberman, & Levin-Sagi, 2006).

Third, this finding could further our understanding of the self-reinforcing nature of power (Magee & Galinsky, 2008). If having power decreases temporal discounting, such heightened self-control on the part of the powerful may help maintain existing power hierarchies. In a United Nations Development Programme (2014) on poverty reduction, Sheehy-Skeffington and Haushofer employed Joshi and Fast’s (2013) finding to suggest that poverty harms one’s chance of long-term success by reducing one’s sense of power and thus one’s self-control.

Given the theoretical and real-world significance of this effect, it is important to examine its reproducibility. Other published experiments on this topic used similar procedures but produced inconsistent results (Duan, Wu, & Sun, 2017, Studies 1 and 3; Heller & Ullrich, 2017; Tost, Wade-Benzoni, & Johnson, 2015, Experiment 2). Duan et al. (2017) found that power reduced temporal discounting for Chinese participants (Study 1), but this effect was specific to participants of Han ethnicity, not participants of Tibetan ethnicity (Study 3). Both Tost et al. (2015, Experiment 2) and Heller and Ullrich (2017) produced null results. These inconsistent findings raise questions about the robustness of the effect of power reducing temporal discounting. However, the existing replication studies have critical problems that limit their conclusiveness. The studies of Tost et al. (2015, Experiment 2) and Duan et al. (2017, Studies 1 and 3) are underpowered (*N*s = 69, 78, and 80, respectively; power to detect the original effect: 64%, 70%, and 71%, respectively; power to detect a medium-sized (d = 0.5) effect: 53%, 59%, and 60%, respectively).[1] Heller and Ullrich's (2017) study suffered from differential attrition between conditions, with significantly more high-power participants (61%) dropping out of the study than low-power (50%) and control (32%) participants. Selective attrition introduces experimental confounds and violates the assumption of random assignment (Zhou & Fishbach, 2016).

Well-powered, rigorous direct replications are needed to test the validity of the original findings (Simons, 2014). To this end, we conducted preregistered close replications of two different experiments in Joshi and Fast (2013).[2] Both of our studies have sample sizes more than 2.5 times the original (see Table 1), as recommended by Simonsohn (2015) for informative replications. The problem of selective attrition was avoided by using an undergraduate student subject pool. Though students have a right to end their participation in a study at any time, they rarely do so.

Study 1 attempted to replicate Study 1 in Joshi and Fast (2013), manipulating real power by assigning participants low or high amounts of control over team members' outcomes, and measuring temporal discounting with money. Study 2 attempted to replicate Study 3 in Joshi and Fast (2013), manipulating power with a well-established recall paradigm (Galinsky, Gruenfeld, & Magee, 2003), and measuring temporal discounting with environmental outcomes. By replicating different paradigms from the original paper, we provide a strong test of the claim that power reduces temporal discounting. Below, we provide an overview of both studies as well as the critical analyses. Further details on the procedures and analyses can be found in the supplemental materials. In both studies, we also tested Joshi and Fast's (2013) proposed mediator of the effect of power on temporal discounting, connection with the future self, measured by how similar and connected participants felt to their self in the future.

Study 1

Both replication studies were run as the first study in a series of studies that lasted for about an hour. Participants completed the studies in individual cubicles in a common room. In Study 1, power was manipulated by assigning participants to a low-power worker or a high-power manager role in a virtual team task. Temporal discounting was measured with a titration procedure followed by a free-response matching question. In the titration procedure, participants made nine choices between receiving a $120 prize that day and receiving $113, $120, $137, $154, $171, $189, $206, $223, or $240 in a year. The free-response matching question asked participants to fill in an amount that made them indifferent between receiving $120 that day and receiving the filled-in amount in a year. Then, as a manipulation check, participants reported the extent to which they had power over other group members in the virtual team task. For this portion of the study, we used the same manipulation and measures as reported in Joshi and Fast (2013, Study 1).

After the original procedure, we added an attention check in which participants indicated to which role they had been assigned. We also measured connection with the future self, as well as multiple potential moderators. In particular, non-naivety of participants has been shown to reduce effect sizes (Chandler, Mueller, & Paolacci, 2014), so we included two questions to probe participants' previous experience with the power manipulation and the discounting measure. The perceived legitimacy of a person's low- or high-power position (i.e., how fair or justified it is) has been shown to moderate the effect of this position on approach tendencies (Lammers, Galinsky, Gordijn, & Otten, 2008, see also Smith, Jost, & Vijay, 2008) and on social distance

Table 1

Comparison of Effects in Joshi and Fast's (2013) Studies 1 and 3 and the Current Studies

Study	*N*	Exclusion rate	Sample	Low-power condition *M* (*SD*)	Control condition *M* (*SD*)	High-power condition *M* (*SD*)	*t*	*p*	Effect size [95% CI]	Power
Joshi & Fast (2013, Study 1)	67	8.2%	Mturk	0.73 (0.42)	-	0.43 (0.30)	-2.32	.023	d = -0.57 [-0.89, -0.24]	62.76%
Current Study 1	342	18.6%	University	0.43 (0.32)	-	0.43 (0.29)	-0.08	.940	d = -0.01 [-0.22, 0.20]	99.95%
Joshi & Fast (2013, Study 3)	78	7.6%	University	0.57_a (0.27)	0.52_a (0.24)	0.40_b (0.27)	-2.32	.023	η^2 = 0.06 [0.00, 0.19]	57.91%
Current Study 2	399	13.8%	University	0.57_a (0.26)	0.49_b (0.28)	0.55_{ab} (0.27)	0.76	.449	η^2 < .001 [0.00, 0.02]	99.96%

Notes. Samples sizes reported are the sample sizes used in each analysis, after exclusions.[3] Mturk sample refers to participants recruited from the Amazon Mechanical Turk platform who completed the study online. University sample refers to university students who completed the study in a laboratory. For the last two rows, the t-test is for a contrast of the high-power condition with an average of the low-power and control conditions (as in Joshi and Fast (2013, Study 3)), and values with different subscripts within a row differ significantly ($p < .05$), as determined by independent-samples t-tests. Power is the probability of detecting an effect of the same size or larger as that found in the relevant study of Joshi and Fast (2013). CI = confidence interval.

(Lammers, Galinsky, Gordijn, & Otten, 2012). As discussed in the Introduction, behavioral approach and social distance are the mechanisms through which power affects temporal discounting according to the approach-inhibition (Keltner et al., 2003) and social distance (Magee & Smith, 2013) theories, respectively. Thus, we included perceived legitimacy as a potential moderator, asking participants to rate how legitimate the role assignment was. We also measured socioeconomic status. Finally, after participants completed the other studies in the hour-long session, we measured two additional potential moderators: participants' trait General Sense of Power (Anderson, John, & Keltner, 2012) and goals related to money.

We followed the same exclusion criteria as Joshi and Fast (2013) and also excluded participants who incorrectly identified their assigned roles. Discount rate was calculated as in the original study with the hyperbolic discounting formula $k = (A/V - 1)/D$, where A = the future amount that made participants indifferent between the present and future rewards, V = the present reward ($120), and D = the delay (1 year). Thus, k indicates how much a participant values present rewards relative to future rewards. To correct for the positive skew of the discount rate distribution, we also excluded participants with discount rates beyond 3 interquartile ranges (Baguley, 2012).

The manipulation check confirmed that the power manipulation was effective: high-power participants (M = 4.71, SD = 1.37) reported having more power over their team members than low-power participants (M = 3.61, SD = 1.41), $t(340) = 7.33$, $p < .001$, d = 0.79, 95% CI [0.57, 1.01]. Table 1 reports the primary statistics for Study 1. Post-exclusions, Study 1 still had over 99% power to detect the original effect. In contrast to Joshi and Fast's (2013) Study 1, we found no significant difference in discount rate between power conditions. According to an equivalence test (Lakens, 2017), these data provide evidence for the null hypothesis of no effect (relative to that of an effect larger than $d = 0.38$),[4] $t(336.85) = -3.42$, $p < .001$. We also ran complementary nonparametric tests on the untrimmed discount rate, as well as discount rates based solely on titration responses or matching responses (alternative temporal discounting calculations used in some research, e.g., Hardisty, Thompson, Krantz, & Weber, 2013), to test the robustness of our findings. These analyses also showed no effect of power. Additionally, power had no effect on connection with the future self.

Out of the five potential moderators we tested, only perceived legitimacy significantly moderated the effect of power on temporal discounting. When the role assignments were perceived as low in legitimacy, we found a pattern consistent with the results of Joshi and Fast (2013): participants in the high-power role discounted less than participants in the low-power role. However, when the assignments were perceived as high in legitimacy, the reverse was true: participants in the high-power role discounted more than participants in the low-power role. Because legitimacy was not experimentally manipulated, and was measured in only one of our studies, we consider this result suggestive but not conclusive. This legitimacy moderation effect is in line with the approach-inhibition theory of power and specifically with past research on how legitimacy moderates the relationship between power and behavioral approach tendencies. Lammers et al. (2008) found that when power was experienced as legitimate, high-power individuals displayed more approach than low-power individuals. However, when power was experienced as illegitimate, high–power individuals displayed the same degree of approach as, or even less approach than, low-power individuals, and such reduced approach tendencies have been associated with less impulsive behavior and greater self-control (Avila, 2001; Keltner et al., 2003; Schmeichel, Harmon-Jones, & Harmon-Jones, 2010).

Study 2

Participants in this study were randomly assigned to write about a situation where they either lacked power (low-power condition) or had power (high-power condition), or about their last trip to the grocery store (control condition). To measure temporal discounting, participants made eight choices between immediate improved air quality for 21 days and improved air quality for 35, 33, 31, 29, 27, 25, 23, or 21 days in one year. As in Joshi and Fast's (2013) Study 3, this measure was counterbalanced with the measure of connection with the future self. After the original procedure, we added a common manipulation check for the recall power manipulation (e.g., Smith, Jostmann, Galinsky, & van Dijk, 2008, Study 3; Tost, Gino, & Larrick, 2012, Experiment 3), asking participants how much power they had in the incident they recalled. The original study did not include a manipulation check. Experience with the tasks was again measured after the original procedure as a potential moderator but was not significant. We used the same discount rate calculation (the hyperbolic discounting formula, as in Study 1) and exclusion criteria as the original study.

The power manipulation was effective: high-power participants (M = 5.35, SD = 1.17) reported having more power in the incident they recalled than low-power participants (M = 2.75, SD = 1.26), $t(256) = 17.09$, $p < .001$, $d = 2.13$, 95% CI [-1.89, 2.38]. Control participants (M = 5.32, SD = 1.29) also reported having more power than low-power participants, $t(277) = -16.84$, $p < .001$, $d = 2.02$, 95% CI [1.78, 2.25]. The difference between high-power and control participants was not significant, $t(259) = 0.20$, $p = .840$, $d = 0.02$, 95% CI [-0.22, 0.27].

Table 1 reports the primary statistics for Study 2. A planned contrast comparing the discount rate of high-power participants with that of the average of low-power and control participants (as in Joshi & Fast (2013, Study 3)) showed no significant effect of condition on temporal discounting. Independent-samples t tests comparing the three conditions found that control participants discounted less than low-power participants, but all other comparisons were nonsignificant. According to an equivalence test (Lakens, 2017) focusing on the difference between the low- and high-power conditions, these data provide evidence for the null hypothesis of no effect (relative to that of an effect larger than d = 0.43), $t(249.06) = -2.96$, $p = .002$.

Consistent with Joshi and Fast (2013), a planned contrast showed that high-power participants (M = 3.85, SD = 1.45) reported being more connected with the future self than the average of low-power (M = 3.43, SD = 1.45) and control (M = 3.58, SD = 1.38) participants, $t(397) = 2.20$, $p = .030$, $\eta^2 = 0.01$, 95% CI [0.00, 0.03]. However, connection with the future self did not mediate the effect of power on temporal discounting, estimated indirect effect = 0.00, 95% bootstrapped confidence intervals of indirect effects [0.00, 0.00], $p = .960$. Thus, even though in Study 2 we found evidence for power affecting connection with the future self, which was Joshi and Fast's (2013) proposed mediator, we found no evidence for power affecting temporal discounting directly, nor for an indirect effect of power on temporal discounting via connection with the future self. Tost et al. (2015, Experiment 2) and Heller and Ulrich (2017) also did not find evidence for connection with the future self acting as a mediator.

Meta-analysis

We conducted a meta-analysis of experiments examining the effect of low versus high power on temporal discounting, which included our two close replications reported here, three additional replication studies of ours reported in the supplemental materials, the four previously published replications (Duan et al., 2017, Studies 1 and 3; Heller & Ullrich, 2017; Tost et al., 2015, Experiment 2), and the target studies of these replications (Joshi & Fast (2013), Studies 1 & 3). Correlational studies (e.g., Joshi & Fast, 2013, Study 4; Duan et al., 2017, Study 2) were not included since our goal was to assess the causal evidence for power affecting temporal discounting. Details of the meta-analysis including the full selection criteria are in the supplemental materials. Figure 1 shows the effect size for each experiment and the overall meta-effect in a forest plot. The overall meta-effect of power, calculated as the standardized effect size (Cohen's d) of the difference in discount rates between the low-power and high-power conditions, was -0.11, 95% CI [-0.25, 0.03]. Looking separately at the two discounting measures, the meta-effect was -0.10, 95% CI [-0.29, 0.08] within the monetary discounting experiments, and -0.15, 95% CI [-0.40, 0.10] within the air-quality discounting experiments. Thus, the evidence overall, as well as within each experimental design, is not consistent with an effect of power on temporal discounting.

The meta-analysis also showed a small to moderate amount of heterogeneity (I^2 = 47.19%) across experiments, Q_{10} = 20.01, p = .03, meaning heterogeneity accounts for 47.19% of the total variability in the data (Hamilton, 2017). The meta-effect was not moderated by whether temporal discounting was measured in the context of money versus air quality, Q_1 = 0.09, p = .77, nor did we identify any other moderators.

General Discussion

Joshi and Fast (2013) presented initial evidence that power reduces temporal discounting. With much larger samples, however, we found no effect of power on temporal discounting in two preregistered close replication studies. Using various methods of calculating temporal discounting and analysis strategies, including those used by the original authors, we never replicated Joshi and Fast's (2013) finding and never found a significant difference between low- and high-power conditions in temporal discounting. Furthermore, a meta-analysis of known replication studies and the target studies showed a nonsignificant effect of power on temporal discounting.

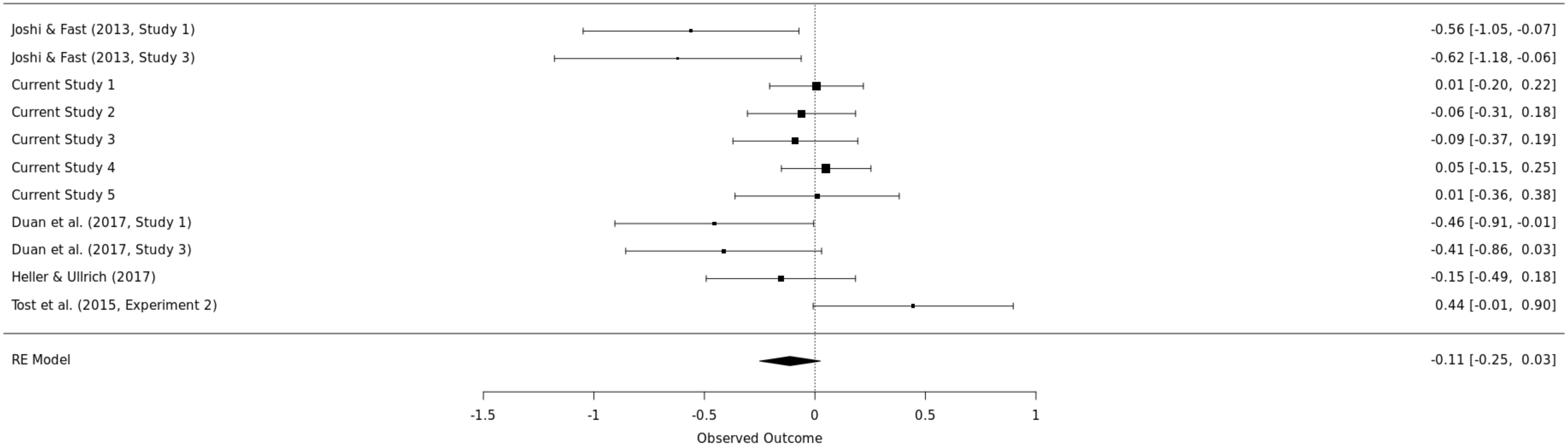

Figure 1. Forest plot of the random-effects meta-analysis. The plot is based on standardized effect sizes of the difference in discount rates between the low-power and high-power conditions. Bars indicate 95% confidence intervals. The size of the square indicates the weight of each study in the meta-analysis.

Why did we fail to find an effect of power on temporal discounting, unlike Joshi and Fast (2013)? One possibility is that our studies differed in small but critical ways from those of Joshi and Fast (2013), and these methodological differences led to our different results. Indeed, Studies 1 and 2 both involved multi-study, group-testing setups: participants completed our study as part of a series of unrelated studies while seated in individual cubicles in a common room. In contrast, participants in Study 3 of Joshi and Fast (2013) only completed that specific study during their session and apparently took part one person at a time. We tried to mitigate the potential influence of the multi-study setting by having our study always be the first one participants completed in each session. A priori, we also have no strong reason to believe that the power manipulations we used would not be successful in a setup such as we employed for Studies 1 and 2. Both the role-based manipulation in Study 1 and the recall manipulation in Study 2 have been successfully employed in previous research with multi-study and/or group-testing setups (e.g., Dubois, Rucker, & Galinsky, 2016, Experiment 3; Galinsky et al., 2003, Study 3; Garbinsky, Klesse, & Aaker, 2014, Experiment 1; Rucker, Dubois, & Galinsky, 2011, Experiment 1; Rus, van Knippenberg, & Wisse, 2010, Studies 1 & 2). Additionally, many studies using these manipulations have been conducted online via Amazon Mechanical Turk, including Study 1 of Joshi and Fast (2013); in these cases it is unknown whether other people were around while the participants did the study, and whether these participants were completing multiple online studies in a row.

However, it is possible that the effects of power manipulations are strongest in a single-study, individual-testing setup. This issue is especially important for researchers to consider as more social psychological research, including research on power, is conducted with online samples. As noted above, researchers have limited ability to control the environment surrounding such participants.

A second possibility is that we failed to manipulate power successfully. It is important to distinguish between failures to manipulate the construct of interest and failures to find an effect of the construct of interest, though even the former can be informative for researchers (e.g., Cheung et al., 2016; Finkel, 2016). For Study 1, the effectiveness of the role-based power manipulation was confirmed with the same manipulation check used by Joshi and Fast (2013, Study 1), in which participants report how much power they had in their role. Such manipulation checks are commonly used with role manipulations of power (e.g., Hildreth & Anderson, 2016, Studies 1a, 1b, 3, & 4; Mooijman, van Dijk, Ellemers, & van Dijk, 2015, Studies 1b & 4c).

For Study 2, though the original study did not include a manipulation check, we included a common manipulation check for the recall power prime, asking participants how much power they had in the episode they recalled (e.g., Smith et al., 2008, Study 3; Tost et al., 2012, Experiment 3). Low-power participants reported having less power in their recalled episode than high-power and control participants, but high-power participants did not report having more power than control participants. Since Joshi and Fast (2013) did not use a manipulation check in their Study 3, we do not know if this indicates any difference in effectiveness between our manipulation and theirs. Notably, this pattern of results does not rule out our ability to find effects of power on temporal discounting. Past researchers have found effects of power, including the recall power prime, on their critical dependent measures even when the control condition did not significantly differ from the low- and/or high-power conditions on a manipulation check (e.g., Schmid, Kleiman, & Amodio, 2015, Study 1). Furthermore, our manipulation check results do suggest that the recall prime successfully produced a power difference between low-power and high-power participants, which is the critical comparison. As discussed in the supplemental materials, even a targeted analysis comparing only the low- and high-power participants in Study 2 yields no effect of power on temporal discounting.

In Studies 1, 3, and 5 (see supplemental materials for the latter two studies), we also measured participants' General Sense of Power (Anderson et al., 2012). In Study 1, this measure came after several intervening studies to avoid any influence of the power manipulation on it. In Studies 3 and 5, participants completed it soon after the power manipulation. Similar to Tost et al. (2015, Experiment 2) and Heller and Ullrich (2017), we did not find any difference in participants' General Sense of Power between power conditions in any of these studies. Heller and Ullrich (2017) interpreted their null finding as a sign of an ineffective power manipulation. However, the version of the General Sense of Power scale employed in all these studies was designed to assess participants' trait-level personal sense of power, asking how much control and influence they had in their relationships with other people in general (Anderson et al., 2012). While the power manipulations employed in these studies may affect participants' momentary feelings of power, these manipulations are unlikely to affect participants' perception of their power and influence in all their social relationships. Indeed, other researchers have found significant effects of power manipulations on manipulation checks and key dependent measures when these manipulations did not affect participants' trait General Sense of Power (e.g., Anderson & Galinsky, 2006, Study 2; Tost et al., 2015, Experiment 2). In short, the General Sense of Power scale as administered in our Studies 1, 3, and 5 is not a manipulation check.

Meanwhile, we do share others' concern that the standard manipulation checks used in the power research literature may be subject to demand effects (e.g., Sturm & Antonakis, 2015). Because many commonly used power manipulations make it clear that the experiment has to do with power, participants may respond to the manipulation check with how they think they are supposed to feel, rather than with how powerful they actually feel. Though this topic is beyond the scope of the present replication attempts, as we were focused on conducting close replications of past work, future power research needs to grapple with this issue.

A third possible explanation for our failure to replicate Joshi and Fast (2013) is that our participants responded to the temporal discounting measure in unusual or extreme ways, or otherwise responded carelessly, interfering with the ability of our power manipulation to have an effect. We believe this is not the case. As can be seen in Table 1, the means and standard deviations for our conditions are very similar to those of Joshi and Fast (2013). Comparing our data to studies involving temporal discounting measures similar to ours (Hardisty et al., 2013; Hardisty & Weber, 2009; Heller & Ulrich, 2017; Tost et al., 2015, Experiment 2), we also confirmed that other aspects of our data (e.g., the percentage of participants who always preferred the immediate option) were not unusual (see supplemental materials for details). Researchers often report skewed discount rate distributions, as we did with Studies 1 and 2, and such skewness and outliers are dealt with in various ways, including dropping outliers, transforming data, and using non-parametric tests (e.g., Hardisty et al., 2013; Hardisty & Weber, 2009; Lempert, Glimcher, & Phelps, 2015; Tost et al., 2015). To ensure the robustness of our findings, we preregistered and reported analyses using multiple common ways of dealing with skewness and outliers in the supplemental material.

In addition, because discount rates are measured and calculated in a variety of ways in the temporal discounting literature, using responses to either titration questions or matching questions by themselves (e.g., Hardisty et al., 2013), or a combination of both (e.g., Hardisty & Weber, 2009), we preregistered and reported discount rate analyses using multiple methods of calculating discount rates in the supplemental material. Regardless of the method used, we found no evidence for an effect of our power manipulations on temporal discounting. We encourage future research on power and temporal discounting to be mindful of all the above methodological differences in eliciting and calculating temporal discounting to increase the robustness of future investigations.

Finally, though we did not replicate Joshi and Fast's (2013) finding that elevated power reduces temporal discounting via increasing connection with the future self, we want to highlight that our power manipulations did affect some key dependent measures. In Study 1, we found a significant interaction between power and participants' perceived legitimacy of the role manipulation on temporal discounting. The pattern was similar to Joshi and Fast's (2013) results when our participants felt their role assignment was not very legitimate, but it reversed when our participants felt the assignment was fairly legitimate. None of the other experiments on power and temporal discounting, either by us or other researchers, have measured or manipulated legitimacy, so this finding is suggestive but tentative. In Study 2, high-power participants reported greater connection with their future selves than low-power and control participants. However, similar to Tost et al. (2015, Experiment 2) and Heller and Ulrich (2017), we found no evidence for connection with the future self as a mediator of the effects of power on temporal discounting.

While we found little evidence for power reducing temporal discounting in our replication studies and in the meta-analysis, in other research power has been shown to affect some behaviors conceptually related to temporal discounting, such as saving (Garbinsky et al., 2014) and delaying consumption (May & Monga, 2014). Our finding of a null effect of power on temporal discounting does not necessarily imply that these other findings are less valid. These behaviors, though related to temporal discounting, are also conceptually distinct from it (e.g., they are not always correlated with temporal discounting: Urminsky & Zauberman, 2015) and are affected by multiple other mechanisms. For instance, Garbinsky et al. (2014) found that high-power individuals were more willing to save because they were motivated to maintain their power by accumulating wealth; this result suggests that power should only reduce temporal discounting for rewards seen as means for maintaining power. These distinctions in construct definitions and causal attributions distinguish these other findings from the effect of power on temporal discounting, highlighting the importance of conducting close replications as well as conceptual ones (Cesario, 2014; Simons, 2014).

Identifying whether and how power affects temporal discounting is important for theory testing because two prominent theories of power, the approach-inhibition theory (Keltner et al., 2003) and the social distance theory (Magee & Smith, 2013), make divergent predictions on this issue. It also has important real-world implications for understanding and improving intertemporal decision-making. There is already excitement in the policy area regarding the implications of Joshi and Fast's (2013) work (e.g., Patton et al., 2016; United Nations Development Programme, 2014). However, as shown in our meta-analysis, the cumulative data are not consistent with an effect of power on temporal discounting. We also found suggestive evidence for a moderation effect of perceived legitimacy in Study 1: relatively illegitimate power tended to decrease temporal discounting, similar to Joshi and Fast (2013), but relatively legitimate power tended to increase it. Future research should investigate this moderation effect with manipulated legitimacy to assess causality. In sum, more high-powered research testing the conditions under which power influences temporal discounting is needed before incorporating this effect into theory or practice.

References

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

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

Avila, C. (2001). Distinguishing BIS-mediated and BAS-mediated disinhibition mechanisms: A comparison of disinhibition models of Gray (1981, 1987) and of Patterson and Newman (1993). *Journal of Personality and Social Psychology*, *80*(2), 311–324. http://doi.org/10.1037/0022-3514.80.2.311

Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences.* Palgrave Macmillan.

Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science*, *9*(1), 40–48. http://doi.org/10.1177/1745691613513470

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers. *Behavior Research Methods*, *46*(1), 112–30. http://doi.org/10.3758/s13428-013-0365-7

Cheung, I., Campbell, L., LeBel, E. P., Ackerman, R. A., Aykutoğlu, B., Bahník, Š., … Yong, J. C. (2016). Registered Replication Report: Study 1 from Finkel, Rusbult, Kumashiro, & Hannon (2002). *Perspectives on Psychological Science*, *11*(5), 750–764. http://doi.org/10.1177/1745691616664694

Duan, J., Wu, S. J., & Sun, L. (2017). Do the powerful discount the future less? The effects of power on temporal discounting. *Frontiers in Psychology*, *8*(June), 1–11. http://doi.org/10.3389/fpsyg.2017.01007

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2016). Dynamics of communicator and audience power: the persuasiveness of competence versus warmth. *Journal of Consumer Research*, *43*(1), 68–85. http://doi.org/10.1093/jcr/ucw006

Finkel, E. J. (2016). Reflections on the commitment-forgiveness registered replication report. *Perspectives on Psychological Science*, *11*(5), 765–767. http://doi.org/10.1177/1745691616664695

Frederick, S., Loewenstein, G., & O’donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401. http://doi.org/10.1257/002205102320161311

Fujita, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal levels and self-control. *Journal of Personality and Social Psychology*, *90*(3), 351–67. http://doi.org/10.1037/0022-3514.90.3.351

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*(3), 453–66. http://doi.org/10.1037/0022-3514.85.3.453

Garbinsky, E. N., Klesse, A.-K., & Aaker, J. (2014). Money in the bank: Feeling powerful increases saving. *Journal of Consumer Research*, *41*(3), 610–623. http://doi.org/10.1086/676965

Gubler, T., & Pierce, L. (2014). Healthy, wealthy, and wise: retirement planning predicts employee health improvements. *Psychological Science*, *25*(9), 1822–30. http://doi.org/10.1177/0956797614540467

Hamilton, W. K.(2017). MAVIS: Meta Analysis via Shiny. R Package., version v1.1.3.

Hardisty, D. J., Thompson, K. F., Krantz, D. H., & Weber, E. U. (2013). How to measure time preferences: An experimental comparison of three methods. *Judgment and Decision Making*, *8*(3), 236–249. http://doi.org/10.1007/s10826-012-9600-6

Hardisty, D. J., & Weber, E. U. (2009). Discounting future green: Money versus the environment. *Journal of Experimental Psychology. General*, *138*(3), 329–340. http://doi.org/10.1037/a0016433

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2010). Individual discount rates and smoking: Evidence from a field experiment in Denmark. *Journal of Health Economics*, *29*(5), 708–717. http://doi.org/10.1016/j.jhealeco.2010.06.006

Heller, S., & Ullrich, J. (2017). Does power increase self-control? Episodic priming may not provide the answer. *Collabra: Psychology*, *3*(1), 3. http://doi.org/10.1525/collabra.48

Hildreth, J. A. D., & Anderson, C. (2016). Failure at the top: How power undermines collaborative performance. *Journal of Personality and Social Psychology*, *110*(2), 261–286. http://doi.org/10.1037/pspi0000045

Joshi, P. D., & Fast, N. J. (2013). Power and reduced temporal discounting. *Psychological Science*, *24*(4), 432–438. http://doi.org/10.1177/0956797612457950

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

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, *128*(1), 78–87. http://doi.org/10.1037/0096-3445.128.1.78

Lakens, D. (2017). Equivalence Tests. *Social Psychological and Personality Science*, *8*(4), 355–362. http://doi.org/10.1177/1948550617697177

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2008). Illegitimacy moderates the effects of power on approach. *Psychological Science*, *19*(6), 558–64. http://doi.org/10.1111/j.1467-9280.2008.02123.x

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2012). Power increases social distance. *Social Psychological and Personality Science*, *3*(3), 282–290. http://doi.org/10.1177/1948550611418679

Lempert, K. M., Glimcher, P. W., & Phelps, E. A. (2015). Emotional arousal and discount rate in intertemporal choice are reference dependent. *Journal of Experimental Psychology: General*, *144*(2), 366–373. http://doi.org/10.1037/xge0000047

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. http://doi.org/10.1080/19416520802211628

Magee, J. C., & Smith, P. K. (2013). The Social distance theory of power. *Personality and Social Psychology Review*, *17*(2), 158–186. http://doi.org/10.1177/1088868312472732

May, F., & Monga, A. (2014). When time has a will of its own, the powerless don’t have the will to wait: Anthropomorphism of time can decrease patience. *Journal of Consumer Research*, *40*(5), 924–942. http://doi.org/10.1086/673384

Mischel, W., Shoda, Y., & Rodriguez, M. (1989). Delay of gratification in children. *Science*, *244*(4907), 933–938. http://doi.org/10.1126/science.2658056

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. http://doi.org/10.1037/pspi0000021

Patton, G. C., Sawyer, S. M., Santelli, J. S., Ross, D. A., Afifi, R., Allen, N. B., … Viner, R. M. (2016). Our future: a Lancet commission on adolescent health and wellbeing. *The Lancet*, *387*(10036), 2423–2478. http://doi.org/10.1016/S0140-6736(16)00579-1

Pronin, E., Olivola, C. Y., & Kennedy, K. A. (2008). Doing unto future selves as you would do unto others: Psychological distance and decision making. *Personality and Social Psychology Bulletin*, *34*(2), 224–236. http://doi.org/10.1177/0146167207310023

Rucker, D. D., Dubois, D., & Galinsky, A. D. (2011). Generous paupers and stingy princes: Power drives consumer spending on self versus others. *Journal of Consumer Research*, *37*(6), 1015–1029. http://doi.org/10.1086/657162

Rus, D., van Knippenberg, D., & Wisse, B. (2010). Leader power and leader self-serving behavior: The role of effective leadership beliefs and performance information. *Journal of Experimental Social Psychology*, *46*(6), 929–933. http://doi.org/10.1016/j.jesp.2010.06.007

Schmeichel, B. J., Harmon-Jones, C., & Harmon-Jones, E. (2010). Exercising self-control increases approach motivation. *Journal of Personality and Social Psychology*, *99*(1), 162–73. http://doi.org/10.1037/a0019797

Schmid, P. C., Kleiman, T., & Amodio, D. M. (2015). Power effects on cognitive control: Turning conflict into action. *Journal of Experimental Psychology: General*, *144*(3), 655–663. http://doi.org/10.1037/xge0000068

Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science*, *9*(1), 76–80. http://doi.org/10.1177/1745691613514755

Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, *26*(5), 559–569. http://doi.org/10.1177/0956797614567341

Smith, P. K., Jost, J. T., & Vijay, R. (2008). Legitimacy crisis? Behavioral approach and inhibition when power differences are left unexplained. *Social Justice Research*, *21*(3), 358–376. http://doi.org/10.1007/s11211-008-0077-9

Smith, P. K., Jostmann, N. B., Galinsky, A. D., & van Dijk, W. W. (2008). Lacking power impairs executive functions. *Psychological Science*, *19*(5), 441–7. http://doi.org/10.1111/j.1467-9280.2008.02107.x

Sturm, R. E., & Antonakis, J. (2015). Interpersonal power: A review, critique, and research agenda. *Journal of Management*, *41*(1), 136–163. http://doi.org/10.1177/0149206314555769

Tost, L. P., Gino, F., & Larrick, R. P. (2012). Power, competitiveness, and advice taking: Why the powerful don't listen. *Organizational Behavior and Human Decision Processes*, *117*(1), 53–65. http://doi.org/10.1016/j.obhdp.2011.10.001

Tost, L. P., Wade-Benzoni, K. A., & Johnson, H. H. (2015). Noblesse oblige emerges (with time): Power enhances intergenerational beneficence. *Organizational Behavior and Human Decision Processes*, *128*, 61–73. http://doi.org/10.1016/j.obhdp.2015.03.003

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–63. http://doi.org/10.1037/a0018963

United Nations Development Programme. (2014). *Human development report 2014: Sustaining human progress: Reducing vulnerabilities and building resilience.* Retrieved from http://hdr.undp.org/sites/default/files/hdr14-report-en-1.pdf

Urminsky, O., & Zauberman, G. (2015). The psychology of intertemporal preferences. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making* (pp. 141–181). Chichester, UK: John Wiley & Sons, Ltd. http://doi.org/10.1002/9781118468333.ch5

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, *111*(4), 493–504. http://doi.org/10.1037/pspa0000056

Footnotes

[1]When discussing Tost et al. (2015, Experiment 2), we only report statistics regarding the personal discounting condition, which replicated Joshi and Fast's (2013) Study 1. They also had an intergenerational discounting condition, which is irrelevant for the current purpose.

[2]Preregistration documents, experimental materials, datafiles, and analysis scripts for Studies 1 and 2 are posted at https:// osf.io/gsv84 and https://osf.io/24mej, respectively. We also conducted three other studies which had some procedural differences from the original studies, two replicating Study 1 of Joshi and Fast (2013) and one replicating Study 3 of Joshi and Fast (2013). These studies also did not find any effect of power on temporal discounting. Their methods and results are reported in the supplemental material.

[3]The percentage of participants excluded in Study 1 is statistically higher than Joshi and Fast's (2013) Study 1, $\chi^2(1) = 4.01$, $p = .045$. This is likely caused by the additional attention check we included for the assigned role in the power manipulation (8% of participants were excluded for failing this attention check). The percentage of participants excluded did not differ significantly between Study 2 and Joshi and Fast's (2013) Study 3, $\chi^2(1) = 1.52$, $p = .217$.

[4]With $N = 67$ in the analysis, Joshi and Fast (2013, Study 1) had 33% power to detect an effect of $d = 0.38$. This is the smallest effect size we aim to detect with our replication, so we use this effect size as the equivalence bound for Study 1. Using the same rule, the equivalence bound for Study 2 is $d = 0.43$.

Author Contributions

M. Zhang designed the experiments and collected and analyzed data under the supervision of P. K. Smith. M. Zhang drafted the manuscript, and P.K. Smith provided critical revisions. Both authors approved the final version of the manuscript for submission.

Discount rates are calculated in a variety of ways in the temporal discounting literature, using responses to either titration questions or matching questions by themselves (e.g., Hardisty, Thompson, Krantz, & Weber, 2013; Jones & Rachlin, 2006; Read, Frederick, & Scholten, 2013), or a combination of both (e.g., Hardisty & Weber, 2009). Titration and matching questions can yield different discount rates due to the different psychological processes evoked by these different elicitation methods (Hardisty et al., 2013). The methodological limitations of one method can be remedied with the additional use of other methods. For example, titration discount rate measures are subject to ceiling effects if many participants have discount rates higher than the scale upper bound, but matching measures do not pose such restrictions. Thus, we preregistered and reported discount rate analyses for every study using multiple methods of calculating discount rates (e.g., using only choice titration, combining choice titration and matching measure).

Data skewness is common in temporal discounting studies, including in studies using comparable delay periods and reward amounts as our studies (e.g., Crockett, Clark, Lieberman, Tabibnia, & Robbins, 2010; Koff & Lucas, 2011; Lempert, Glimcher, & Phelps, 2015; Peters & Büchel, 2010; Wang & Dvorak, 2010). Researchers use data trimming (e.g., Tost, Wade-Benzoni, & Johnson, 2015, Experiment 2), data transformation (e.g., Lempert et al., 2015), or non-parametric statistical tests (e.g., Hardisty et al., 2013) to deal with skewed data and outliers. We pre-registered and reported results using all three methods. Reporting our results using these different methods also provides evidence regarding their robustness. All preregistration plans, materials, data, and analysis scripts can be found at osf.io/gsv84 (Study 1), osf.io/24mej (Study 2), osf.io/ze5ig (Studies 3 and 5), and osf.io/3yr59 (Study 4).

Study 1

Method

Participants. 420 undergraduate students completed the study and reported their demographic information (225 women, 195 men). Their average age was 21.30 years (*SD* = 3.10). Participants received course credit for their participation.

We excluded 58 participants from the main analyses per our preregistered criteria. 35 participants (21 in the low-power condition, 14 in the high-power condition) incorrectly remembered in the attention check to which role they had been assigned (worker versus manager). Power condition did not affect the likelihood of failing the attention check, $\chi^2(1) = 0.94$, $p = .332$. 12 participants (4 in the low-power condition, 8 in the high-power condition) showed no temporal discounting (i.e., preferred $120 or less in a year to $120 today). Power condition did not affect the likelihood of showing no discounting, $\chi^2(1) = 0.76$, $p = .384$. Five participants (3 in the low-power condition, 2 in the high-power condition) responded inconsistently in the titration temporal discounting measures (e.g., chose $137 in a year over $120 today but $120 today over $206 in a year). Six participants (3 in the low-power condition, 3 in the high-power condition) responded inconsistently between the titration and the matching measure (i.e., preferred $120 today over $240 in a year in the titration measure but reported an indifference point lower than $240 in the matching measure).

In addition, because the discount rate distribution in our experiment had a positive skew and some extreme outliers (skewness = 19.02, kurtosis = 361.94), we excluded 20 participants (10 in the low-power condition, 10 in the high-power condition) whose discount rates were more than 3 interquartile ranges from the median. Thus, the final sample size was 342 (187 women, 155 men; $M_{age} = 21.21$ years, $SD_{age} = 3.10$).

We used 3 interquartile ranges as an exclusion criterion because data points outside of 3 interquartile ranges are considered extreme values (Baguley, 2012, p. 312). We had pre-registered another common exclusion method, excluding data points beyond three standard deviations, to deal with outliers. However, this exclusion was not effective for our data because the standard deviation (SD = 525.50) was heavily inflated by one extreme outlier (indifference point = $120000 in a year for $120 today). After excluding discount rates more than three standard deviations from the mean, which meant excluding only the one extreme outlier mentioned above, the distribution was still highly skewed (skewness = 14.42, kurtosis = 226.25). The interquartile range exclusion method yields a discount rate distribution similar to the original paper (see Table 1 in manuscript), while keeping most participants in the analyses (4.76% exclusion rate). In addition to trimming outliers, we also report below two preregistered ways to deal with outliers: transformation and a nonparametric test.

Procedure. This study was run in two parts. The first part was the first in a series of studies that lasted for about an hour in total.[1] Participants completed the study in individual cubicles in a common room in groups of 5 to 25. The first part began with the same procedure as Joshi and Fast's (2013) Study 1, with material obtained from P. Joshi. First, participants' power level was manipulated with a virtual team task. Then their discount rate was measured with a titration measure and a matching measure. Finally, participants answered a manipulation-check question for the power manipulation.

After the procedure of the original study, we included additional measures. First participants completed an attention check followed by measures of perceived legitimacy and connection with the future self. Next, as part of the cover story for the virtual team power manipulation, low-power participants received either a general knowledge task or an anagram task to complete. High-power participants received information about three bogus team workers and assigned tasks to them. After finishing these tasks, participants answered questions about their previous experience with similar power manipulations and temporal discounting measures and completed measures of socioeconomic status, in that order. Finally, participants concluded the first part of the study by reporting what they thought the purpose of the study was and any suspicions they had about it.

Participants went on to complete other studies for most of the remainder of the hour. They then completed the second part of our study, which consisted of a measure of money goals and the General Sense of Power scale (Anderson, John, & Keltner, 2012). The order of these two measures was counterbalanced. We were interested in testing whether individual differences in money goals and the general sense of power would moderate the effect of power on discount rate. To avoid any influence of the power manipulation on participants' responses to these two measures, we separated the measures from the first part of the study.

Virtual team task. Participants were randomly assigned to one of two conditions: low power or high power. Participants were told they were participating in a study of virtual teams. All participants read the following text:

> We are interested in examining the functioning of virtual organizations and how individuals communicate with each other in groups and teams. You will only be permitted to communicate with your group members virtually. At specific intervals in the questionnaire, you may be given a chance to send messages to your group members via the internet. Each group will be consisting of 4 members. Each group will consist of 1 manager and 3 workers. The workers will follow directions given by the team manager. The manager will select tasks for the workers to perform.

Then, participants were informed about their assigned role as either a worker (low-power condition) or a manager (high-power condition). Participants learned that managers would evaluate team workers' performance. Workers, however, would not to evaluate the manager's performance. Participants were also told that 3 of the 50 participating groups with the best team performance would win a prize. The manager would choose how the money was distributed among their workers if their group won.

Temporal discounting measures. Participants completed two temporal discounting measures: a titration procedure and a free-response matching question. In the titration procedure, participants made nine choices between lottery prizes to be given out either on that day or in one year. Specifically, participants chose between receiving a $120 prize on that day and receiving a prize of $113, $120, $137, $154, $171, $189, $206, $223, and $240 in a year. All nine questions were presented on the same page. The free-response matching question was presented on a separate page and asked participants to fill in the number that made them indifferent between $120 today and $_____ in one year.

Power manipulation check. Participants reported the extent to which they had power over other group members on a 7-point scale (the points were unnumbered but labeled Strongly Disagree, Disagree, Somewhat Disagree, Neither Agree nor Disagree, Somewhat Agree, Agree, and Strongly Agree).

Attention check for power manipulation. Participants indicated whether they were assigned to the manager or the worker role in the virtual team task.

Perceived legitimacy of role assignment. Participants rated how legitimate they thought the role assignment was on a 7-point scale (1 = *Not at all*, 7 = *Very much*).

Connection with the future self. Participants reported how connected and how similar they felt to themselves in 10 years with a scale of overlapping circles representing the overlap between their current self and themselves in 10 years, adapted from the Inclusion of Other in the Self scale (Aron, Aron, & Smollan, 1992) and provided by Joshi and Fast.

Experience with similar tasks. Participants answered two separate questions about how many times they had done tasks similar to the virtual team task and the lottery choice task, in that order, with their answers restricted to a point estimate between 0 and 10000.

Socioeconomic status. Participants reported their demographic information, including three questions measuring their objective and subjective socioeconomic status (Anderson, Kraus, Galinsky, & Keltner, 2012). The first two questions asked participants to indicate their father's and mother's highest level of education; the provided categories were did not finish high school, high school graduate or general education diploma (GED), some college, college graduate, and postgraduate degree or degrees (e.g., Masters, PhD, JD, MD). The next question asked about their family's annual household income (before taxes); the provided categories were < $15,000; $15,001-$25,000; $25,001-$35,000; $35,001-$50,000; $50,001-$75,000; $75,001-$100,000; $100,001-$150,000; and > $150,000. The last question measured subjective SES. This measure consisted of a picture of a ladder with 10 rungs representing people with different levels of education, income, and occupation status. Participants were instructed to select the rung where they felt they stood relative to other people in the United States. Each rung of the ladder was given a number between 1 and 10, with higher numbers indicating higher placement on the ladder (Kraus & Keltner, 2009).

Goals related to money. Participants reported, in two separate questions, how important it was for them to earn money right now and to save money for the future, both on 7-point scales

(1 = *Not at all important*, 7 = *Very important*). The order of these two questions was counterbalanced.

General Sense of Power. Participants completed the General Sense of Power Scale (Anderson, John, et al., 2012).

Results

All analyses reported below were preregistered unless specifically noted.

Power manipulation check. High-power participants (M = 4.71, SD = 1.37) reported having more power over their team members than low-power participants (M = 3.61, SD = 1.41), $t(340) = 7.33$, p < .001, d = 0.79, 95% CI [0.57, 1.01].

Discount rate from choice titration supplemented with matching measure. P. Joshi informed us that in the original studies, participants' indifference points were calculated as the point at which they switched from preferring the present option to preferring the future option (personal communication, March 24, 2017). This is different from the calculation method specified in our preregistration, which took an average of the point of switching and the point right before switching as the indifference point. Both methods are commonly used in the temporal discounting literature (e.g., Hardisty et al., 2013; Hardisty & Weber, 2009; Kim, Schnall, & White, 2013; Weber et al., 2007). Using our preregistered calculation method does not change the pattern of results reported below. When participants chose $120 today over all future prizes (i.e., maxed out the titration scale; 34 participants, 18 in the low-power condition, 16 in the high-power condition, 9% of the final sample)[2], their answer in the free-response matching question was used as their indifference point, as was done in Joshi and Fast's (2013) Study 1 (P. Joshi, personal communication, March 24, 2017). The percentage of participants maxing out the titration scale did not differ between power conditions, $\chi^2(1) = 0.04$, p =.845.

For all analyses reported in the supplemental materials, each participant's discount rate was calculated as in the original study with a hyperbolic discounting formula $k = (A/V - 1)/D$, where A = the future amount that made participants indifferent between the present and the future reward, V = the present value ($120), and D = the delay (1 year).

A t-test on the trimmed discount rates showed no effect of power (M_{LP} = 0.43, SD_{LP} = 0.32; M_{HP} = 0.43, SD_{HP} = 0.29), $t(340) = -0.08$, $p = .940$, $d = -0.01$, 95% CI [-0.22, 0.20].

We also ran an equivalence test (Lakens, 2017) to test for evidence supporting a null effect, or no difference in discount rate between the low-power and high-power conditions. With $N = 67$ in their analyses, Joshi and Fast (2013, Study 1) had 33% power to detect an effect of d = 0.38. This is the smallest effect size we aim to detect with our replication. Thus, we set the equivalence bound as d = [-0.38, 0.38], meaning any effect within this range would be considered equivalent to a null effect. The equivalence test result provided evidence for the null hypothesis of no effect, $t(336.85) = -3.42$, $p < 0.001$.

We also preregistered and used two other methods to deal with the skewness of discount rates. In the first method, we performed the three most common transformations on the discount rates: square root transformation (after transformation skewness = 18.69, kurtosis = 353.24), log transformation (after transformation skewness = 8.35, kurtosis = 98.80), and inverse transformation (after transformation skewness = 1.17, kurtosis = 1.46). We used the inverse transformation in our analyses because it brought the discount rates closest to a normal distribution, which has a skewness of 0 and a kurtosis of 3. A t-test on the inverse-transformed discount rates showed no effect of power (M_{LP} = 1.22, SD_{LP} = 0.22; M_{HP} = 1.22, SD_{HP} = 0.20), $t(360) = 0.02$, $p = .981$, $d = 0.00$, 95% CI [-0.20, 0.21].

In the second method, we ran a Wilcoxon rank sum nonparametric test on the untransformed discount rates, which again showed no effect of power (Mdn_{LP} = 0.28, Mdn_{HP} = 0.43), W = 15999, p = .697.

Below we report discount rate analyses based on other calculations of discount rates: 1) the discount rate from the choice titration alone and 2) the discount rate from the matching measure.

Discount rate from only choice titration. This analysis was not preregistered. For participants who maxed out the titration scale (i.e., always chose $120 today), we followed the method used for Joshi and Fast's (2013) Study 3 (P. Joshi, personal communication, March 24, 2017), which assumes that maxed-out participants would switch over to choose the future option if the titration scale were extended one step further. This is a common way of dealing with maxed-out participants (e.g., Jones & Rachlin, 2006; Kim, Schnall, & White, 2013). Thus, maxed-out participants were assigned an indifference point of $257. Consistent with the analyses reported above, power did not affect the discount rate calculated with this indifference point (M_{LP} = 0.47, SD_{LP} = 0.34; M_{HP} = 0.48, SD_{HP} = 0.32), $t(366)$ = 0.13, p = .899, d = 0.01, 95% CI [-0.19, 0.22].

Discount rate from matching measure. Participants' responses to the matching discount measure were also used independently to calculate discount rates. 72 participants showing no temporal discounting (i.e., reporting an indifference point lower than or equal to $120 on the free-response measure) were excluded from this analysis. 313 participants remained (160 women, 153 men, M_{age} = 21.18, SD_{age} = 2.75).

Because the matching-based discount rate distribution was positively skewed (skewness = 13.98, kurtosis = 206.57), we performed the standard transformations: square root transformation (after transformation skewness = 11.90, kurtosis = 148.45), log transformation (after transformation skewness = 7.17, kurtosis = 60.52), and inverse transformation (after transformation skewness = 0.99, kurtosis = 0.55). We used the inverse transformation in our analyses because it brought the discount rates closest to a normal distribution. A t-test showed no effect of power on the matching-based discount rate, (M_{LP} = 1.31, SD_{LP} = 0.24; M_{HP} = 1.30, SD_{HP} = 0.22), t(360) = -0.31, p = .753, d = 0.04, 95% CI [-0.19, 0.26]. A Wilcoxon rank sum nonparametric test on the untransformed matching-based discount rates also showed no effect of power (Mdn_{LP} = 0.33, Mdn_{HP} = 0.42), W = 12281.00, p = .966.

In addition to the above analyses on discount rates, we analyzed our data using another standard measure of temporal discounting, the proportion of times participants chose the immediate reward ($120 today) in the titration procedure (Read, Frederick, & Scholten, 2013). On average, low-power participants chose $120 today over future rewards in 48% (SD = 26%) of their choices, and high-power participants also chose $120 today in 48% ($SD$ = 25%) of their choices. A logistic regression showed no significant difference between conditions, z(366) = 0.20, p = .844, OR = 1.01, 95% CI[0.88, 1.16].

In sum, we did not find an effect of power on temporal discounting with any of the discount rate calculations, or with a proportion of choice analysis.

Connection with the future self. The ratings of “connected” and “similar” were averaged for an index of connection with the future self (α = 0.53). There was no difference in connection to the future self between high-power (M = 3.75, SD = 1.27) and low-power (M = 3.81, SD = 1.29) conditions, t(340) = 0.49, p = .625, d = 0.05, 95% CI [-0.20, 0.34].

We conducted mediation analyses with connection with the future self as the mediator of the effect of power on temporal discounting. All mediation analyses reported in this paper were conducted with the *mediation* package in R (Tingley, Yamamoto, Hirose, Keele, & Imai, 2014). All simulations used 5,000 samples, after which quasi-Bayesian confidence intervals were calculated (Imai, Keele, & Tingley, 2010). Power condition (coded as 0 = low power, 1 = high power, for all regression analyses for Study 1) did not predict temporal discounting, $\beta = -0.00$, $t(340) = -0.08$, $p = .940$, 95% CI [-0.07, 0.06], or connection with the future self, $\beta = 0.07$, $t(340) = 0.49$, $p = .625$, 95% CI [-0.20, 0.34]. Connection with the future self did not predict temporal discounting, $\beta = 0.11$, $t(340) = 0.69$, $p = .490$, 95% CI [-0.02, 0.04]. When connection with the future self and power were both included in the model, neither predicted temporal discounting: for connection, $\beta = 0.01$, $t(339) = 0.69$, $p = .489$, 95% CI [-0.02, 0.04]; for power condition, $\beta = -0.00$, $t(339) = -0.09$, $p = .926$. There was no evidence for mediation, estimated indirect effect = -.00, 95% bootstrapped confidence intervals of indirect effects [-0.00, 0.00], $p = .980$.

Potential moderators. We report full results using the trimmed titration-plus-matching discount rates because is the original study also used titration-plus-matching discount rates. At the end of each analysis, we also summarize the results based on the other three calculations of discount rates: 1) the inverse-transformed discount rate from choice titration supplemented with matching measure (from here on, referred to as the inverse-transformed titration-plus-matching discount rate); 2) the titration-based discount rate; and 3) the inverse-transformed matching-based discount rate (from here on, referred to as the matching-based discount rate). The full results with these three other discount rate calculations can be found at osf.io/gsv84. Since we do not have any reason to believe one discount rate calculation method is superior to the others, we base our conclusions on results consistent across different calculation methods and take caution in interpreting any inconsistent results.

Perceived legitimacy of role assignment. Perceived legitimacy of role assignment might moderate the effect of power on temporal discounting, through moderating the effect of power on either approach tendency or social distance (Lammers, Galinsky, Gordijn, & Otten, 2008; see also Smith, Jost, & Vijay, 2008). Lammers et al. (2008) showed that high power led to more approach and less inhibition than low power when it was experienced as legitimate. However, when power was perceived as illegitimate, high power led to either similar levels of approach and inhibition as low power, or sometimes less approach and more inhibition (i.e., the opposite pattern to legitimate power). Less approach and more inhibition are associated with less impulsive, disinhibited behavior and greater self-control (Avila, 2001; Keltner, Gruenfeld, & Anderson, 2003; Schmeichel, Harmon-Jones, & Harmon-Jones, 2010). Therefore, according to the approach-inhibition theory of power, legitimate power should increase temporal discounting while illegitimate power may not affect temporal discounting, or may even reduce it.

Legitimacy also moderates the effect of power on social distance: legitimate power increases social distance, whereas illegitimate power decreases social distance (Lammers, Galinsky, Gordijn, & Otten, 2012). Increased social distance leads to higher construal level (Magee & Smith, 2013), which reduces temporal discounting (Fujita & Carnevale, 2012). Thus, the social distance theory of power (Magee & Smith, 2013) would predict that legitimate power should reduce temporal discounting while illegitimate power should increase it.

To test these competing hypotheses, we regressed discount rate on perceived legitimacy, power condition, and their interaction. We report standardized regression coefficients for this and all other regression results. The main effects of perceived legitimacy, $\beta = -0.03$, 95% CI [-0.09, 0.16], $t(338) = -1.34$, $p = .181$, and power, $\beta = -0.00$, 95% CI [-0.07, 0.06], $t(338) = -0.09$, $p = .931$, were not significant. There was a significant interaction, $\beta = 0.06$, 95% CI [0.01, 0.14],

t(338) = 2.24, p = .026. We next conducted an exploratory simple slope analysis to test the effect of power at low (-1 SD) versus high (+1 SD) levels of perceived legitimacy. When perceived legitimacy was high, power did not affect discount rate significantly, β = 0.07, t(338) = 1.53, p = .127. When perceived legitimacy was low, high-power participants discounted marginally less than low-power participants, β = - 0.08, t(338) = -1.67, p = .096.

It is important to note that low-power and high-power participants did not differ in how legitimate they perceived their assigned role to be (M_{LP} = 4.31, SD_{LP} = 1.19; M_{HP} = 4.48, SD_{HP} = 1.46), t(340) = 1.21, p = .227, d = 0.13, 95% CI [-0.08, 0.34], although this test was not pre-registered.

The perceived legitimacy and power condition interaction effect was significant for all the other discount rate calculations as well. The direction of the simple slope effects was always the same as reported above: power tended to increase temporal discounting when perceived legitimacy was high and reduce temporal discounting when perceived legitimacy is low. However, the significance of these simple slopes varied between different discount rate calculations. Both simple slopes reach statistical significance for the inverse-transformed titration-plus-matching discount rate, but neither reached statistical significance for the titration-based discount rate. For the matching-based discount rate, only the simple slope for low perceived legitimacy reached statistical significance.

Experience with similar tasks. Participants estimated they had done tasks similar to the virtual team task 0.26 times (SD = 0.94) on average. To test whether experience with this task moderated the effect of power on discount rate, we regressed participants' discount rate on their experience with the virtual team task, power condition, and their interaction. There was no main effect of experience, β = -0.01, 95% CI [-0.05, 0.04], t(338) = -0.26, p = .796; no main effect of power condition, $\beta = 0.00$, 95% CI [-0.07, 0.06], $t(338) = -0.02$, $p = .985$; and no interaction, $\beta = 0.04$, 95% CI [-0.02, 0.11], $t(338) = 1.28$, $p = .201$. We also recoded participants' experience with the virtual team task variable into a binary variable (0 = never did similar tasks, 1 = did similar tasks at least once before). 296 participants (86.55%) reported they had never done tasks similar to the virtual team task. A 2 (experience with virtual team task) by 2 (power condition) between-subjects ANOVA on discount rates showed no main effect of experience, $F(1, 338) = 0.22$, $p = .643$, η2 = .001; no main effect of power condition, $F(1, 338) = 0.01$, $p = .940$, $\eta^2 = .00$; and no interaction, $F(1, 338) = 0.02$, $p = .878$, $\eta^2 = .00$.

Participants estimated they had done tasks similar to the titration temporal discounting measure 1.00 time ($SD = 1.85$) on average. Regression analysis showed no main effect of participants' experience, $\beta = 0.01$, 95% CI [-0.04, 0.07], $t(338) = 0.53$, $p = .599$; no main effect of power condition, $\beta = 0.00$, 95% CI [-0.07, 0.06], $t(338) = -0.07$, $p = .946$; and no interaction, $\beta = -0.01$, 95% CI [-0.08, 0.06], $t(338) = -0.30$, $p = .768$. 189 participants (55.26%) reported they had never done tasks similar to the titration temporal discounting measure. A 2 (experience with titration temporal discounting task) by 2 (power condition) between-subjects ANOVA on discount rates showed no main effect of experience, $F(1, 338) = 2.17$, $p = .141$, $\eta^2 = .006$; no main effect of power condition, $F(1, 338) = 0.01$, $p = .940$, $\eta^2 = .000$; and no interaction, $F(1, 338) = 0.27$, $p = .605$, $\eta^2 = .001$.

Analyses based on other forms of discount rate calculation also showed null results. There was no evidence that participants’ experience with either of these tasks moderated the effect of power on temporal discounting.

Objective and subjective socioeconomic status. Following prior work (e.g., Kraus, Piff, & Keltner, 2009), parental education ratings for both parents were assigned a number from 1 to

5, with higher numbers indicating greater educational attainment. Family income was assigned a number from 1 to 8, with higher numbers indicating greater household income. Family household income, maternal education, and paternal education were standardized separately and then summed together to create a composite measure of objective SES.

We regressed discount rate on objective SES, power condition, and their interaction. There was no main effect of objective SES, $\beta = 0.01$, 95% CI [-0.04, 0.05], $t(338) = 0.34$, $p = .737$; no main effect of power condition, $\beta = 0.00$, 95% CI [-0.07, 0.06], $t(338) = -0.04$, $p = .970$; and no interaction effect, $\beta = -0.04$, 95% CI [-0.11, 0.02], $t(338) = -1.29$, $p = .199$.

We ran a similar analysis with subjective SES. There was no main effect of subjective SES, $\beta = 0.01$, 95% CI [-0.04, 0.05], $t(338) = 0.33$, $p = .740$; no main effect of power condition, $\beta = 0.00$, 95% CI [-0.07, 0.06], $t(338) = -0.05$, $p = .956$, and no interaction effect, $\beta = 0.01$, 95% CI [-0.05, 0.08], $t(338) = 0.35$, $p = .726$.

Analyses based on other forms of discount rate calculation also showed null results. Neither objective nor subjective SES moderated the effect of power on temporal discounting.

Goals related to money. Although not pre-registered, we tested whether power affected participants' goals related to money. Power did not affect the importance of earning money right now ($M_{LP} = 5.42$, $SD_{LP} = 1.52$; $M_{HP} = 5.25$, $SD_{HP} = 1.56$), $t(340) = -1.00$, $p = .320$, $d = 0.11$, 95% CI [-0.32, 0.11]. Eight participants did not report the importance of saving money for the future, so they were excluded from analyses of the importance of saving money for the future. Power did not affect the importance of saving money for the future ($M_{LP} = 6.06$, $SD_{LP} = 1.12$; $M_{HP} = 6.15$, $SD_{HP} = 1.06$), $t(332) = 0.79$, $p = .430$, $d = 0.09$, 95% CI [-0.13, 0.30]. We also examined the difference in importance between future saving goals and immediate earning goals (i.e., the response to the earning goals question was subtracted from the response to the saving goals question). Power did not affect this difference (*M_{LP}* = 0.65, *SD_{LP}* = 1.46; *M_{HP}* = 0.90, *SD_{HP}* = 1.53), $t(332) = 1.55$, $p = .120$, $d = 0.17$, 95% CI [-0.05, 0.39].

Having established that power did not affect the importance of money goals, we ran three regression analyses to test the moderating effect of goals related to money. First, we regressed discount rate on the importance of earning money right now, power condition, and their interaction. The regression analysis showed no effect of the earning goal, β = 0.02, 95% CI [-0.02, 0.07], $t(338) = 1.06$, $p = .291$, no effect of power condition; β = 0.00, 95% CI [-0.06, 0.06], $t(338) = 0.00$, $p = .999$; and no interaction effect, β = 0.00, 95% CI [-0.07, 0.06], $t(338) = -0.11$, $p = .911$.

Second, we regressed discount rate on the importance of saving money for the future, power condition, and their interaction. The regression analysis showed that the more important the saving goal was to participants, the less they discounted, β = -0.05, 95% CI [-0.09, 0.00], $t(330) = -2.04$, $p = .042$. There was no effect of power condition, β = 0.00, 95% CI [-0.06, 0.07], $t(330) = 0.12$, $p = .905$; and no significant interaction, β = 0.06, 95% CI [0.00, 0.13], $t(330) = 1.85$, $p = .065$.

Third, we regressed discount rate on the difference in goal importance, power condition, and their interaction. There was a significant main effect of the difference in goal importance: participants who considered saving goals more important relative to earning goals discounted less, β = -0.06, 95% CI [-0.11, -0.01], $t(330) = -2.45$, $p = .015$. There was no effect of power condition; β = 0.01, 95% CI [-0.06, 0.07], $t(330) = 0.26$, $p = .797$; and no significant interaction, β = 0.05, 95% CI [-0.02, 0.11], $t(330) = 1.45$, $p = .148$.

The main effect of the saving goal and the main effect of the difference in goal importance reported above were not significant when analyzed with any other form of discount rate calculations. There was no moderation effect found with any other discount rate calculations.

General Sense of Power. Four participants did not complete the General Sense of Power scale and so were excluded from these analyses. First we tested whether power condition affected participants' sense of power (SOP). We predicted that power condition would not have an effect on SOP because SOP assesses a general sense of power rather than state-level feelings of power, and past research has not found an effect of temporary power manipulations on this measure (e.g., Anderson & Galinsky, 2006). Furthermore, the SOP measure was separated from the power manipulation by at least 20 minutes of other studies, so we would expect any effects of the manipulation to have dissipated by then. Indeed, a t-test showed no effect of power condition on SOP, (M_{LP} = 4.98, SD_{LP} = 0.79; M_{HP} = 5.00, SD_{HP} = 0.86), $t(336) = 0.27$, $p = .790$, $d = 0.03$, 95% CI [-0.18, 0.24].

Having shown that our manipulation did not affect SOP, we then tested whether SOP moderated the effect of power on temporal discounting. We regressed discount rate on SOP, power condition, and their interaction. Participants with a higher general sense of power had higher discount rates, $\beta = 0.08$, 95% CI [0.03, 0.12], $t(334) = 3.27$, $p = .001$. There was no effect of power condition, $\beta = 0.00$, 95% CI [-0.06, 0.07], $t(334) = 0.09$, $p = .928$; and no interaction effect, $\beta = -0.03$, 95% CI [-0.10, 0.03], $t(334) = -1.04$, $p = .298$.

The SOP main effect did not reach statistical significance with any other discount rate calculations. There was no moderation effect of SOP with any other discount rate calculation.

Gender. We also conducted an exploratory analysis to test whether gender moderated the effect of power on temporal discounting. Joshi and Fast (2013, Footnote 1) reported that participant gender did not affect discount rate in their studies. A 2 (participant gender) by 2 (power condition) between-subjects ANOVA on discount rates showed no effect of gender, $F(1, 338) = 0.45$, $p = .503$, $\eta^2 = .001$; no effect of power condition, $F(1, 338) = 0.01$, $p = .922$, $\eta^2 = .000$; and no interaction effect, $F(1, 338) = 0.22$, $p = .637$, $\eta^2 = .001$.

Analysis of the inverse-transformed titration-plus-matching discount rate showed a significant main effect of gender: men discounted more than women. But none of the other discount rate analyses showed a significant effect of gender. There was no moderation effect of gender with any discount rate calculation.

To summarize, we found that perceived legitimacy of the role assignment moderated the effect of power on temporal discounting consistently across different data analyses strategies. Power increased temporal discounting when perceived legitimacy was higher and decreased it when perceived legitimacy was lower. None of the other potential moderators we tested (experience with tasks, subjective and objective SES, goals related to money, SOP, gender) showed evidence for moderation.

Study 2

Method

Participants. 463 undergraduate students (250 women, 213 men) completed the study. Their average age was 20.97 years (*SD* = 2.39). Participants received course credit for participating in this experiment.

We excluded 64 participants from our main analyses per our preregistered criteria. Six participants (1 in the low-power condition, 1 in the control condition, 4 in the high-power condition) were excluded because they did not complete the titration temporal discounting measure. 20 participants responded inconsistently in the titration temporal discounting measure. There was no significant difference between power conditions in responding inconsistently (6 in the low-power condition, 3 in the control condition, 11 in the high-power condition), χ^2 (2) = 4.98, p = .083. 38 participants (19 in the low-power condition, 15 in the control condition, 23 in the high-power condition) showed no temporal discounting (i.e., preferring improved air quality for 21 days or less in a year to improved air quality for 21 days immediately) in the titration measure. Power condition did not affect the likelihood of showing no discounting, χ^2 (2) = 1.42, p =.491. Thus, the final sample size was 399 (211 women, 188 men, M_{age} = 20.94 years, SD_{age} = 2.43).

Procedure. This study was the first in a series of studies that lasted for about an hour in total. Participants completed the study in individual cubicles in a common room in groups of 5 to 25. Participants first completed the procedure of Study 3 in Joshi and Fast (2013). Power was manipulated through an episodic recall task (Galinsky, Gruenfeld, & Magee, 2003). Temporal discounting was measured with an environmental preference survey. Participants also completed the measure of connection with the future self. The order of the connection with future self measure and the temporal discounting measure was counterbalanced as in the original study. The order of these two measures did not have any effect on temporal discounting or connection with the future self, so it will not be discussed further. After the original procedure, participants completed a manipulation check question for the power manipulation and two questions about participants' experience with the experimental tasks.

Power manipulation. Following Joshi and Fast (2013) and Galinsky et al. (2003), participants were randomly assigned to one of three conditions: low power, control, or high power. Low-power participants wrote about a situation when they lacked power, whereas high-power participants wrote about a situation when they had power. Control participants wrote about their last trip to the grocery store.

Connection with the future self. The same connection with the future self measure was used as in Study 1.

Temporal discounting measures. Participants were given the following instructions:

> Imagine that your county department is considering a temporary change in its emission policy to study the effects of air quality on human health and local wildlife. In order to study the effects of air quality, the particulate output of nearby factories and power plants would be immediately reduced for a period of three weeks, after which time the air quality would return to its former level, but the government is also considering making the change 1 year in the future, for a different length of time.

The measure consisted of a titration procedure and a free-response matching question. In the titration procedure, participants made eight choices between immediate improved air quality for 21 days and improved air quality one year from now for 35, 33, 31, 29, 27, 25, 23, and 21 days. All choices were presented on the same page, in the above order. The free-response matching question was presented on a separate page. It asked participants to fill in the number that made the following two options equally attractive: "Improved air quality immediately for 21 days or Improved air quality one year from now for __________days."

Manipulation check. After the procedure of the original experiment described above, participants reported how much power they had in the incident they recalled on a 7-point scale (1 = *Very little*, 7 = *A lot*).

Experience with similar tasks. Participants answered two separate questions about how many times they had done tasks similar to the recall task and the discounting task, in that order, with their answers restricted to a point estimate between 0 and 10000.

Results

All analyses reported below were preregistered unless specifically noted.

Manipulation check. A one-way ANOVA showed a significant difference between power conditions in self-reported power in the recalled incident, $F(2, 396) = 194.82$, $p < .001$, $\eta^2 =.50$ [0.43, 0.55]. Independent samples t-tests showed that high-power participants ($M = 5.35$, $SD = 1.17$) reported having more power than low-power participants ($M = 2.75$, $SD = 1.26$), $t(256) = 17.09$, p < .001, $d = 2.13$, 95% CI [-1.89, 2.38]. Control participants ($M = 5.32$, $SD = 1.29$) also reported having more power than low-power participants, $t(277) = -16.84$, p < .001, $d = 2.02$, 95% CI [1.78, 2.25]. The difference between high-power and control participants was not significant, $t(259) = 0.20$, $p = .840$, $d = 0.02$, 95% CI [-0.22, 0.27].

Discount rate from only choice titration. Participants' indifference points were calculated as in Study 1. Using our preregistered indifference point calculation method does not change the pattern of results reported below. For participants who maxed out the titration scale (i.e., always chose 21 days of immediate air improvement; 211 participants, 81 in the low-power condition, 64 in the control condition, 66 in the high-power condition, 53% of final sample), we followed the method used in Joshi and Fast's (2013) Study 3 and extended the titration scale one step further, assigning them an indifference point of 37 days (P. Joshi, personal communication, March 24, 2017). This is a common way of dealing with maxed-out participants (e.g., Jones & Rachlin, 2006). The percentage of participants maxing out the titration scale did not differ between power conditions, $\chi^2 (2) = 2.46$, $p =.293$.

A planned contrast analysis comparing high-power participants ($M = 0.55$, $SD = 0.27$) to the average of low-power ($M = 0.57$, $SD = 0.26$) and control participants ($M = 0.49$, $SD = 0.28$) was not significant, $t(397) = 0.76$, $p = .450$, $\eta2 = 0.00$, 95% CI [0.00, 0.02]. However, a one-way ANOVA showed a significant effect of power condition, $F(2, 396) = 3.09$, $p = .047$, $\eta^2 = .015$.

Independent samples t-tests showed that low-power participants discounted more than control participants, $t(277) = 2.35$, $p = .020$, $d = 0.28$, 95% CI [0.05, 0.52], whereas high-power participants did not differ from low-power participants, $t(256) = -0.48$, $p = .630$, $d = 0.06$, 95% CI [-0.18, 0.30], or control participants, $t(259) = 1.77$, $p = .080$, $d = 0.22$, 95% CI [-0.03, 0.47]. This pattern differs from the results of Joshi and Fast's (2013) Study 3, in which high-power participants discounted less than both low-power and control participants, but the latter two groups did not differ.

We also ran an equivalence test (Lakens, 2017) to test for evidence supporting a null effect between the low-power and high-power conditions. With $N = 52$ in the low-power and high-power conditions, Joshi and Fast (2013, Study 3) had 33% power to detect an effect of $d = 0.43$. This is the smallest effect size we aim to detect with our replication. Thus, we set the equivalence bound as $d = [-0.43, 0.43]$, meaning any effect within this range would be considered equivalent to a null effect. The equivalence test result provided evidence for the null hypothesis of no effect, $t(249.06) = -2.96$, $p = .002$.

Below we report results from other calculations of discount rates: 1) the discount rate from choice titration supplemented with the matching measure and 2) the discount rate from the matching measure.

Discount rate from choice titration supplemented with matching measure. When combining the titration and matching measures, we followed the method used in Study 1: when participants maxed out the titration scale (i.e., always chose immediate air quality improvement for 21 days to all future options), their answers to the free-response matching question were used as their indifference point. Among these participants, three reported "infinity" as their indifference point and 60 reported an indifference point of fewer than 35 days in a year (21 in the low-power condition, 20 in the control condition, 19 in the high-power condition), which is inconsistent with their titration response. We excluded these 63 participants in addition to the 64 participants excluded from the titration procedure. Thus, 336 participants remained in the following analyses (176 women, 160 men; M_{age} = 20.89 years, SD_{age} = 2.33).

The discount rate distribution had a highly positive skew and some extreme outliers (skewness = 18.65, kurtosis = 348.00). As in Study 1, we report three methods used to address this issue. In the first method, we excluded 42 participants (15 in the low-power condition, 12 in the control condition, 15 in the high-power condition) whose discount rates were more than 3 interquartile ranges from the median. The likelihood of exclusion did not differ between power conditions, $\chi^2(2) = 1.34$, $p = .51$. Thus, 294 participants were used in this analysis (149 women, 145 men; M_{age} = 20.88, SD_{age} = 2.36). A one-way ANOVA did not show a significant effect of power (M_{LP} = 0.78, SD_{LP} = 0.69; M_{Ctr} = 0.62, SD_{Ctr} = 0.64; M_{HP} = 0.81, SD_{LP} = 0.81), $F(2, 291) = 2.08$, $p = .126$, $\eta^2 = .014$.

In the second method, we followed our preregistered analysis plan and performed the three most common transformations on the discount rates: square root transformation (after transformation skewness = 18.65, kurtosis = 347.93), log transformation (after transformation skewness = 8.76, kurtosis = 103.03), and inverse transformation (after transformation skewness = 0.48, kurtosis = -0.87). We used the inverse transformation in our analyses because it brought the discount rates closest to a normal distribution. The one-way ANOVA was not significant (M_{LP} = 1.40, SD_{LP} = 0.30; M_{Ctr} = 1.32, SD_{Ctr} = 0.30; M_{HP} = 1.40, SD_{LP} = 0.32), $F(2, 333) = 2.69$, $p = .070$, $\eta^2 = .016$.

In the third method, we ran a Kruskal-Wallis rank sum nonparametric test on the untransformed discount rates and found the effect of power condition was not significant (Mdn_{LP} = 0.90, Mdn_{Ctr} = 0.48, Mdn_{HP} = 0.67), χ^2 (2) = 5.94, p = .051.

Discount rate from matching measure. Four participants who did not provide a numerical indifference point and 126 participants (41 in the low-power condition, 46 in the control condition, 39 in the high-power condition) showing no temporal discounting (i.e., reporting an indifference point less than or equal to 21 days on the free-response measure) were excluded from this analysis. The proportion of no discount participants did not differ between power conditions, χ^2 (2) = 0.50, p = .779. Thus, 333 participants remained (174 women, 159 men; M_{age} = 20.86 years, SD_{age} = 2.02).

The discount rate distribution had a highly positive skew and some extreme outliers (skewness = 18.22, kurtosis = 332.20), so we followed our preregistered analysis plan and performed the square root transformation (after transformation skewness = 18.24, kurtosis = 333.20), log transformation (after transformation skewness = 6.97, kurtosis = 60.21), and inverse transformation (after transformation skewness = 0.08, kurtosis = -0.83). We used the inverse transformation in our analyses since it brought the discount rates closest to a normal distribution. A one-way ANOVA showed a significant effect of power condition, F(2, 330) = 4.71, p = .010, η^2 = .028. Independent samples t-tests showed that control participants (M = 1.44, SD = 0.28) discounted less than low-power participants (M = 1.51, SD = 0.27), t(221) = 1.85, p = .070, d = 0.25, 95% CI [-0.02, 0.51], and high-power participants (M = 1.55, SD = 0.29), t(219) = 3.06, p < .001, d = 0.41, 95% CI [0.14, 0.67], but the latter two groups did not differ , t(222) = 1.30, p = .200, d = 0.17, 95% CI [-0.10, 0.43].

A Kruskal-Wallis rank sum test on the untransformed discount rates also showed a significant effect of power condition (Mdn_{LP} = 1.00, Mdn_{Ctr} = 0.48, Mdn_{HP} = 0.90), χ^2 (2) = 8.93, *p* = .011. Follow-up exploratory Wilcoxon rank sum tests showed the same pattern of results as the titration only discount rate analyses: control participants discounted less than low-power participants, *W* = 5254, *p* = .050, and high-power participants, *W* = 7399, *p* = .004, and the latter two groups did not differ, *W* = 6745, *p* = .270.

In addition to the above analyses on discount rates, we analyzed the proportion of times participants chose the immediate reward (21 days of improved air quality now) in the titration procedure. A logistic regression model showed an overall significant effect of power condition, χ^2 (2) = 2.56, *p* < .001. Control participants chose the immediate reward (*M* = 64%, *SD* = 37%) less often then low-power participants (*M* = 74%, *SD* = 34%), *z*(277) = -5.11, *p* < .001, *OR* = 0.62, 95% CI[0.52, 0.75], and high-power participants (*M* = 72%, *SD* = 35%), *z*(259) = -3.86, *p* < .001, *OR* = 0.69, 95% CI[0.57, 0.83]. There was no difference between low-power and high-power participants, *z*(256) = -1.06, *p* = .288, *OR* = 0.90, 95% CI[0.74, 1.09].

In sum, across different discount rate calculations and statistical tests, high-power and low-power participants did not differ in temporal discounting. Control participants discounted less than the other two groups, though these differences were not always statistically significant.

Connection with the future self. The ratings of "connected" and "similar" were averaged for an index of connection with the future self (α = 0.76). A planned contrast analysis showed that high-power participants (*M* = 3.85, *SD* = 1.45) were more connected to their future self than the average of low-power (*M* = 3.43, *SD* = 1.45) and control participants (*M* = 3.58, *SD* = 1.38), *t*(397) = 2.20, *p* = .030, η^2 = 0.01, 95% CI [0.00, 0.03]. However, a one-way ANOVA showed that the effect of power condition did not reach significance, *F*(2, 396) = 2.77, *p* = .064,

$\eta^2 = .014$. Even though the omnibus test was nonsignificant, we ran exploratory independent samples t-tests to investigate whether any power conditions differed significantly. High-power participants ($M = 3.85$, $SD = 1.45$) reported being more connected to their future self than low-power participants ($M = 3.43$, $SD = 1.45$), $t(256) = 2.30$, $p = .020$, $d = 0.29$, 95% CI [0.04, 0.53]. Control participants ($M = 3.58$, $SD = 1.38$) did not differ from low-power participants, $t(277) = -0.87$, $p = .390$, $d = 0.10$, 95% CI [-0.13, 0.34], or high-power participants, $t(259) = 1.53$, $p = .130$, $d = 0.19$, 95% CI [-0.06, 0.43].

We also conducted mediation analyses with connection with the future self as the mediator of the effect of power on temporal discounting. For this and all other regression analyses for Study 2, power condition was coded as 1 = low power, 2 = control, and 3 = high power. Power condition did not predict temporal discounting, $\beta = -0.01$, 95% CI [-0.04, 0.02], $t(397) = -0.57$, $p = .571$. Connection with the future self also did not predict temporal discounting, $\beta = 0.00$, 95% CI [-0.02, 0.02], $t(397) = -0.20$, $p = .844$. When both were included in the regression model, neither power, $\beta = -0.01$, 95% CI [-0.04, 0.02], $t(396) = -0.55$, $p = .584$, nor connection, $\beta = 0.00$, 95% CI [-0.02, 0.02], $t(396) = -0.13$, $p = .895$, predicted temporal discounting. There was no evidence for mediation, estimated indirect effect = 0.00, 95% bootstrapped confidence intervals of indirect effects [-0.00, 0.00], $p = .960$.

Though we found some evidence that high-power participants felt the most connected with their future self in Study 2, we did not find this effect in the other two studies (Studies 1 and 5) in which we measured connection with the future self. Combined with the null effects in Tost et al. (2015, Experiment 2) and Heller and Ulrich (2017), there is limited evidence for an effect of power on connection with the future self.

Potential moderators. We report full results using the titration only discount rates as this was the discount rate calculation method used in the original study. At the end of each analysis, we also summarize the results based on the other three calculations of discount rates: 1) the discount rate from choice titration supplemented with matching measure with three interquartile range exclusion (from here on, referred to as the trimmed titration-plus-matching discount rate) 2) the inverse-transformed discount rate from choice titration supplemented with matching measure (from here on, referred to as the inverse-transformed titration-plus-matching discount rate), 3) inverse-transformed matching-based discount rate (from here on, referred to as the matching-based discount rate). The full results with these three other discount rate calculations can be found at osf.io/24mej. Again, we base our conclusions on results consistent across different calculation methods.

Experience with similar tasks. Two participants did not answer the experience questions and so were excluded from these analyses. The remaining participants estimated they had done tasks similar to the recall manipulation 1.38 times (*SD* = 5.69) on average. To test whether experience with this task moderated the effect of power on discount rate, we regressed participants' discount rate on their experience with the recall manipulation, power condition, and their interaction. There was no main effect of experience, $\beta = -0.05$, 95% CI [-0.19, 0.08], $t(393) = -0.76$, $p = .448$; no main effect of power condition, $\beta = -0.01$, 95% CI [-0.04, 0.03], $t(393) = -0.47$, $p = .638$; and no interaction, $\beta = 0.03$, 95% CI [-0.03, 0.10], $t(393) = 0.96$, $p = .337$. We also recoded participants' experience with the recall manipulation variable into a binary variable (0 = never did similar tasks, 1 = did similar tasks at least once before). 261 participants (65.74%) reported they had never done tasks similar to the recall manipulation. A 2 (experience with recall manipulation) by 3 (power condition) between-subjects ANOVA on discount rates showed no main effect of experience, $F(1, 391) = 1.50$, $p = .222$, $\eta^2 = .004$; a significant main effect of power condition, $F(2, 391) = 3.12$, $p = .045$, $\eta^2 = .016$; and no interaction, $F(2, 391) = 0.78$, $p = .459$, $\eta^2 = .004$.

Participants estimated they had done tasks similar to the titration temporal discounting measure 0.23 times ($SD = 0.89$) on average. Regression analysis showed no main effect of participants' experience, $\beta = -0.05$, 95% CI [-0.16, 0.06], $t(393) = -0.88$, $p = .377$; no main effect of power condition, $\beta = -0.01$, 95% CI [-0.04, 0.03], $t(393) = -0.47$, $p = .642$; and no interaction, $\beta = 0.02$, 95% CI [-0.07, 0.02], $t(393) = -1.03$, $p = .303$. 355 participants (89.42%) reported they had never done tasks similar to the titration temporal discounting measure. A 2 (experience with titration temporal discounting task) by 3 (power condition) between-subjects ANOVA on discount rates showed no main effect of experience, $F(1, 391) = 0.16$, $p = .690$, $\eta^2 = .000$; a significant main effect of power condition, $F(2, 391) = 3.12$, $p = .045$, $\eta^2 = .016$; and no interaction, $F(2, 391) = 1.45$, $p = .236$, $\eta^2 = .007$.

Consistent with the results reported above, analyses with trimmed titration-plus-matching discount rates did not show any main effect or moderation effect of experience with tasks on temporal discounting. Analyses with the inverse-transformed titration-plus-matching discount rates and the inverse-transformed matching-based discount rates both showed a main of experience with the recall task: participants who had done tasks similar to the recall task before discounted more than participants who had never done similar tasks. Analyses with these two discount rate calculations also showed a significant interaction between experience with the recall task (as a continuous variable) and power condition. Exploratory simple slope analyses showed that power increased temporal discounting for participants who had more experience with the recall task but did not affect temporal discounting for participants who had less experience with the recall task. Since this interaction was not replicated with other methods of calculating discount rates or when treating experience as a dichotomous variable, we hesitate to draw conclusions from these two significant results.

Overall, we do not find clear evidence for moderation effects of experience with either the recall task or the temporal discounting measure.

Gender. A 2 (participant gender) by 3 (power condition) between-subjects ANOVA on discount rates showed no effect of gender, $F(1, 393) = 2.00$, $p = .158$, $\eta^2 = .005$; a significant main effect of power condition, $F(2, 393) = 3.14$, $p = .044$, $\eta^2 = .016$; and no interaction effect, $F(2, 393) = 0.36$, $p = .695$, $\eta^2 = .002$.

Analysis of the inverse-transformed titration-plus-matching discount rate showed a significant main effect of gender: men discounted less than women. However, this effect is in the opposite direction to the significant gender effect with inverse-transformed titration-plus-matching discount rates found in Study 1. Furthermore, no other significant effects were found in analyses with other discount rate calculations. There was no moderation effect of gender with any discount rate calculation.

To summarize, we tested experience with tasks and gender in Study 2 as potential moderators, but we did not find any consistent moderation effects.

Additional replication studies

Below, we report three additional replication studies. Studies 3 and 4 attempted to replicate Study 1 in Joshi and Fast (2013). Study 5 was an attempt to replicate Study 3 in Joshi and Fast (2013). We learned after conducting these studies and submitting our manuscript (P. Joshi, personal communications, September 29, 2016; October 5, 2016) that the procedures used in these three studies (based on two Qualtrics surveys provided by P. Joshi) differed from the original studies, so these studies are not discussed in the main manuscript.

Study 3

Method

Participants. 224 workers on Amazon Mechanical Turk (Mturk) gave their consent to participate in the study. We recruited only US workers with a Mturk HIT approval rate greater than or equal to 95%. 13 workers (5.80%) did not complete the study. Four dropped out before seeing the power manipulation. Seven participants in the low-power condition and two participants in the high-power condition dropped out after being assigned to their roles. Power condition was unrelated to dropout rate, $\chi^2(1) = 1.85$, $p = .173$. Thus, 211 participants completed the study (68 women, 143 men; $M_{age} = 32.05$, $SD_{age} = 9.81$). Each participant received a payment of $0.50 for completing the study.

We excluded 18 participants from the analyses per our preregistered criteria. 16 participants (5 in the low-power condition, 11 in the high-power condition) incorrectly remembered in the attention check to which role they had been assigned (worker versus manager). Power condition did not affect the likelihood of failing the attention check, $\chi^2(1) = 1.44$, $p = .229$. Two participants (1 in the low-power condition, 1 in the high-power condition) responded inconsistently in the titration temporal discounting measure. Thus, the final sample size was 193 (63 women, 130 men; $M_{age} = 32.62$, $SD_{age} = 9.98$).

Procedure. Following the Qualtrics Survey provided by P. Joshi (personal correspondence, November 13, 2014), participants in Study 3 were asked to complete the study when they were alone and not in a group situation. Study 3 began with the same procedure as Joshi and Fast's (2013) Study 1. First, participants' power level was manipulated with a virtual team task. Then their discount rate was measured with a titration measure and a matching measure. The matching measure was different from the one used in the original study, as detailed below.

Participants next completed two more sets of items included in the Qualtrics survey provided by P. Joshi but not reported in the original paper. First, participants completed the General Sense of Power scale (Anderson, John, et al., 2012). Then, they rated nine items about the virtual team task, including one manipulation check for the power manipulation and other items assessing their motivation to engage in the virtual team task and their perceived fit with their assigned roles.

Virtual team task. Power was manipulated with the same virtual team task as in Study 1.

Temporal discounting measures. Participants completed a titration procedure and a free-response matching question. In the titration procedure, participants were told that one participant would actually receive one of their choices, so they should choose carefully. These were the instructions used in the Qualtrics survey provided by P. Joshi, but they do not match Joshi and Fast's (2013) description of the methods for their Study 1.

Next participants completed a matching temporal discounting measure included in the Qualtrics survey provided by P. Joshi. Participants filled in the amount of money they would need to receive in 1, 5, 10, 20, and 30 years to be equivalent to receiving $100 tomorrow. This matching measure is different from the one used in original study (and thus from what we used in Study 1). Because the smaller, sooner reward in this matching measure ($100 tomorrow) is different from the one used in the titration measure ($120 today), discount rates using the matching response to supplement the titration response could not be calculated. Thus, we only report the titration-based and matching-based discount rates.

General Sense of Power. Participants completed the General Sense of Power Scale (Anderson, John, et al., 2012).

Power manipulation check. They also rated nine items about the virtual team task on 7-point scales (the points were unnumbered but labeled Strongly Disagree, Disagree, Somewhat Disagree, Neither Agree nor Disagree, Somewhat Agree, Agree, and Strongly Agree). One of these items was the manipulation check for the virtual team power manipulation: “I have power over my group members.”

Motivation to engage in the virtual team task. Three of these items measured participants’ motivation to engage in the virtual team task: “I look forward to working with my group members,” “I am motivated to see my group succeed,” and “I want my group to win $40.”

Perceived role fit. Two of these items measured participants’ perceived fit with their assigned role: “I was happy with the role I was given” and “I am capable of meeting the requirements of my role in this group task.”

The rest of the items were: “I like my other group members,” “My group members will perform to the best of their ability,” and “My group members have power over me.” We will not discuss participants’ responses to the last three items as they are irrelevant for our purpose.

At the end of the procedure, we also included the attention check for the power manipulation and the experience with similar tasks questions as in Study 1.

Results

Power manipulation check. High-power participants (M = 5.28, SD = 1.05) reported having more power over their team members than low-power participants (M = 3.45, SD = 1.50), $t(191) = 7.79$, $p < .001$, d = 1.41, 95% CI [1.09, 1.73].

Discount rate from only choice titration. We used the same calculation for discount rate from only choice titration as in Study 1. 79 participants (41 in the low-power condition, 38 in the high-power condition, 41% of the final sample) maxed out the titration scale. The percentage of participants maxing out the titration scale did not differ between power conditions, χ^2 (1) = 0.05, p =.816. Power did not affect the discount rate (M_{LP} = 0.81, SD_{LP} = 0.37; M_{HP} = 0.78, SD_{HP} = 0.35), t(191) = -0.62, p = .539, d = 0.09, 95% CI [-0.37, 0.20].

We also ran an equivalence test (Lakens, 2017) to test for evidence supporting a null effect, meaning no difference in discount rate between the low-power and high-power conditions. Using the same equivalence bound as Study 1, the equivalence test result provided evidence for the null hypothesis of no effect, t(190.10) = -2.00, p = .023.

Discount rate for matching measure. Matching-based discount rates were calculated as in Study 1. In addition to excluding the 16 participants who failed the attention check, we excluded 17 participants (8 in the low-power condition, 9 in the high-power condition) because they showed no temporal discounting (i.e., reporting an indifference point lower than or equal to $100). We also excluded one participant who reported an unusually high indifference point of 10^{64} for 30 years' delay, which was 10^{55} times larger than any other indifference point reported. For the remaining 177 participants, we averaged together their discount rates for the five time delays (Hardisty et al., 2013). An independent samples t-test on the average matching-based discount rates showed no effect of power (M_{LP} = 47.01, SD_{LP} = 13.16; M_{HP} = 44.29, SD_{HP} = 10.43), t(174) = -1.52, p = .129, d = -0.23, 95% CI [-0.53, 0.07].

An analysis of the proportion of times participants chose the immediate reward ($120 today) in the titration procedure also showed no difference between power conditions (M_{LP} =

74%, SD_{LP} = 27%, M_{HP} = 72%, SD_{HP} = 29%), $z(191) = -1.16$, $p = .247$, $OR = 0.88$, 95% CI[0.71, 1.09].

In sum, although we were unable to use the same method for calculating discount rates as the original study, we did not find any effect of power on temporal discounting with these other calculation methods, or with a proportion of choices analysis.

Potential moderators. We report the full moderator analyses results with the titration-based discount rates and summarize the results with the matching-based discount rates. The full results with the matching-based calculation can be found at https://osf.io/ze5ig/.

Experience with similar tasks. Participants estimated they had done tasks similar to the virtual team task 4.70 times ($SD = 17.45$) on average. For this and all other regression analyses for Study 3, power condition was coded as 0 = low power, 1 = high power. Regression analysis showed no main effect of experience, $\beta = 0.02$, 95% CI [-0.04, 0.07], $t(189) = 0.66$, $p = .511$; no main effect of power condition,$\beta = -0.04$, 95% CI [-0.15, 0.06], $t(189) = -0.80$, $p = .427$; and no interaction, $\beta = -0.12$, 95% CI [-0.31, 0.07], $t(189) = -1.27$, $p = .207$. We also recoded participants' experience with the virtual team task variable into a binary variable (0 = never did similar tasks, 1 = did similar tasks at least once before). 105 participants (54.40%) reported they had never done tasks similar to the virtual team task. A 2 (experience with virtual team task) by 2 (power condition) between-subjects ANOVA on discount rates showed no main effect of experience, $F(1, 189) = 0.66$, $p = .417$, $\eta^2 = .003$; no main effect of power condition, $F(1, 189) = 0.38$, $p = .540$, $\eta^2 = .002$; and no interaction, $F(1, 189) = 0.19$, $p = .667$, $\eta^2 = .001$.

Participants estimated they had done tasks similar to the titration temporal discounting measure 6.41 times ($SD = 18.39$) on average. Regression analysis showed no main effect of participants' experience, $\beta = 0.03$, 95% CI [-0.04, 0.11], $t(189) = 0.85$, $p = .398$; no main effect of power condition, β = -0.03, 95% CI [-0.14, 0.07], $t(189) = -0.65$, $p = .519$; and no interaction, β = -0.03, 95% CI [-0.13, 0.08], $t(189) = -0.52$, $p = .600$. 111 participants (57.51%) reported they had never done tasks similar to the titration temporal discounting measure. A 2 (experience with titration temporal discounting task) by 2 (power condition) between-subjects ANOVA on discount rates showed no main effect of experience, $F(1, 189) = 0.72$, $p = .398$, $\eta^2 = .004$; no main effect of power condition, $F(1, 189) = 0.38$, $p = .540$, $\eta^2 = .002$; and no interaction, $F(1, 189) = 0.26$, $p = .613$, $\eta^2 = .001$.

Analyses with the matching-based discount rates also showed no effect of experience with either task.

General Sense of Power. First we tested whether power condition affected participants' sense of power (SOP). A t-test showed no effect of power condition, ($M_{LP} = 4.80$, $SD_{LP} = 1.09$; $M_{HP} = 4.77$, $SD_{HP} = 1.00$), $t(191) = -0.15$, $p = .880$, $d = 0.02$, 95% CI [-0.31, 0.26].

Having shown that our manipulation did not affect SOP, we then tested whether SOP moderated the effect of power on temporal discounting. We regressed discount rate on SOP, power condition, and their interaction. The effect of SOP was not significant, β = 0.03, 95% CI [-0.04, 0.09], $t(189) = 0.71$, $p = .480$. There was no effect of power condition, β = -0.03, 95% CI [-0.13, 0.07], $t(189) = -0.61$, $p = .545$; and no interaction effect, β = -0.02, 95% CI [-0.12, 0.09], $t(189) = -0.35$, $p = .729$.

Analyses with the matching-based discount rates also showed no effects of SOP.

Motivation to engage in the virtual team task. We averaged ratings on the three motivation items to create a motivation index (α = 0.71). Power condition did not affect participants' motivation to engage in the virtual team task ($M_{LP} = 5.73$, $SD_{LP} = 0.80$; $M_{HP} = 5.81$, $SD_{HP} = 0.81$), $t(191) = 0.67$, $p = .510$, $d = 0.10$, 95% CI [-0.19, 0.38]. We also examined whether participants' motivation moderated the effect of power on discount rate. We regressed participants' discount rate on participants' motivation, power condition, and their interaction. The effect of motivation was not significant, $\beta = 0.02$, 95% CI [-0.06, 0.09], $t(189) = 0.44$, $p = .661$. There was no effect of power condition, $\beta = -0.03$, 95% CI [-0.14, 0.07], $t(189) = -0.66$, $p = .513$; and no interaction effect, $\beta = 0.01$, 95% CI [-0.09, 0.12], $t(189) = 0.26$, $p = .797$.

Analyses with the matching-based discount rates also showed no effects of motivation to engage in the virtual team task.

Perceived role fit. We averaged ratings on the two perceived role fit items to create an index of perceived role fit ($\alpha = 0.59$). High-power participants ($M = 5.69$, $SD = 1.06$) perceived greater role fit than low-power participants in the low power condition ($M = 5.33$, $SD = 0.97$), $t(191) = 2.48$, $p = .010$, $d = 0.36$, 95% CI [0.07, 0.64]. Then, we regressed participants' discount rate on participants' perceived role fit, power condition, and their interaction. The effect of perceived role fit was not significant, $\beta = 0.05$, 95% CI [-0.03, 0.13], $t(189) = 1.23$, $p = .219$. There was no effect of power condition, $\beta = -0.05$, 95% CI [-0.15, 0.06], $t(189) = -0.90$, $p = .371$; and no interaction effect, $\beta = -0.01$, 95% CI [-0.11, 0.09], $t(189) = -0.18$, $p = .854$.

Analyses with the matching-based discount rates also showed no effects of perceived role fit. However, since power affected perceived role fit, this is not a good test of fit's moderating effect.

Gender. A 2 (participant gender) by 2 (power condition) between-subjects ANOVA on discount rates showed no effect of gender, $F(1, 189) = 1.09$, $p = .299$, $\eta^2 = .006$; no effect of power condition, $F(1, 189) = 0.29$, $p = .589$, $\eta^2 = .002$; and no interaction effect, $F(1, 189) = 1.36$, $p = .244$, $\eta^2 = .007$.

Analyses with the matching-based discount rates also showed no effects of gender.

To summarize, we tested experience with tasks, SOP, motivation to engage in the virtual team task, perceived role fit, and gender as potential moderators, but did not find any moderation effects in Study 3.

Study 4

Method

Participants. 424 Mturk workers gave their consent to participate in the study. Again, we recruited only US workers with a Mturk HIT approval rate greater than or equal to 95%. 20 workers (4.72%) did not complete the study. Four dropped out before seeing the power manipulation. 6 participants in the low-power condition and 10 participants in the high-power condition dropped out after being assigned to their power roles. Power condition was unrelated to dropout rate, $\chi^2(1) = 0.58$, $p = .444$. Thus, 404 workers (185 women, 219 men; $M_{age} = 34.25$, $SD_{age} = 10.49$) completed the study. Each received a payment of $1 for completing the study.

We excluded 29 participants from the analyses per our preregistered criteria. 18 participants (9 in the low-power condition, 9 in the high-power condition) incorrectly remembered in the attention check to which role they had been assigned (worker versus manager). 9 participants (2 in the low-power condition, 7 in the high-power condition) responded inconsistently in the titration temporal discounting measure. 2 participants (1 in the low-power condition, 1 in the high-power condition) showed no temporal discounting. Thus, the final sample size was 375 (178 women, 197 men; $M_{age} = 34.44$, $SD_{age} = 10.43$).

Procedure. Procedures were the same as Study 3 except for the following changes. First, we modified the instructions for the titration temporal discounting measure so the choices were for hypothetical lotteries, using the instructions from Hardisty and Weber (2009) as described in the methods for Study 1 in Joshi and Fast’s (2013) paper. Second, we did not include the matching temporal discounting measure, the general sense of power scale, or questions about the virtual team task other than the manipulation check question about participants' level of power. Finally, after the question about level of power, we measured participants' optimism about getting the prize if they won the hypothetical lottery described in the temporal discounting measure. To measure this optimism, participants rated the following three items (adapted from Duan, Wu, & Sun, 2017) on 7-point scales (the points were unnumbered but labeled Strongly Disagree, Disagree, Somewhat Disagree, Neither Agree nor Disagree, Somewhat Agree, Agree, and Strongly Agree): "Although there is some risk of not getting the lottery prize in a year, I have the ability to deal with it," "I am optimistic that I will get the lottery prize in a year if my survey was chosen," and "I do not think there is an unexpected situation where I cannot get the lottery prize if my survey was chosen." The average of these three items was used as an index of optimism ($\alpha = 0.53$). Duan, Wu, and Sun (2017, Study 1) found that optimism about getting the prize mediated the effect of power on temporal discounting for Chinese participants, but the mediation effect was specific to participants of Han ethnicity, not participants of Tibetan ethnicity (Study 3). We thus included it as a potential mediator.

Results

Power manipulation check. High-power participants ($M = 5.89$, $SD = 1.05$) reported having more power over their team members than low-power participants ($M = 2.82$, $SD = 2.08$), $t(373) = 23.07$, $p < .001$, $d = 2.38$, 95% CI [2.12, 2.65].

Discount rate from only choice titration. Since we did not include the matching measure, we were only able to calculate a discount rate using choice titration alone with the same method as in Studies 1 and 3. 123 participants (60 in the low-power condition, 63 in the high-power condition, 33% of the final sample) maxed out the titration scale. The percentage of participants maxing out the titration scale did not differ between power conditions, χ^2 (1) = 0.13, p =.721. Power did not affect the discount rate (M_{LP} = 0.75, SD_{LP} = 0.35; M_{HP} = 0.77, SD_{HP} = 0.35), t(373) = .50, p = .618, d = 0.05, 95% CI [-0.15, 0.25].

We also ran an equivalence test (Lakens, 2017) to test for evidence supporting a null effect between the low-power and high-power conditions. Using the same equivalence bound as Study 1, the equivalence test result provided evidence for the null hypothesis of no effect, t(371.80) = 3.15, p < .001.

An analysis of the proportion of times participants chose the immediate reward ($120 today) in the titration procedure also showed no difference between power conditions (M_{LP} = 69%, SD_{LP} = 27%, M_{HP} = 71%, SD_{HP} = 27%), z(373) = 0.89, p = .374, OR = 1.07, 95% CI [0.92, 1.24].

Experience with similar tasks. Two participants did not answer the experience questions and thus were excluded from these analyses. The rest of the participants estimated they had done tasks similar to the virtual team task 3.90 times (SD = 31.24) on average. For this and all other regression analyses for Study 3, power condition was coded as 0 = low power, 1 = high power. Regression analysis showed no main effect of experience, β = 0.01, 95% CI [-0.03, 0.05], t(369) = 0.46, p = .645; no main effect of power condition,β = 0.01, 95% CI [-0.06, 0.09], t(369) = 0.38, p = .705; and no interaction, β = 0.01, 95% CI [-0.08, 0.10], t(369) = 0.27, p = .790. We also recoded participants' experience with the virtual team task variable into a binary variable (0 = never did similar tasks, 1 = did similar tasks at least once before). 301 participants (80.70%) reported they had never done tasks similar to the virtual team task. A 2 (experience with virtual team task) by 2 (power condition) between-subjects ANOVA on discount rates showed no main effect of experience, $F(1, 369) = 0.51$, $p = .475$, $\eta^2 = .001$; no main effect of power condition, $F(1, 369) = 0.13$, $p = .723$, $\eta^2 = .000$; and no interaction, $F(1, 369) = 2.08$, $p = .150$, $\eta^2 = .006$.

Participants estimated they had done tasks similar to the titration temporal discounting measure 8.95 times ($SD = 56.90$) on average. Regression analysis showed no main effect of participants' experience, $\beta = 0.01$, 95% CI [-0.03, 0.05], $t(369) = 0.58$, $p = .565$; no main effect of power condition, $\beta = 0.02$, 95% CI [-0.05, 0.09], $t(369) = 0.50$, $p = .616$; and no interaction, $\beta = 0.06$, 95% CI [-0.09, 0.20], $t(369) = 0.75$, $p = .455$. 240 participants (64.34%) reported they had never done tasks similar to the titration temporal discounting measure. A 2 (experience with titration temporal discounting task) by 2 (power condition) between-subjects ANOVA on discount rates showed a main effect of experience. Participants who had done the titration discount measure at least once ($M = 0.81$, $SD = 0.33$) discounted more than participants who had never done the titration discount measure ($M = 0.73$, $SD = 0.36$), $F(1, 369) = 4.51$, $p = .034$, $\eta^2 = .012$. There was no effect of power condition, $F(1, 369) = 0.13$, $p = .722$, $\eta^2 = .000$; and no interaction, $F(1, 369) = 0.48$, $p = .490$, $\eta^2 = .001$.

Gender. A 2 (participant gender) by 2 (power condition) between-subjects ANOVA on discount rates showed that the main effect of gender was not significant, $F(1, 371) = 2.07$, $p = .151$, $\eta^2 = .006$. There was no effect of power condition, $F(1, 371) = 0.12$, $p = .729$, $\eta^2 = .000$; and no interaction effect, $F(1, 371) = 0.69$, $p = .406$, $\eta^2 = .002$.

Optimism. Power did not affect participants' optimism about getting the prize if they won the lottery described in the temporal discounting measure ($M_{LP} = 4.90$, $SD_{LP} = 1.06$; $M_{HP} = 4.83$, $SD_{HP} = 1.05$), $t(373) = -0.64$, $p = .520$, $d = 0.07$, 95% CI [-0.13, 0.27]. Participants who were more optimistic about getting the prize when they win the lottery discounted less, $\beta = -0.09$, 95% CI [-0.14, -0.04], $t(371) = -3.54$, $p < .001$. We ran a mediation analysis testing if optimism mediated the effect of power on temporal discounting. There was no evidence for mediation, estimated indirect effect = 0.00, 95% bootstrapped confidence intervals of indirect effects [-0.01, 0.02], p = .500.

Past research on power and optimism has found that power increased optimism regarding general life events (Anderson & Galinsky, 2006) and time estimates (Weick & Guinote, 2010), but these are very different from our measure of participants' optimism about getting the prize if they won the lottery described in the temporal discounting measure. It is possible that because the lottery was hypothetical, the optimism measure, as it focused on getting the lottery, was less meaningful for participants. Furthermore, the optimism measure came at the very end of the study, when the power manipulation would presumably be weakest. In Duan, Wu, and Sun (2017, Studies 1 and 3), optimism was measured either immediately after the power manipulation or as the second dependent measure (i.e., right after the temporal discounting measure).

Study 5

Method

Participants. 199 undergraduate students (93 women, 106 men) completed the study. Their average age was 21.21 years (SD = 1.56). Participants received course credit for participating in this experiment.

We excluded 26 participants from our main analyses per our preregistered criteria. 10 participants (5 in the low-power condition, 2 in the control condition, 3 in the high-power condition) responded inconsistently in the titration temporal discounting measure. Power condition did not affect the likelihood of responding inconsistently, χ^2 (2) = 1.44, p =.514. 16 participants (2 in the low-power condition, 4 in the control condition, 10 in the high-power condition) showed no temporal discounting in the titration measure. More high-power participants showed no discounting than control and low-power participants, χ^2 (2) = 7.45, p =.02. The final sample size was 173 (77 women, 96 men, M_{age} = 21.22, SD_{age} = 1.55).

Procedure. This study was part of a series of studies that lasted for about an hour in total. Participants completed the study in individual cubicles in a common room in groups of 5 to 25. First participants completed the same power recall manipulation as in Study 2. They then completed the temporal discounting measure and the connection with the future self measure, with the order of these counterbalanced. The connection with the future self measure was the same as in Study 2. The temporal discounting measure differed in a few key ways from the measure used in Study 2, as we specify below. As in Study 2, the counterbalancing did not affect discount rate or connection with the future self, so it will not be discussed further. Finally, participants completed the General Sense of Power scale and reported their demographic information, as well as any suspicions they had about the study.

Temporal discounting measures. Temporal discounting was measured with an environmental preference survey, which consisted of a titration procedure and a free-response matching question. The instructions were the same as in Study 2, but the measures themselves differed in three ways. First, in the titration procedure, participants made nine choices between immediate improved air quality for 21 days and improved air quality one year from now for 35, 33, 31, 29, 27, 25, 23, 21, and 19 days. The last choice was not in the original study or in Study 2. Because participants with an indifference point less than or equal to 21 (i.e., who showed no temporal discounting) were excluded from all analyses, having this extra choice does not change the discount rate calculation. Second, after the choice titration, participants chose between immediate improved air quality for 21 days and $250. This question was included in the

Qualtrics survey obtained from P. Joshi but was not reported in the original paper. We will not discuss the result of this question since it is irrelevant to the current research. Third, the titration questions and the free-response matching question (which was phrased as in Study 2) were presented on the same page. In the original study and in Study 2, the matching question was presented on a separate page from the titration questions. Showing the matching question on the same page as the titration questions might have changed the way participants respond to the titration questions, such as by encouraging participants to indicate an indifference point beyond the range of the titration scale (e.g., Tost et al., 2015).

Results

Discount rate from only choice titration. Discount rate from only choice titration was calculated as in Study 2. 83 participants (48% of the final sample, 26 in the low-power condition, 30 in the control condition, 27 in the high-power condition) maxed out the titration scale. The percentage of participants maxing out the titration scale did not differ between power conditions, χ^2 (2) = 0.49, p =.782. A planned contrast analysis comparing high-power participants (M = 0.51, SD = 0.29) to the average of low-power (M = 0.51, SD = 0.28) and control participants (M = 0.52, SD = 0.27) was not significant, t(171) = -0.06, p = .950, η^2 = 0.00, 95% CI [0.00, 0.01].

A one-way ANOVA also did not show any effect of power condition, F(2,170) = 0.03, p = .972, η^2 = .000.

We also ran an equivalence test (Lakens, 2017) to test for evidence supporting a null effect, meaning no difference in discount rate between the low-power and high-power conditions. Using the same equivalence bound as Study 2, the equivalence test result provided evidence for the null hypothesis of no effect, t(106.45) = 2.21, p = .014.

Below we report results from other calculations of discount rates: 1) the discount rate from choice titration supplemented with the matching measure, and 2) the discount rate from the matching measure.

Discount rate from choice titration supplemented with matching measure. Discount rates were calculated as in Study 2. When participants maxed out the titration scale, their answers to the free-response matching question were used as their indifference point. Among these participants, 10 (1 in the low-power condition, 5 in the control condition, 4 in the high-power condition) reported an indifference point of fewer than 35 days in a year, which is inconsistent with their titration response. We excluded these 10 participants in addition to the 20 participants excluded from the titration procedure. Thus, 163 participants remained in the following analyses (71 women, 92 men; M_{age} = 21.2 years, SD_{age} = 1.53).

The discount rate distribution had a positive skew (skewness = 3.57, kurtosis = 17.94). As in the previous studies, we first excluded 24 participants (7 in the low-power condition, 10 in the control condition, and 7 in the high-power condition) whose discount rates were more than 3 interquartile ranges from the median. The likelihood of exclusion did not differ between power conditions, $\chi^2(2) = 0.82$, $p = .663$. Thus, 139 participants remained for this analysis (62 women, 77 men; M_{age} = 21.20 years, SD_{age} = 1.52). A one-way ANOVA did not show a significant effect of power ($M_{LP} = 0.91$, $SD_{LP} = 1.02$; $M_{Ctr} = 0.90$, $SD_{Ctr} = 1.12$; $M_{HP} = 0.72$, $SD_{LP} = 0.76$), $F(2, 136) = 0.54$, $p = .582$, $\eta^2 = .008$.

Second, we performed the three most common transformations on the discount rates: square root transformation (after transformation skewness = 2.16, kurtosis = 4.61), log transformation (after transformation skewness = 1.39, kurtosis = 0.92), and inverse transformation (after transformation skewness = 0.34, kurtosis = -1.21). Since square root transformation and log transformation both corrected the distribution adequately, we used both transformations in our analyses. Below we report the results based on the square root transformation. Using log transformed discount rates did not change any of these results. A one-way ANOVA did not show any effect of power condition (M_{LP} = 1.63, SD_{LP} = 1.04; M_{Ctr} = 1.78, SD_{Ctr} = 1.11; M_{HP} = 1.62, SD_{HP} = 0.96), $F(2, 160) = 0.44$, p = .645, η^2 = .005.

Third, we conducted a non-parametric test, the Kruskal-Wallis rank sum test, on the untransformed discount rates. This also showed no effect of power condition (Mdn_{LP} = 1.25, Mdn_{Ctr} = 1.25, Mdn_{HP} = 1.25), χ^2 (2) = 0.52, p = .770.

Discount rate from matching measure. 37 participants (11 in the low-power condition, 9 in the control condition, 17 in the high-power condition) showing no temporal discounting were excluded from this analysis. The proportion of no discount participants did not differ between power conditions, $\chi^2(2) = 3.84$, p = .146. Thus, 162 participants remained in this analysis (71 women, 91 men; M_{age} = 21.23 years, SD_{age} = 1.56).

The discount rate distribution had a positive skew (skewness = 3.11, kurtosis = 12.25), so we first performed the three most common transformations on the discount rates: square root transformation (after transformation skewness = 1.79, kurtosis = 3.04), log transformation (after transformation skewness = 0.97, kurtosis = -0.14), and inverse transformation (after transformation skewness = -0.16, kurtosis = -1.18). We used the square root transformation in our analyses because it brought the discount rates closest to a normal distribution. A one-way ANOVA did not show any effect of power condition (M_{LP} = 1.99, SD_{LP} = 1.30; M_{Ctr} = 2.07, SD_{Ctr} = 1.19; M_{HP} = 1.96, SD_{HP} = 1.28), $F(2, 159) = 0.11$, p = .895, η^2 = .001.

Second, we conducted a non-parametric test, the Kruskal-Wallis rank sum test, on the untransformed discount rates. This showed no effect of power condition (Mdn_{LP} = 1.53, Mdn_{Ctr} = 1.40, Mdn_{HP} = 1.53), χ^2 (2) = 0.54, p = .760.

A logistic regression with the proportion of times participants chose the immediate reward (21 days of improved air quality now) in the titration procedure also showed no difference between power conditions (M_{LP} = 71%, SD_{LP} = 33%, M_{Ctr} = 72%, SD_{Ctr} = 32%, M_{HP} = 71%, SD_{HP} = 34%), χ^2 (2) = 0.02, p = .872.

To summarize, we did not find any effect of power on temporal discounting with any of the discount rate calculations, or with a proportion of choices analysis.

Connection with the future self. The ratings of "connected" and "similar" were averaged for an index of connection with the future self (α = 0.71). A planned contrast analysis comparing high-power participants (M = 3.44, SD = 1.42) to the average of low-power (M = 3.59, SD = 1.34) and control participants (M = 3.91, SD = 1.44) was not significant, $t(171)$ = -1.33, p = .190, η^2 = 0.01, 95% CI [0.00, 0.06]. A one-way ANOVA also showed no effect of power condition, $F(2, 170)$ = 1.67, p = .191, η^2 = .019.

We also conducted mediation analyses with connection with the future self as the mediator of the effect of power on temporal discounting. Power condition (coded as 1 = low power, 2 = control, and 3 = high power, for all regression analyses for Study 5) did not predict temporal discounting, β = 0.00, 95% CI [-0.05, 0.05], $t(171)$ = 0.07, p = .943. Connection with the future self did not predict temporal discounting either, β = 0.00, 95% CI [-0.03, 0.03], $t(171)$ = 0.23, p = .821. When connection and power were both included in the model, neither power condition, β = 0.00, 95% CI [-0.05, 0.05], $t(170)$ = 0.08, p = .936, or connection, β = 0.00, 95% CI [-0.03, 0.03], $t(170)$ = 0.23, p = .819, predicted temporal discounting. There was no evidence for mediation, estimated indirect effect = 0.00, 95% bootstrapped confidence intervals of indirect effects [0.00, 0.00], p = .960.

Potential moderators. We report the moderator results with the titration-based discount rates below and summarize the results with the other discount rate calculations. The full results with other discount rate calculations can be found at https://osf.io/ze5ig/.

General Sense of Power. A one-way ANOVA found no effect of power condition on SOP (M_{LP} = 4.18, SD_{LP} = 0.26; M_{Ctr} = 4.11, SD_{Ctr} = 0.26; M_{HP} = 4.15, SD_{HP} = 0.33), $F(2, 170)$ = 1.01, p = .365, η^2 = .012.

Having shown that our manipulation did not affect SOP, we then tested whether SOP moderated the effect of power on temporal discounting. We regressed discount rate on SOP, power condition, and their interaction. The effect of SOP was not significant, β = 0.01, 95% CI [-0.11, 0.13], $t(169)$=0.17, p =.867. There was no effect of power condition, β = 0.00, 95% CI [-0.05, 0.06], $t(169)$=0.08, p=.940; and no interaction effect, β = -0.01, 95% CI [-0.06, 0.05], $t(169)$=-0.32, p=.751.

Other discount rate calculations also did not show any effect of SOP.

Gender. We also conducted an exploratory analysis to test whether gender moderated the effect of power on temporal discounting. A 2 (participant gender) by 3 (power condition) between-subjects ANOVA on discount rates showed that the main effect of gender was not significant, $F(1, 167)$ = 0.58, p = .447, η^2 = .003. There was no effect of power condition, $F(2, 167)$ = 0.06, p = .938, η^2 = .001; and no interaction effect, $F(2, 167)$ = 0.26, p = .770, η^2 = .003.

Other discount rate calculations also did not show any effect of gender.

Meta-analysis

To obtain a better estimate of the effect of power on temporal discounting, we conducted a meta-analysis. We set the scope of the meta-analysis to be experimental studies that manipulated power and measured temporal discounting with a titration and/or matching procedure. This allowed us to estimate the casual effect of power on temporal discounting. Correlational studies (Duan, et al., 2017, Study 2; Joshi & Fast, 2013, Study 4; May & Monga, 2014, Studies 1-4) were not included since our goal was to assess the causal evidence for power affecting temporal discounting. Additionally, we focused the meta-analysis on the difference between low-power and high-power conditions, as this allowed us to include the largest number of studies, and the supporting theories discussed by Joshi and Fast (2013) focused on differences between low-power and high-power individuals. We searched PsycINFO, Open Science Framework, PsychFileDrawer, and Arxiv with the keywords *Power* and *Temporal Discounting*/*Discount Rate* for studies that fit these criteria.

Studies that manipulated power and measured behaviors conceptually related to temporal discounting, such as saving (Garbinsky, Klesse, & Aaker, 2014) and delaying consumption (May & Monga, 2014, Study 5) were not included in the meta-analysis. Although participants in these studies were asked to make decisions regarding delaying rewards, the amount of the rewards and the length of the delays were not specified, preventing any estimation of discount rates. Furthermore, as discussed in the main manuscript, the effects of power in these two papers do not necessarily involve temporal discounting, and in fact both papers provided evidence for mechanisms other than temporal discounting.

Two additional experiments (Lee, Malkoc, & Rucker, 2013, Experiments 2-3) were brought to our attention by a reviewer. In each experiment, power was manipulated, and then participants indicated how much money they would need to delay the receipt of a reward for two different periods of time, to measure present bias (the tendency to discount more for a shorter period of delay than for a longer period of delay; Urminsky & Zauberman, 2015). Though these experiments found a marginally significant tendency for high-power participants to show reduced present bias, they found no main effect of power on temporal discounting (S. Malkoc, personal communication, September 20, 2017). These two experiments do not fit our selection criteria, as both experiments included only high-power and control conditions, so they were not included in the meta-analysis.

We found three papers reporting a total of four experiments that satisfied these criteria (Duan et al., 2017, Studies 1 and 3; Heller & Ullrich, 2017; Tost et al., 2015, Experiment 2). We also included Studies 1 and 3 of Joshi and Fast (2013), the two direct replications reported in our paper (Studies 1 and 2), and the three replications with some procedural differences from the target studies reported in the supplemental materials (Studies 3-5). Table 1 contains key information for all 11 experiments included in the meta-analysis. All experiments except for Duan et al. (2017, Studies 1 and 3) and Heller and Ullrich (2017) were conducted with U.S. participants. Duan et al.'s (2017) studies were conducted with Chinese participants of Han and Tibetan ethnicity. Heller and Ullrich's (2017) study was conducted with German and Swiss participants.

The experiments included in this meta-analysis manipulated power with either a role-based virtual team task or with a recall prime, and measured temporal discounting either with a monetary reward scenario or with an air quality scenario. Considering the differences in study design, as well as the different currencies involved in the monetary discount rate measures, we

Table 1

Experiments Included in the Meta-Analysis of Power's Effect on Temporal Discounting

Experiment	Sample	N	Power Manipulation	Temporal Discounting Measure
Joshi & Fast (2013, Study 1)	Mturk	67	Roles in a Virtual Team	Monetary Scenario
Current Study 1	University	342	Roles in a Virtual Team	Monetary Scenario
Current Study 3	Mturk	193	Roles in a Virtual Team	Monetary Scenario
Current Study 4	Mturk	375	Roles in a Virtual Team	Monetary Scenario
Tost et al. (2015, Exp 2)	Mturk	77	Roles in a Virtual Team	Monetary Scenario
Duan et al. (2017, Study 1)	University	78	Recall Writing Prime	Monetary Scenario
Duan et al. (2017, Study 3)	University	80	Recall Writing Prime	Monetary Scenario
Joshi & Fast (2013, Study 3)	University	52	Recall Writing Prime	Air Quality Scenario
Current Study 2	University	258	Recall Writing Prime	Air Quality Scenario
Current Study 5	University	112	Recall Writing Prime	Air Quality Scenario
Heller & Ullrich (2017)	Online	138	Recall Writing Prime	Air Quality Scenario

Notes. Samples sizes reported are the number of participants in the low-power and high-power conditions in each study. Mturk sample refers to participants recruited from the Amazon Mechanical Turk platform who completed the study online. University sample refers to university students who completed the study in a laboratory. Heller and Ullrich (2017) used a convenience sample recruited online consisting of roughly half students and half professionals.

conducted the meta-analysis with a random-effects model using standardized effect size estimates (Cohen's d; Cumming, 2014) with the MARVIS R package (Hamilton, 2017).

We calculated discount rates as in Joshi and Fast (2013), using the titration-plus-matching discount rate for monetary discount rate measures (as in their Study 1), and the titration-only discount rate for air quality discount rate measures (as in their Study 3). However, since the titration-plus-matching discount rate was not available for our Studies 3 and 4, we used the titration-only discount rate for these two studies. Another exception is Tost et al. (2015, Experiment 2), which used a monetary discount rate measure. For the titration-plus-matching calculation, they excluded participants with discount rates more than three standard deviations away from the mean. However, their reported mean discount rates and standard deviations (M_{LP} = 1.85, SD_{LP} = 2.62; M_{HP} = 4.32, SD_{HP} = 14.15) were still very different from the other studies in the meta-analysis. Thus, in the meta-analysis we used the titration-only discount rate (M_{LP} = 0.73, SD_{LP} = 0.29; M_{HP} = 0.85, SD_{HP} = 0.24) reported by Tost et al. (2015, Experiment 2), which produces descriptive statistics comparable to the other studies. We then calculated the standardized effect sizes (Cohen's d) of the discount rate differences between the low-power and high-power conditions for each study in the meta-analysis.

Figure 1 in the main text shows the standardized effect sizes for each study, with 95% confidence intervals, plus the resultant random-effect meta-effect. The meta-effect of power on discount rate was -0.11, 95% CI [-0.25, 0.03]. Within the monetary discounting studies, the meta-effect of power was -0.10, 95% CI [-0.29, 0.08]. Within the air quality discounting studies, the meta-effect of power was -0.15, 95% CI [-0.40, 0.10]. Thus, the evidence overall, as well as within either study design, is not consistent with an effect of power on temporal discounting.

The meta-analysis showed a small to moderate amount of heterogeneity (I^2 = 47.19%) across studies, Q_{10} = 20.01, p = .03, meaning heterogeneity accounts for 47.19% of the total variability in the data (Hamilton, 2017). While heterogeneity suggests the existence of moderators, in this case it was difficult to identify such moderators. The heterogeneity between monetary and air quality studies was not significant, Q_1 = 0.09, p = .77. Other notable differences between the studies are the different samples and power manipulations used. However, inspection of the forest plot (see Figure 1 in main text) does not reveal any sample or power manipulation effects.

References

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

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

Anderson, C., Kraus, M. W., Galinsky, A. D., & Keltner, D. (2012). The local-ladder effect: social status and subjective well-being. *Psychological Science*, *23*(7), 764–71. http://doi.org/10.1177/0956797611434537

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

Avila, C. (2001). Distinguishing BIS-mediated and BAS-mediated disinhibition mechanisms: A comparison of disinhibition models of Gray (1981, 1987) and of Patterson and Newman (1993). *Journal of Personality and Social Psychology*, *80*(2), 311–324. http://doi.org/10.1037/0022-3514.80.2.311

Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences.* Palgrave Macmillan.

Crockett, M. J., Clark, L., Lieberman, M. D., Tabibnia, G., & Robbins, T. W. (2010). Impulsive choice and altruistic punishment are correlated and increase in tandem with serotonin depletion. *Emotion*, *10*(6), 855–862. http://doi.org/10.1037/a0019861

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. http://doi.org/10.1177/0956797613504966

Duan, J., Wu, S. J., & Sun, L. (2017). Do the powerful discount the future less? The effects of power on temporal discounting. *Frontiers in Psychology*, *8*(June), 1–11. http://doi.org/10.3389/fpsyg.2017.01007

Fujita, K., & Carnevale, J. J. (2012). Transcending Temptation Through Abstraction: The Role of Construal Level in Self-Control. *Current Directions in Psychological Science*, *21*(4), 248–252. http://doi.org/10.1177/0963721412449169

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*(3), 453–66. http://doi.org/10.1037/0022-3514.85.3.453

Garbinsky, E. N., Klesse, A.-K., & Aaker, J. (2014). Money in the bank: Feeling powerful increases saving. *Journal of Consumer Research*, *41*(3), 610–623. http://doi.org/10.1086/676965

Hardisty, D. J., Thompson, K. F., Krantz, D. H., & Weber, E. U. (2013). How to measure time preferences: An experimental comparison of three methods. *Judgment and Decision Making*, *8*(3), 236–249. http://doi.org/10.1007/s10826-012-9600-6

Hardisty, D. J., & Weber, E. U. (2009). Discounting future green: Money versus the environment. *Journal of Experimental Psychology. General*, *138*(3), 329–340. http://doi.org/10.1037/a0016433

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. http://doi.org/10.1037/a0020761

Jones, B., & Rachlin, H. (2006). Social Discounting. *Psychological Science*, *17*(4), 283–286. http://doi.org/10.1111/j.1467-9280.2006.01699.x

Joshi, P. D., & Fast, N. J. (2013). Power and reduced temporal discounting. *Psychological Science*, *24*(4), 432–438. http://doi.org/10.1177/0956797612457950

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

Kim, H., Schnall, S., & White, M. P. (2013). Similar psychological distance reduces temporal discounting. *Personality and Social Psychology Bulletin*, *39*(8), 1005–16. http://doi.org/10.1177/0146167213488214

Koff, E., & Lucas, M. (2011). Mood moderates the relationship between impulsiveness and delay discounting. *Personality and Individual Differences*, *50*(7), 1018–1022. http://doi.org/10.1016/j.paid.2011.01.016

Kraus, M. W., & Keltner, D. (2009). Signs of Socioeconomic Status. *Psychological Science*, *20*, 99–106. http://doi.org/10.1111/j.1467-9280.2008.02251.x

Kraus, M. W., Piff, P. K., & Keltner, D. (2009). Social class, sense of control, and social explanation. *Journal of Personality and Social Psychology*, *97*(6), 992–1004. http://doi.org/10.1037/a0016357

Lakens, D. (2017). Equivalence Tests. *Social Psychological and Personality Science*, *8*(4), 355–362. http://doi.org/10.1177/1948550617697177

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2008). Illegitimacy moderates the effects of power on approach. *Psychological Science*, *19*(6), 558–64. http://doi.org/10.1111/j.1467-9280.2008.02123.x

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2012). Power increases social distance. *Social Psychological and Personality Science*, *3*(3), 282–290. http://doi.org/10.1177/1948550611418679

Lempert, K. M., Glimcher, P. W., & Phelps, E. A. (2015). Emotional arousal and discount rate in intertemporal choice are reference dependent. *Journal of Experimental Psychology:*

General, *144*(2), 366–373. http://doi.org/10.1037/xge0000047

Magee, J. C., & Smith, P. K. (2013). The Social distance theory of power. *Personality and Social Psychology Review*, *17*(2), 158–186. http://doi.org/10.1177/1088868312472732

May, F., & Monga, A. (2014). When time has a will of its own, the powerless don’t have the will to wait: Anthropomorphism of time can decrease patience. *Journal of Consumer Research*, *40*(5), 924–942. http://doi.org/10.1086/673384

Peters, J., & Büchel, C. (2010). Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. *Neuron*, *66*(1), 138–48. http://doi.org/10.1016/j.neuron.2010.03.026

Read, D., Frederick, S., & Scholten, M. (2013). DRIFT: An analysis of outcome framing in intertemporal choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 573–588. http://doi.org/10.1037/a0029177

Schmeichel, B. J., Harmon-Jones, C., & Harmon-Jones, E. (2010). Exercising self-control increases approach motivation. *Journal of Personality and Social Psychology*, *99*(1), 162–73. http://doi.org/10.1037/a0019797

Smith, P. K., Jost, J. T., & Vijay, R. (2008). Legitimacy crisis? Behavioral approach and inhibition when power differences are left unexplained. *Social Justice Research*, *21*(3), 358–376. http://doi.org/10.1007/s11211-008-0077-9

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software*, *59*(5), 1–38. http://doi.org/10.18637/jss.v059.i05

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

Decision Processes, *128*, 61–73. http://doi.org/10.1016/j.obhdp.2015.03.003

Urminsky, O., & Zauberman, G. (2015). The psychology of intertemporal preferences. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making* (pp. 141–181). Chichester, UK: John Wiley & Sons, Ltd. http://doi.org/10.1002/9781118468333.ch5

Wang, X. T., & Dvorak, R. D. (2010). Sweet Future. *Psychological Science*, *21*(2), 183–188. http://doi.org/10.1177/0956797609358096

Weber, E. U., Johnson, E. J., Milch, K., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science*, *18*(6), 516–524. http://doi.org/10.1111/j.1467-9280.2007.01932.x

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

Footnotes

[1]In the lab where Studies 1, 2, and 5 took place, the norm is for study sessions to be approximately one hour in length and involve multiple studies. Thus, our participants are accustomed to hour-long study sessions. They are dismissed once the hour is up, regardless of where they are in the series of studies, and they are not penalized if they do not complete all studies. Thus, there is little pressure to rush through the studies.

[2]For transparency, we report the number of participants who maxed out the titration scale in each study. Although researchers report using different ways to determine maxed-out participants' discount rates (e.g. Hardisty & Weber, 2009; Jones & Rachlin, 2006), they rarely report the number of participants who maxed out. Joshi and Fast (2013) also did not report this information. Joshi (personal communication, February 1, 2017) indicated that the max-out rates in their studies were less than 15%, but did not provide specific percentages for each study. We were able to gather information on what percentage of participants maxed out from two other replication studies, Tost et al. (2015, Experiment 2; L. Tost, personal communication, June 27, 2017) and Heller and Ullrich (2017; publicly available data at https://osf.io/dqr4m/), as well as from two papers that used discounting measures comparable to ours, Hardisty and Weber (2009; publicly available data at http://davidhardisty.info/cv.php) and Hardisty, Thompson, Krantz, and Weber (2013; publicly available data at http://davidhardisty.info/cv.php), for comparison. Though we do not know if these studies are representative of temporal discounting research with such measures, they at least provide a coarse benchmark for the range of max-out rates that occur normally. Below we compare rates from these four papers to the rates in our studies to assess if the rates in our studies are atypical. Since the rate of participants maxing out the scale is affected by the upper bound of the titration scale, and the upper bound used can vary greatly from study

to study, we report these upper bounds for all studies and adjust our calculations accordingly to provide rates that are appropriate to compare between studies.

Our Studies 1, 3, and 4 used a monetary discount rate measure. The upper bound of the monetary titration scale was at a discount rate of 1 (i.e., $k = 1$), meaning that participants whose discount rates were higher than 1 maxed out the scale. In these three studies, 9%, 41%, and 33% of the final samples, respectively, maxed out the titration scale. Tost et al. (2015, Experiment 2) used the same monetary titration scale with an upper bound of $k = 1$, and had max-out rates of 45% in the low-power condition and 58% in the high-power condition. Hardisty and Weber's (2009) Studies 1 and 2 used a similar monetary titration scale, except that the upper bound of their scale was $k = 0.67$ for both studies. Their max-out rates were 6% (Study 1) and 27% (Study 2). Hardisty et al. (2013) involved monetary titration scales varying in time delay from 1 year to 50 years in Study 1, and 6 months to 10 years in Study 2. Since our monetary titration scale involved a delay of 1 year, we focused on the data for the titration scale with a 1-year delay. Hardisty et al.'s (2013) 1-year-delay monetary titration scales had upper bounds as high as $k = 282.33$ (Study 1) and $k = 32.33$(Study 2), making their max-out rates (4% and 0%, respectively) difficult to compare with ours. As a proxy, we looked at the percentage of participants who had discount rates above the upper bound used in our studies (i.e., $k > 1$). In Hardisty et al.'s (2013) Study 1, 47% of participants had a discount rate greater than 1, and in Study 2 it was 61%. In sum, the data we were able to access from studies using identical or comparable monetary discount rate measures showed substantial variability in max-out rates (6% to 61%), and the max-out rates we had in Studies 1, 3, and 4 fit within this range.

Our Studies 2 and 5 used an air-quality discount rate measure. The upper bound of the air-quality titration scale was at a discount rate of 0.67 (i.e., $k = 0.67$), meaning that participants whose discount rates were higher than 0.67 maxed out the scale. In these two studies, 53% and 48% of the final samples, respectively, maxed out the titration scale. Heller and Ullrich (2017) used the same air-quality titration scale with an upper bound of $k = 0.67$, and had a max-out rate of 43%. Hardisty and Weber's (2009) Study 1 used a similar air-quality titration scale, except that the upper bound of their scale was $k = 0.76$. Their max-out rate was 19%. Hardisty and Weber's (2009) Study 3 used a similar air-quality titration scale except for two differences. First, they varied whether the time delay was 1 year or 10 years. Since our air-quality titration scale involved a delay of 1 year, we focused on the data for the titration scale with a 1-year delay. Second, the upper bound of their scale for the 1-year delay was $k = 4$, making their max-out rate (9%) difficult to compare with ours. As a proxy, we looked at the percentage of participants who had discount rates above the upper bound used in our studies (i.e., $k > 0.67$), which was 58% of participants in Hardisty and Weber's (2009) Study 3. In sum, the data we were able to access from studies using identical or comparable air-quality discount rate measures showed substantial variability in max-out rates (19% to 58%), and the max-out rates we had in Studies 2 and 4 were within but at the upper end of this range.

Overall, from the limited data we could obtain on max-out rates with discount rate measures similar or identical to the ones we used, we did not find evidence that the max-out rates in our own studies were unusual. It is worth noting that maxing out the titration scale did not appear to be considered an irrational behavior in Joshi and Fast (2013), or in any of the other studies cited above. When researchers feel that participants have responded irrationally to titration scales, such as by showing multiple indifference points, these participants are normally excluded from further analyses. Data from maxed-out participants were never reported to be excluded in the studies we examined. Instead, researchers either used a matching discount measure to supplement the titration measure to calculate a discount rate for maxed-out participants (e.g., Hardisty & Weber, 2009; Joshi & Fast, 2013, Studies 1 and 2; Tost et al., 2015, Experiment 2), or assigned these participants a discount rate one step above the titration scale upper bound (e.g., Heller & Ullrich, 2017; Jones & Rachlin, 2006; Joshi & Fast, 2013, Study 3). These methods were used both in studies with low max-out rates (e.g., 6% in Hardisty & Weber, 2009, Study 1, monetary scenario) and in studies with high max-out rates (e.g., 43% in Heller & Ullrich, 2017, air-quality scenario). We preregistered and reported the results of both methods for all our studies.