People are More Skeptical of Others' Public Virtue Motivations than Their Own in Separate (but

not Joint) Evaluations

*Kyle Fiore Law[1], Jordan Wylie[2], Gordon Kraft-Todd[3], Nathan Liang[4], Liane Young[2,5], &

Stylianos Syropoulos[1]

[1]School of Sustainability, College of Global Futures, Arizona State University, Tempe, AZ

[2]Department of Psychology and Neuroscience, Boston College, Chestnut Hill, MA

[3]Ker-twang

[4]Department of Psychology, Cornell University, Ithaca, NY

[5]The Schiller Institute for Integrated Science and Society, Boston College, Chestnut Hill, MA

*Corresponding Author

Kyle Fiore Law: kflaw@albany.edu; https://orcid.org/0000-0001-5911-4772
Gordon Kraft-Todd: gordon@ker-twang.com; https://orcid.org/0000-0003-1220-9269
Jordan Wylie: wyliej@bc.edu; https://orcid.org/0000-0001-5142-8515
Nathan Liang: nl453@cornell.edu; https://orcid.org/0000-0002-5308-6368
Liane Young: liane.young@bc.edu
Stylianos Syropoulos: syropoul@bc.edu; https://orcid.org/0000-0001-5622-1417

CRediT Concept	KFL	JW	NL	GKT	LY	SS
Conceptualization	X	X		X	X	X
Methodology				X		X
Software	X					X
Validation						X
Formal analysis						X
Investigation				X		X
Resources					X	
Data Curation						X
Writing - Original Draft	X					X
Writing - Review & Editing	X	X	X	X	X	X
Visualization	X		X	X		
Supervision						X
Project administration	X					X
Funding acquisition				X	X	

Abstract

Public acts of virtue can promote prosocial norms yet are often met with moral skepticism—a phenomenon known as virtue discounting. What psychological processes might underlie people's propensity to both discount others' public virtue and also engage in it themselves? We examine one possible explanation: whether people expect their own public virtuous behavior to be judged more favorably than others' similar actions. Across four pre-registered studies (N = 2,511), we tested for self-serving asymmetries in moral expectations. In three between-subjects experiments, participants either anticipated how others would evaluate their own actions (meta-perceptions) or judged the actions of another person (third-party judgments). Study 1 found no asymmetry in moral goodness. But in Studies 2 and 3, participants expected their own public virtue to be judged as more principled (and more morally good, in Study 2), less reputation-driven, and more trustworthy. Study 3 showed these asymmetries held across multiple perspectives. In contrast, Study 4 used a within-subjects design and found that self-serving asymmetries disappeared when judgments were made side by side. Together, these findings clarify how self-enhancement shapes moral expectations under naturalistic conditions and extend research on moral self-enhancement beyond trait judgments to public virtue and its perceived motivation.

Keywords: *motivation*, *morality*, *virtue discounting*, *reputation*, *social perception*

Statement of Contribution

What is already known on this subject?

- Public virtuous acts inspire prosocial norms but are often judged more harshly than private acts.

- This reputational skepticism is linked to perceived motivation (e.g., signaling vs. sincerity).

- People show self-enhancement bias, tending to believe they are more moral on average than others.

What does this study add?

- Demonstrates self-serving asymmetries in how people expect others will judge their public virtue.

- Reveals that self-enhancement shapes perceived motivation more than perceived moral goodness.

- Shows these asymmetries emerge in separate-evaluations but disappear under direct comparison.

People Are More Skeptical of Others' Public Virtue Motivations than Their Own in

Separate (but not Joint) Evaluations

From bumper stickers and yard signs to online activism and conspicuous displays of generosity, people routinely communicate their good deeds to others (Miller, 2019). Public acts of virtue serve an important adaptive purpose within a cooperative society, establishing prosocial norms that encourage collective good by inspiring others to act morally (Berkowitz, 1972; Levy, 2021; Nook et al., 2016; Sparkman et al., 2022). Nonetheless, these behaviors are often met with skepticism, as observers discount their moral worth by attributing self-serving reputational motives rather than morally principled ones (Johnson & Park, 2021; Jordan & Rand, 2019; Jung et al., 2017; Konuk & Otterbring, 2024), a phenomenon known as virtue discounting (Kraft-Todd et al., 2023). It is well-established that moral judgment plays a role in regulating social behavior (Curry et al., 2019; Ellemers & van den Bos, 2012; Graham et al., 2011), and people modify their actions in response to perceived approval or condemnation (Day et al., 2014; Ostermaier & Uhl, 2017). What psychological processes might underlie people's propensity to discount others' public virtue and also engage in it themselves?

We do not claim to resolve this paradox by identifying its causal mechanism. Instead, our aim is to investigate whether a self-serving asymmetry in moral expectations exists: that is, whether people tend to anticipate that their own public virtue will be judged differently than others' similar acts tend to be. People routinely show self-serving asymmetries in other domains, such as competence (Dunning et al., 1989), happiness (Klar & Giladi, 1997), and morality (Aquino & Reed II, 2002), but it remains an open question whether similar biases extend to expectations about public virtue. Documenting such an asymmetry would not explain, definitively, why people engage in public virtue despite discounting others, but it would clarify one psychological factor that may help illuminate how people navigate the reputational stakes of

moral behavior. In particular, if people expect more charitable evaluations of their own actions than they typically afford to others, this perceptual gap could help explain why the reputational risks of public virtue may feel less discouraging from the actor's perspective than they appear from the observer's.

Virtue Discounting and the Second Paradox of Public Virtuosity

Public displays of virtue can be contagious (Berkowitz, 1972; Cialdini & Trost, 1998; Nook et al., 2016). They can signpost to others that good deeds are common within one's social context and reinforce norms that encourage broader engagement in prosocial behavior. A rich literature on social influence has consistently demonstrated the power of norms in shaping actions that serve to benefit the collective good, like charitable giving (Zaki, 2020), pro-environmentalism (Cole et al., 2022; Law et al., 2024), civic engagement (Gerber & Rogers, 2009), public health compliance (Mladenović et al., 2023), and workplace ethics (Appelbaum et al., 2005). Accordingly, public virtuosity holds an important place within societies, helping to promote cohesion and cooperation within large-scale groups (Allison, 1992; Axelrod, 1986; Gintis, 2003; Levy, 2021).

Despite the social value of public virtuosity, recent research by Kraft-Todd et al. (2023) highlights a critical paradox: while visible moral acts can inspire others, people frequently doubt their authenticity. This "virtue discounting" occurs when observers attribute public generosity to reputational motives rather than genuine intentions. Across multiple studies, individuals who judged morally virtuous actions performed in public rated them as less sincere and morally inferior on average compared to individuals who judged identical actions performed in private. One explanation is that public virtue challenges the Aristotelian ideal (Aristotle, 1998) that "true" virtue is pursued for its *own sake*, not for *external validation* (Cokelet & Fowers, 2019). Empirical evidence supports this perspective (Kraft-Todd et al., 2023), showing that visibility

shifts attributions from principled motivation to reputation signaling, leading to harsher moral judgments.

Yet, virtue discounting presents a second paradox: the same individuals who at one time scrutinize others' public virtuous acts may at other times engage in virtuous acts themselves, sometimes conspicuously in public settings (Johnson & Park, 2021; Jung et al., 2017; Miller, 2019). If people consistently discount the sincerity of others' acts of public virtue, one might expect that people, on average, anticipate the same skepticism from others toward their own actions, diminishing any perceived reputational benefits of public moral behavior. However, this assumes that people's "meta-perceptions" (i.e., how they believe others will judge them) align with typical "third-party judgments" (i.e., how people judge others). In practice, however, people may operate under different evaluative standards depending on perspective, tending to assume others will view their own virtue as principled, even despite general tendencies to judge others' virtue as reputationally motivated. This asymmetry may reflect a self-enhancing bias that remains largely unexplored.

Self vs. Others: Do People Anticipate Less Virtue Discounting for Their Own Actions?

Self Enhancement and Self-Serving Biases

While the emerging literature on virtue discounting suggests that observers devalue public virtuous acts as motivated by self-interest (Kraft-Todd et al., 2023; see also Berman & Silver, 2022), research on self-perception suggests that individuals may not apply this same skepticism to themselves. Self-enhancement, the motivation to maintain a positive self-image and high self-esteem, has been conceptualized as a fundamental psychological need (Brookes, 2015; Brown et al., 1988; Gebauer et al., 2013; Sedikides & Gregg, 2008; Wojciszke & Białobrzeska, 2014). One way that people have been found to satisfy this need is by showing self-serving biases in how they evaluate themselves versus others. For instance, the "Better-

Than-Average Effect" (BTAE) refers to the tendency for people to see themselves as superior to

the average person (Alicke et al., 1995; Alicke & Govorun, 2005; Brown, 1986; Dunning et al.,

1989; Zell et al., 2020). Similar self-serving biases have been observed in how people evaluate

their own versus others' competence (Dunning et al., 1989; Kim & Han, 2023), their relative

happiness (Klar & Giladi, 1997, 1999), and even the future prospects of their own group

compared to others (Berntsen & Rubin, 2024). Yet, among the many domains in which people

exhibit inflated self-perceptions, the "illusion of moral superiority" stands out as particularly

pronounced (Aquino & Reed II, 2002; Ding & Sugiura, 2024; Goodwin et al., 2014; Tappin &

McKay, 2017). Even convicted criminals rate themselves as more moral and prosocial than both

their incarcerated peers and the general public (Sedikides et al., 2014; Taylor et al., 2021).

Merging Insights from Self-Enhancement and Virtue Discounting

While psychological inquiry into self-enhancement has revealed robust tendencies to

engage in self-serving asymmetries, two important distinctions have received less attention. First,

much of the self-enhancement literature, especially in the moral domain, has focused on how

people evaluate themselves versus others on stable character traits such as honesty, fairness, and

generosity (Aquino & Reed, 2002; Ding & Sugiura, 2024; Goodwin et al., 2014; Tappin &

McKay, 2017). In contrast, less attention has been paid to self-serving biases in the context of

moral *behaviors* and their underlying *motivations*. Yet, behaviors and their motivations remain

especially psychologically relevant in the context of virtue discounting, where moral evaluations

hinge less on characterological traits and more on judgments of specific actions and the sincerity

behind them (Kraft-Todd et al., 2023). And, virtue discounting remains both practically

consequential, since public displays can promote virtuous norms (Berkowitz, 1972; Cialdini &

Trost, 1998; Nook et al., 2016), and theoretically puzzling, as people typically avoid engaging in

behaviors that trigger the very kind of moral judgment that public virtue often provokes (Kraft-Todd et al., 2023; Miller, 2019).

Second, although self-enhancement is typically demonstrated using within-subjects comparisons between the self and others (e.g., the BTAE; Alicke et al., 1995; Alicke & Govorun, 2005; Brown, 1986; Dunning et al., 1989; Zell et al., 2020), real-world opportunities to judge virtue may not always afford the opportunity for explicit comparison. That is, people often evaluate others' public actions, such as witnessing a charitable post online (Kraft-Todd et al., 2023; Miller, 2019), without simultaneously considering how others would judge their own similar behavior, or how they themselves would judge the same behavior if it was performed in private. This reflects what Hsee and colleagues (1996) describe as a "separate evaluation" context, in which options are assessed in isolation rather than side by side, and judgments tend to rely on more salient or easily evaluated attributes, like publicity. More generally, excepting explicit moral dilemmas (e.g., Pizarro & Bloom, 2003), people often arrive at moral judgments quickly and intuitively, without rationally comparing behavior to objective standards or tangible benchmarks (Ditto et al., 2009; Greene & Haidt, 2002).

Reflecting this naturalistic structure, emerging research on virtue discounting has used between-subjects designs in which participants evaluate public *or* private behaviors, but not both, in a separate evaluation context (Kraft-Todd et al., 2023). Our first three studies follow this ecologically valid approach to test whether self-serving asymmetries in moral perception emerge in the domain of virtue discounting under these more realistic conditions, where only a single type of judgment is made. In particular, our primary focus is on whether "meta-perceptions" of virtuous behaviors, or how people expect their own behaviors and motives to be judged, systematically diverge from "third-party judgments" of virtuous behaviors, or how people tend to judge others' behaviors and motives. To date, despite using more naturalistic separate evaluation

contexts, research on virtue discounting has focused exclusively on third-party judgments, leaving open how people expect their own virtue to be evaluated. Yet individuals routinely occupy both observer and actor roles in everyday moral life, making this comparison especially relevant to understanding real-world moral perception. This distinction, between how people believe *they* will be judged and how people tend to judge *others*, is especially important in contexts where moral actions are publicly observable and subject to reputational scrutiny. We reason that, although observers often discount others' public virtue by attributing it to reputational motives rather than principled intent (Johnson & Park, 2021; Jordan & Rand, 2019; Jung et al., 2017; Konuk & Otterbring, 2024), people may not expect their own public virtue to be viewed quite as negatively.

Of course, the salience of virtue discounting may exert a stronger influence on judgments of public moral acts than self-enhancement, effectively nullifying any self-serving bias in this domain. While people generally see themselves as more moral than others, heightened skepticism of public virtue may override this bias, leading individuals to assume that public acts of goodness will be perceived as reputation-driven rather than genuinely principled, regardless of who performs them. Prior research suggests that people are highly attuned to the reputational stakes of being observed, often adjusting their behavior to appear more prosocial (Kraft-Todd et al., 2015; Kraft-Todd et al., 2023). If people are aware of this tendency in themselves and others, they might reasonably assume that others recognize it as well. As a result, they could expect their own public virtue to be met with the same skepticism they apply when evaluating others' public virtue, thereby eliminating any asymmetry between meta-perceptions and third-party perceptions in the context of virtue discounting. As such, in our final study, we examine whether the same patterns emerge in a within-subjects "joint evaluation" context (Hsee, 1996), where people's

judgments of others may serve as a clearer baseline for forming expectations about how they themselves will be judged.

Collectively, our investigation draws from both the literature on self-enhancement and the emerging paradigm of virtue discounting (Kraft-Todd et al., 2023). By integrating insights from both literatures and employing a mix of between-subjects designs that mirror real-world evaluative contexts and a within-subjects design that makes self-other contrasts salient to participants, we test whether self-serving asymmetries extend to how people anticipate being judged by others for their own morally virtuous actions. Together, our studies not only clarify whether, and under what circumstances, self-serving asymmetries manifest across observer perspectives, but also offer practical insight into one explanation, albeit a potential one, as to which psychological processes might underlie continued engagement in public virtue despite reputational skepticism.

The Current Studies

Across four primary studies, we test whether a self-serving bias emerges in the context of virtue discounting when people imagine how others will morally judge and attribute underlying motivations to their actions. Specifically, we examine whether people expect their own public acts of virtue to be judged as morally superior, more principled, and less reputation-driven than people tend to, on balance, judge the public virtuous acts of others. To capture differences in moral judgment across the numerous contexts in which they might arise, our studies distinguish among four judgment contexts: meta-perceptions (how people believe others will judge their actions and motivations), third-party judgments (how people judge others' actions and motivations), "self-judgments" (how people judge their own actions and motivations), and "second-order perceptions" (how people believe others will judge someone else's actions and

motivations). Because this framework can be conceptually dense, we include Table 1 and Figure

1 to support clarity.

Table 1

Summary of asymmetries found in the context of pubic acts of virtue across Studies 1-3.

Study	Comparison(s)	Outcome(s)	Asymmetry Effects for Public Virtue
Study 1 (Separate Evaluation)	Meta-perceptions vs. third-party judgments (i.e., how participants expect to be judged by observers vs. how participants judge others)	Moral Goodness	Participants, on average, expected their own public virtuous actions to be judged as no more morally good than participants, on average, judged others'.
Study 2 (Separate Evaluation)	Meta-perceptions vs. third-party judgments (i.e., how participants expect to be judged by observers vs. how participants judge others)	Moral Goodness	Participants, on average, expected their own public virtuous actions to be judged as more morally good than participants, on average, judged others'.
		Principled Motivation	Participants, on average, expected observers to judge their motivations for public virtuous actions as more principled than the motivations participants, on average, attributed to others.
		Reputation Signaling	Participants, on average, expected observers to judge their motivations for public virtuous actions as less reputation-driven than the motivations participants, on average, attributed to others.
		Trust	Participants, on average, expected observers to see them as more trustworthy when they engage in public virtuous actions than participants, on average, perceived others who do the same.
		Normative Expectations	Participants, on average, expected others to see their public virtuous actions as less normative (i.e., more unique) than participants, on average, saw others' virtuous actions.
Study 3 (Separate Evaluation)	*Meta-perceptions/self-perceptions vs. third-party judgments/second-order perceptions (i.e., how participants expect to be judged by observers/how participants judge themselves vs. how participants judge others/how participants expect others to judge others)	Moral Goodness	Participants, on average, saw their own public virtuous actions as no more morally good than participants, on average, judged others' (and expected observers to make the same judgments).
		Principled Motivation, Reputation Signaling, Trust	Study 3 replicated the patterns from Study 2, showing that participants, on average, viewed their own public virtue motivations as more principled, less reputation-driven, and themselves as more trustworthy. In contrast, participants, on average, perceived others' motivations as less principled, more reputation-driven, and others as less trustworthy (and expected observers to make the same judgments).
		Normative Expectations	Mirroring Study 2, participants, on average, expected others to see their public virtuous actions as less normative (i.e., more unique) than participants, on average, saw others' virtuous actions (and expected observers to make the same judgments).
Study 4 (Joint Evaluation)	Meta-perceptions vs. third-party judgments (i.e., how participants expect to be judged by observers vs. how participants judge others)	Moral Goodness, Principled Motivation, Reputation Signaling	Here, where the same participants reported meta-perceptions and third-party judgments side by side in a joint evaluation context, no self-serving asymmetry emerged for any outcome.

Note. No differences were found between meta-perceptions and self-perceptions nor between third-party perceptions and second-order perceptions in Study 3.

Figure 1

Conceptual distinctions between judgment contexts.
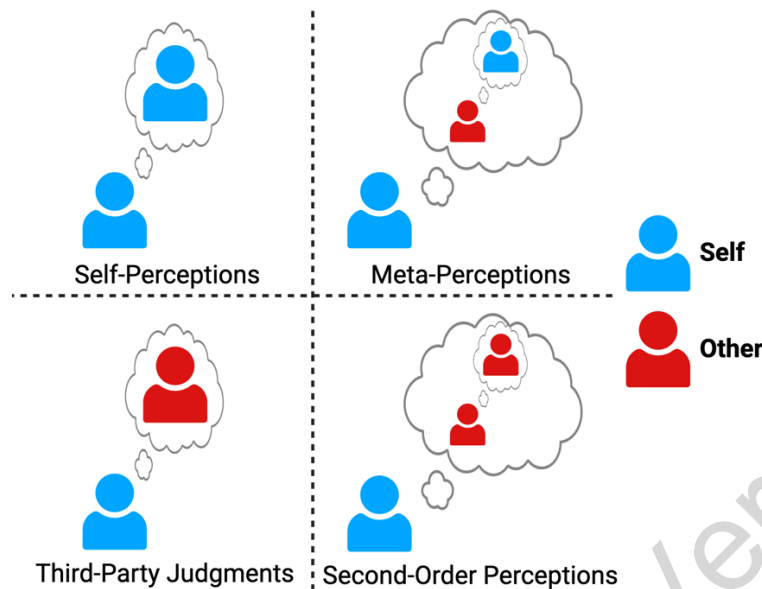
Note. A visual depiction of the distinctions between the four judgment contexts used as stimuli across studies: "self-judgments" (how people judge their own actions and motivations), meta-perceptions (how people believe others will judge their actions and motivations), third-party judgments (how people judge others' actions and motivations), and "second-order perceptions" (how people believe others will judge someone else's actions and motivations).

In a supplementary study (see Supplementary Study S1), we offer a first-pass test of self-serving asymmetries in meta-perceptions of moral character using a joint evaluation context (i.e., within-subjects design). Specifically, we contrast meta-perceptions with third-party judgments to assess whether people *expect their own moral character to be judged* more favorably by others than how *they themselves judge others' moral character*. Study 1 tests for meta-perceptual self-serving asymmetries in the focal context of virtue discounting. This contrast, between how people expect to be judged and how people tend to judge others, has not previously been examined in the virtue discounting literature, which has relied exclusively on third-party evaluations. Specifically, we test whether people, on average, apply greater virtue discounting when judging the moral goodness of others' public virtuous actions compared to tendencies in how people expect their own public moral actions to be judged. Importantly, Studies 1-3 evaluate

self-enhancement in the separate evaluation context (i.e., between-subjects designs). This reflects how moral judgments often occur in everyday life (i.e., one target at a time, without direct comparison), but also contrasts with traditional self-enhancement literature, which often uses joint-evaluation designs (an approach we adopt in Study 4 to test the boundary conditions of these effects).

Study 2 deepens inquiry from Study 1 by investigating motivational attributions as a potential explanatory mechanism. Namely, we assess whether people expect their own public virtue motivations to be judged as more principled and less reputation-driven, on average, than how people tend to judge others' public virtue motivations, and examine how these motivational attributions relate to moral goodness judgments of the virtuous actions themselves. Study 3 replicates and extends Study 2 by addressing a key methodological limitation: prior studies compared meta-perceptions (how participants believe others judge them) to third-party judgments (how participants judge others), but did not measure self-perceptions (how participants judge their own virtue) or "second-order perceptions" (how participants believe others judge others). Study 3 fully crosses actor (self vs. other), observability (public vs. private), and observer perspective (self vs. other), revealing a pattern of results that largely replicates Study 2. Finally, in Study 4, we test whether the same effects observed in earlier studies emerge in a within-subjects joint-evaluation context, where the same people make both third-party judgments and meta-perceptions side by side, allowing direct comparison within individuals.

Beyond testing for self-enhancement in the context of virtue discounting, we also seek to replicate and extend both effects independently. Specifically, we assess whether self-serving bias extends to the separate evaluation context (Studies 1-3), meta-perceptions of moral character and behavior (Study 1), underlying motivations (Studies 2–4), and broader character judgments (Study S1). Additionally, we replicate prior findings on virtue discounting, which show that

public acts of virtue are perceived as lower in moral goodness due to attributions linking them to

reputational concerns rather than principled commitment (Kraft-Todd et al., 2023). Previous

research has demonstrated that observability influences moral goodness through shifts in

motivational attributions, with public virtue seen as more reputation-driven and less sincere than

private virtue. Here, we examine whether these attributional mechanisms extend beyond third-

party perceptions to meta-perceptions (Studies 2 and 4), as well as self-perceptions and second-

order perceptions (Study 3) across different actor and observer contexts. The data, code, and

survey materials for all studies are available on the Open Science Framework,

https://osf.io/ea58u/?view_only=b481f36a30544f3ea3f29a785f278a1e. Descriptive information

for all studies is provided in Table 2. Analyses were conducted in SAS 9.4. Multilevel mediation

models were estimated in Mplus 8.

Table 2

Descriptive information for each study.

	Study S1	Study 1	Study 2	Study 3	Study 4
Total N	163	457	903	901	250
Male	73	225	448	439	106
Female	87	225	436	438	142
Nonbinary/other	1	7	19	24	2
White	117	358	690	705	177
Black or African American	20	58	95	70	46
Asian or Asian American	16	26	63	57	25
Other racial/ethnic identity	19	15	55	69	17
Democrat/Liberal	64	206	430	434	53
Independent/Moderate	41	132	249	285	106
Republican/Conservative	53	109	197	152	81
M_{age}	44.81	42.59	42.32	39.23	42.94
SD_{age}	15.82	12.79	13.22	13.38	13.19

Study 1

Supplementary Study S1 (see SOM) demonstrated that people believe they are seen as

more moral than others, extending previous work studying self-enhancement in the context of

specific moral traits (e.g., Tappin & McKay, 2017) to *meta-perceptions* of broader moral character. In Study 1, we investigated whether self-serving asymmetries emerge in judgments of moral goodness for generous actions performed either publicly or privately (i.e., anonymously) using a separate evaluation context. Specifically, we tested whether participants expect their own virtuous acts to be judged more favorably (in meta-perceptions), on average, than a separate group of participants tend to judge others' acts (in third-party perceptions), and whether this asymmetry is especially pronounced in the public condition. If so, this would suggest that people tend to anticipate less virtue discounting for their own public actions than people tend to apply when evaluating others' public virtue.

Method

Participants

We recruited 457 participants via CloudResearch (Litman et al., 2016). A sensitivity analysis using G*Power (Faul et al., 2007) suggested that, for our sample size and design (i.e., a between-within subjects ANOVA), we could detect effect sizes as small as $f = .044$ with 80% power.

Materials and Procedure

Participants were randomly assigned to one of four conditions in a 2 *actor* (other vs. self) × 2 *observability* (public vs. private) design. Each participant read the same 17 vignettes, presented in a randomized order, depicting generous behaviors (e.g., "lend money to a friend"; Kraft-Todd et al., 2023). In the "other" condition, participants imagined someone else performing the act, while in the "self" condition, they imagined themselves doing so. Additionally, in the "public" condition, the action was visible to others, whereas in the "private" condition, it was performed anonymously. After each scenario, participants rated the behavior's moral goodness on a 1–7 Likert scale. Those in the "other" condition evaluated the morality of

another person's actions (i.e., third-party perceptions), while those in the "self" condition rated

how they believed others would judge their own actions (i.e., meta-perceptions).

Results

To examine whether there were differences based on the *actor* (self vs. other) and the

observability (public vs. anonymous) of generous actions, we conducted a mixed-ANOVA

accounting for one repeated-measures factor (accounting for the 17 vignettes), two between-

subjects factors (to account for the effects of actor and observability), and a between-subjects

interaction (observability × actor). Descriptives for each of the 17 behaviors across each of the

four conditions are displayed in Table 3.

Table 3

Descriptive statistics for virtue discounting by vignette split by observability and target agent.

| Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	5.57 (1.20)	5.86 (1.16)	4.94 (1.27)	4.44 (1.43)
Pay colleague a compliment	5.40 (1.11)	5.53 (1.10)	4.97 (1.25)	4.98 (1.36)
Hold door for someone	5.61 (1.15)	5.84 (0.98)	5.33 (1.12)	5.44 (1.13)
Fix someone's flat tire	6.29 (0.78)	6.30 (0.81)	5.91 (1.09)	5.75 (1.17)
Walk elderly person across street	6.18 (1.01)	6.35 (0.85)	6.10 (0.94)	5.85 (1.11)
Tip waiter generously	5.72 (0.97)	5.96 (1.05)	5.17 (1.20)	4.98 (1.29)
Buy groceries for someone in need	6.31 (0.90)	6.31 (0.98)	5.87 (1.09)	5.76 (1.06)
Buy food for homeless person	6.24 (0.91)	6.27 (0.99)	5.81 (1.00)	5.62 (1.22)
Donate car to charity	6.20 (0.94)	6.16 (1.01)	4.97 (1.29)	5.00 (1.25)
Mow friend's lawn	5.89 (1.03)	5.94 (1.12)	5.37 (1.17)	5.36 (1.24)
Walk friend's dog	5.55 (1.13)	5.64 (1.11)	5.00 (1.12)	4.97 (1.29)
Donate to an artist	5.61 (1.05)	5.66 (1.19)	4.71 (1.20)	4.81 (1.26)
Pick up trash	6.10 (0.94)	6.15 (0.98)	5.66 (1.13)	5.61 (1.12)
Donate to charity	6.19 (0.96)	6.25 (0.88)	5.07 (1.25)	5.03 (1.22)
Buy friend an expensive gift	5.27 (1.30)	5.56 (1.26)	4.42 (1.32)	4.39 (1.29)
Let someone in a rush skip the line	5.64 (1.16)	5.79 (1.06)	4.95 (1.42)	5.17 (1.46)
Help coworker after hours	5.94 (0.99)	5.99 (1.11)	5.81 (0.98)	5.71 (1.10)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Although ratings on the 17 vignettes differed significantly, $F(16, 7248)=85.64$, $p < .001$,

$\eta_p^2=.159$, they were positively correlated with one another (see Table S1 in the Supplementary

Materials). Intriguingly, no significant observability × actor interaction was observed, $F(1,$

$453)=1.46$, $p=.227$, $\eta_p^2=.003$, suggesting no self-serving attenuation of virtue discounting in the

context of meta-perceptions vs. third-party judgments. Further, no significant main effect of actor was observed, $F(1, 453)=0.07$, $p=.780$, $\eta_p^2=.0001$, suggesting a lack of a self-serving asymmetry altogether. However, a significant main effect of observability was observed, $F(1, 453)=79.61$, $p < .001$, $\eta_p^2=.149$; see Figure 2, whereby participants, on average, rated behaviors that were performed in public (regardless of the actor) as significantly less morally good compared to behaviors which were performed in private, replicating prior work demonstrating the phenomenon of virtue discounting (Kraft-Todd et al., 2023).

Figure 2

Average ratings for moral goodness across all vignettes *for the four conditions with 95% CI*
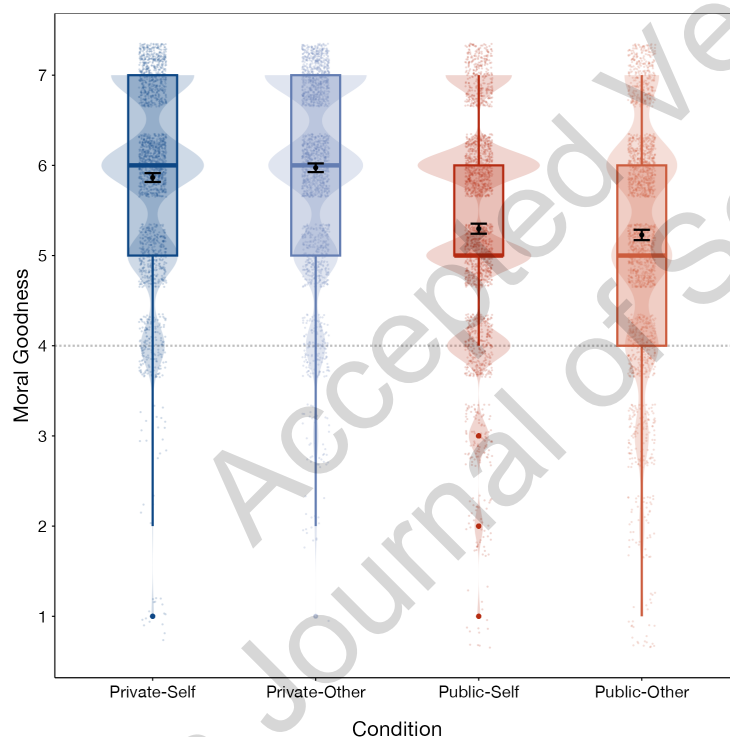
Discussion

Our first study suggested that, across 17 different acts of generosity, people, on average, consistently rated these acts as more morally good when they were performed anonymously rather than in public. No evidence was observed for a self-serving asymmetry in moral goodness

ratings of the actions, nor for a self-serving attenuation of virtue discounting for public acts of virtue when the actor was oneself versus someone else. In our subsequent studies, we shifted our focus to include judgments of the actor and their underlying motivations rather focusing narrowly on judgments of actions alone. That is, while Study 1 did not find evidence of self-enhancement for moral judgments of actions in the separate evaluation context, we examined whether self-serving asymmetries might emerge in judgments about the actors themselves—specifically, in attributions of their motivation and assessments of their trustworthiness. This approach allowed us to investigate whether differences in motivational attributions and trust assessments influence moral judgments and whether, in contexts where a self-serving asymmetry emerges for motivation and trust, it might also extend to evaluations of the action itself.

Study 2

While Study 1 did not find clear evidence of self-enhancement in moral judgments of virtuous actions, Study 2 reexamined this effect with a larger sample and expanded the investigation to include judgments of actors and their underlying motivations. Study 2 was pre-registered (https://aspredicted.org/blind.php?x=FHH_4NJ) and had three key objectives. First, as in Study 1, we tested whether a self-serving asymmetry would emerge in meta-perceptions versus third-party perceptions, extending the analysis beyond moral judgments of virtuous actions to actor-based judgments of motivational attributions and trustworthiness. We hypothesized that participants assigned to report meta-perceptions would, on balance, expect others to judge their actions as more morally good and see them as more motivated by principle and more trustworthy, relative to participants assigned to make third-party judgments, who would, on balance, evaluate the same actions as less morally good and see the actors as more motivated by reputation and less trustworthy.

Second, we sought to replicate Study 1 and prior findings on virtue discounting by testing whether moral judgments and motivational attributions of virtuous acts would vary based on observability (Kraft-Todd et al., 2023), and whether differences in motivational attributions would predict judgments of moral goodness. Finally, and perhaps most critically, we investigated the interaction between actor (self vs. other) and observability (public vs. private). We hypothesized that a self-serving asymmetry would attenuate virtue discounting in the 'self' condition, leading individuals to expect more favorable moral judgments, motivational attributions, and trust assessments for their own public virtuous actions compared to how individuals in the 'other' condition judge, attribute motivation, and assess trust for the public virtuous actions of others.

Method

Participants

We recruited 903 participants via CloudResearch. Demographic information for this sample is provided in Table 1. An a priori power analysis for a repeated measures ANOVA with between-subjects factors, for four groups (2 private vs. public × 2 self vs. other), five measurements (5 randomly selected vignettes out of 17), 90% power, correlation between measurements of $r = 0.50$, and a small effect size $f = 0.10$, suggested a sample of 856 participants.

Materials and Procedure

Participants were randomly assigned to one of four conditions. In each condition, a random assortment of 5 of the same 17 vignettes displayed in Study 1 were presented in a randomized order. The process was otherwise identical to Study 1, with the only differences being the addition of new measures that were presented after each vignette (i.e., after moral goodness was measured) to capture correlates of virtue discounting identified in prior research

(e.g., motivational attributions, trust, normative expectations; Kodipady et al., 2022; Kraft-Todd

et al., 2023). All measures were captured on a 1-7 Likert scale (1=strongly disagree - 7=strongly

agree). See Table 4 for a full breakdown of the measures.

Table 4

Item information for all outcomes in Study 2

Outcome	Other Conditions	Self Conditions
	What do you think of the person's behavior?	*What do you think others would think of your behavior?*
Moral Goodness	1. This person's behavior is morally good	1. My behavior is morally good
Trust	1. I trust this person	1. They would trust me
	To what extent do you think that this person engaged in this behavior because...	*Now consider the degree to which others would think that you engaged in this behavior because...*
Principled Motivation (Cronbach's *a*=0.82)	1. They truly wanted to help others 2. They are a virtuous person 3. This behavior reflects their values	1. You truly wanted to help others 2. You are a virtuous person 3. This behavior reflects your values
Reputation Signaling (Cronbach's *a*=0.89)	1. They wanted to look more generous than they are in reality 2. They wanted to gain status 3. They wanted to enhance their reputation 4. They wanted to attract attention 5. They wanted to feel good about themselves	1. You wanted to look more generous than you are in reality 2. You wanted to gain status 3. You wanted to enhance your reputation 4. Your wanted to attract attention 5. You wanted to feel good about yourself
Normative Expectations (Cronbach's *a*=0.64)	1. They think it is common to behave this way 2. They believe that others think people should behave this way	1. You think it is common to behave this way 2. You believe that others think people should behave this way

Results

Condition Differences

Moral goodness and trust. On average, public acts of generosity were rated as less

morally good, and public actors as less trustworthy, than private acts and actors (see Table 5 for

statistics and Figure 3 for visualization). However, while no differences emerged in meta-

perceptions versus third-party perceptions for private acts, a self-serving asymmetry was

observed in public contexts: 'self' condition participants expected others to judge their own

public actions as more morally good on average than 'other' condition participants judged

others' public actions.

Motivational attributions. On average, public actions were perceived as less driven by

principled motives and more by reputation signaling compared to private actions. Again, no

differences emerged in meta-perceptions versus third-party perceptions for private acts.

However, in public contexts, 'self' condition participants expected others to see their own actions as more principled and less reputation-driven on average than 'other' condition participants judged others' public actions—further evidence of a self-serving asymmetry.

Normative expectations. In public contexts, 'self' condition participants expected others to see their own acts of generosity as less normative (i.e., more unique) on average than 'other' condition participants judged others' public acts. In contrast, for private acts, the pattern reversed. 'Self' condition participants expected their own behaviors to be seen as more normative (i.e., more typical) on average than 'other' condition participants judged others' private behaviors.

Collectively, these findings suggest that self-enhancement extends across multiple dimensions of moral evaluation in public settings, reducing virtue discounting when oneself is the actor while skepticism toward others' public virtue remains high.

Table 5

Multilevel regression model testing the effect of observability (public vs. private) and actor (self vs other) on the five outcomes across the 17 scenarios. Participants were only shown 5 randomly selected scenarios at a time.

Parameter	Moral Goodness Estimate (*SE*)	Trust Estimate (*SE*)	Principled Motivation Estimate (*SE*)	Reputation Signaling Estimate (*SE*)	Normative Expectation Estimate (*SE*)
Intercept (Private, Other)	5.94*** (.05)	5.46*** (.06)	5.73*** (.05)	2.92*** (.07)	4.72*** (.06)
Setting: Public=1 (Private=0)	-0.64*** (.08)	-.64*** (.09)	-.66*** (.07)	1.49*** (.10)	.16* (.08)
Actor: Self=1 (Other=0)	-0.10 (.08)	0.06 (.09)	-0.10 (.07)	.22* (.10)	.20* (.08)
Interaction (Public, Self)	.34** (.11)	.29* (.12)	.30** (.10)	-.55*** (.14)	-.35** (.11)

Note. Scores can be interpreted as differences relative to the intercept.
*p < .05, **p < .01, ***p < .001.

Figure 3

*Ratings for all outcomes in Study 2 by observability and agent with 95% CI*
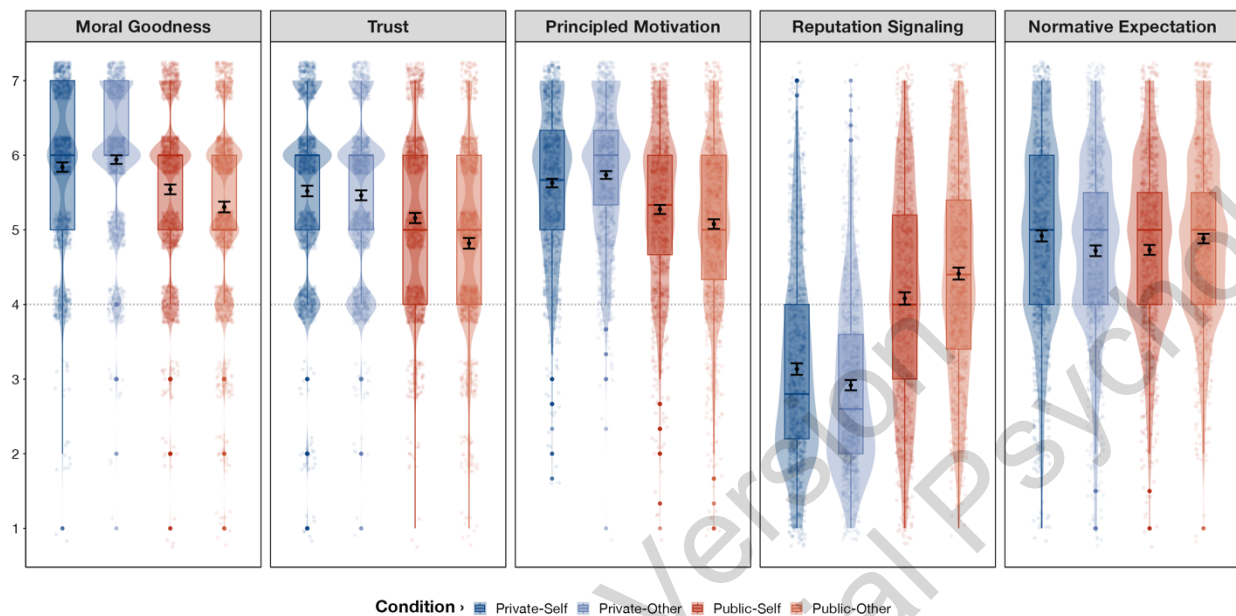
Indirect Effects

  We estimated a 2-1-1 multilevel mediation model in Mplus using a Bayesian Estimator to examine whether trust, principled motivation, reputation signaling, and normative expectations mediated the effect of the manipulation on moral goodness ratings. Instead of *p*-values, 95% Credible Intervals (Cr.I.) were used, with statistical significance determined by intervals non-inclusive of zero. The model included two dummy-coded predictors: actor (self=0, other=1) and observability (private=0, public=1), with the four mechanisms as parallel mediators and moral goodness as the outcome. Prior research on moral perceptions of virtue suggests that principled motivation, reputation signaling, and normative expectations influence moral judgments (Kraft-Todd et al., 2023), while trust plays a key role in perceptions of morality more generally (Evans & van de Calseyde, 2018). All associations between the mediators and moral goodness were estimated as random slopes.

Results showed significant differences for trust, principled motivation, and reputation signaling. Public acts, regardless of the actor, were perceived as less principled, more reputation-driven, and less trustworthy than private acts. Even after accounting for these mediators, public acts were rated as less morally good. Trust, principled motivation, and normative expectations positively predicted moral goodness, while reputation signaling negatively predicted it (see Figure 4). All indirect effects were significant except for those via normative expectations (see Table 6).

Figure 4

Multilevel mediation (2-1-1) testing the indirect effect of observability (public=1, private=0) for ratings of others, and meta-perceptions of others' ratings, on moral goodness, via the four proposed mechanisms.
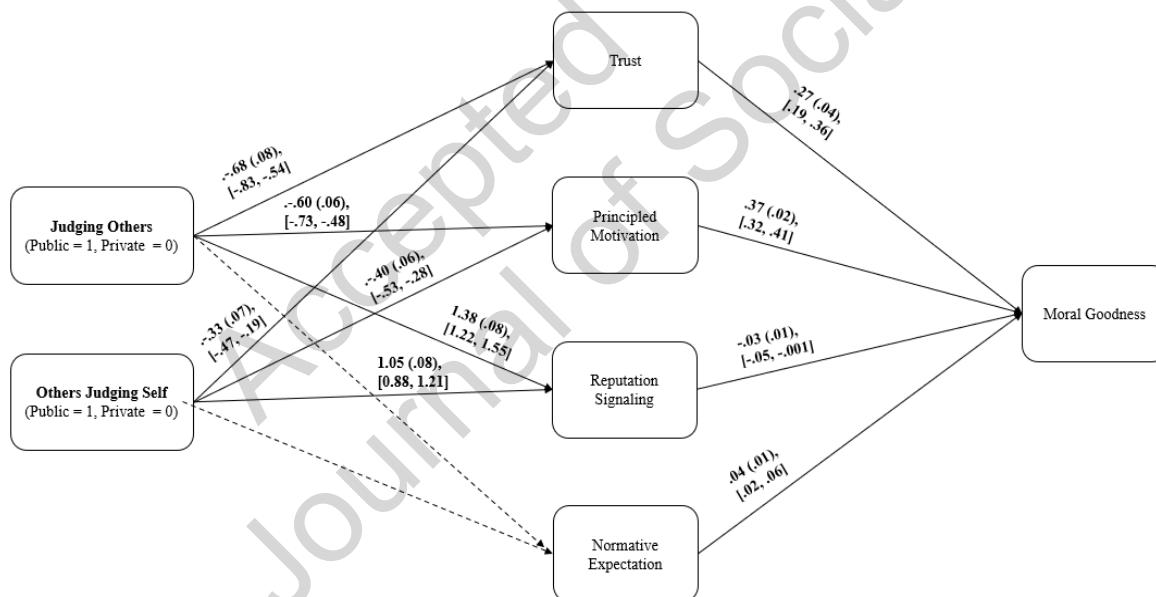
Note. The Bayesian Estimator was used, and findings are displayed as *b*, (*SE*), [95% Credible Interval]. Values that do not include 0 are significant, and non-significant effects are not displayed. The main effects of the two dummy-coded predictors on ratings of moral goodness were significant: judging others for public actions (compared to private): *b*=-.58, *SE*=.07, 95% Cr.I. [-.72, -.45]; judging the self for public actions (compared to private): *b*=-.34, *SE*=.07, 95% Cr.I. [-.72, -.45].

Table 6

Indirect effects tested in Figure 3.

			95% Cr. I.	
Judging the actions of others	*b*	*SE*	Lower	Upper
Public (= 1) → Trust → Moral Goodness	**-0.28**	**0.03**	**-0.35**	**-0.22**
Public (= 1) → Principled Motivation → Moral Goodness	**-0.22**	**0.03**	**-0.28**	**-0.17**
Public (= 1) → Reputation Signaling → Moral Goodness	**-0.04**	**0.02**	**-0.07**	**-0.001**
Public (= 1) → Normative Expectation → Moral Goodness	0.002	0.003	-0.003	0.009
Perceptions of others' judging participants actions	*b*	*SE*	Lower	Upper
Public (= 1) → Trust → Moral Goodness	**-0.14**	**0.03**	**-0.20**	**-0.08**
Public (= 1) → Principled Motivation → Moral Goodness	**-0.15**	**0.02**	**-0.20**	**-0.10**
Public (= 1) → Reputation Signaling → Moral Goodness	**-0.03**	**0.01**	**-0.05**	**-0.001**
Public (= 1) → Normative Expectation → Moral Goodness	-0.003	0.003	-0.01	0.002

Note. Cr. I.=Credible Interval. Bolded effects are significant (Cr. I. does not include 0).

Discussion

Our second study noted several important insights. First, as previous research has illustrated (e.g., Kraft-Todd et al., 2023), public generous actions are subject to virtue discounting, a finding emerging both for the moral goodness of the action itself, and the motivations behind it. Second, instances of a self-serving asymmetry were noted, but particularly for generous actions conducted in public, suggesting that self-enhancement attenuates virtue discounting when the actor is oneself. This attenuation emerged for moral goodness, principled motivation, reputation signaling, and trust. In line with prior research (Kraft-Todd et al., 2023), principled motivation, reputation signaling, and trust were significant predictors of discounting the moral goodness of virtuous actions. Similarly, normative expectations were positively correlated with ratings of goodness, aligning with prior findings suggesting that when public virtuous behaviors are perceived as more common, they are less likely to be seen as reputation-driven (Kodipady et al., 2022). Surprisingly, however, participants in the 'self' condition expected their own *public* virtuous behaviors to be judged as less normative on average than participants in the 'other condition' judged others', while the opposite asymmetry manifested in

the context of *private* behaviors (i.e., participants in the 'self' condition expected their own

private virtuous behaviors to be judged as more normative than participants in the 'other'

condition judged others'). As we elaborate in the General Discussion, one possible explanation is

that people expect observers to view their public virtue as optimally distinct, striking a balance

between uniqueness and social fit (Brewer, 1991).

Study 3

In Study 3, we sought to replicate and extend Study 2 by addressing a methodological

limitation: prior studies compared meta-perceptions (how participants believe others judge them)

to third-party judgments (how participants judge others), but they did not measure self-

perceptions (how participants judge their own virtue) or "second-order perceptions" (how

participants believe others judge others). Meanwhile, typical research on self-enhancement has

generally compared self-perceptions to third-party judgments (e.g., Tappin & McKay, 2017;

Sedikides et al., 2014). In Study 3, we fully-crossed actor (self vs. other), observability (public

vs. private), and observer (self vs. other) in a 2 × 2 × 2 design to test the robustness of the effects

observed in Study 2. If the same pattern of results emerges—where participants, on average,

expect less virtue discounting for their own public acts in both self-perceptions and meta-

perceptions, but greater virtue discounting in third-party and second-order perceptions—it

strengthens the argument that this asymmetry is not merely an artifact of comparing meta-

perceptions to third-party judgments. Instead, it would suggest that people tend to both see their

own public virtue as morally superior and expect others to recognize it as such, while

simultaneously discounting others' public virtue and assuming others will do the same. This

study was pre-registered, https://aspredicted.org/V5D_8MY.

Method

Participants

We recruited 901 participants via Prolific. Demographic information for this sample is provided in Table 2. Our sample size was determined using the same rationale as Study 2 and was adequately powered for our design.

Materials and Procedure

A 2 × 2 × 2 between-subjects design with five repeated measures was implemented. Participants were randomly assigned to one of eight conditions, in which they were presented with a randomly selected set of five vignettes (out of the same 17 used in Study 1), displayed in a randomized order. The procedure was otherwise identical to Study 2, with one key difference: the vignettes now varied along three dimensions: the observability of the generous action (public vs. private), the actor performing the action (self vs. other), and the observer making the judgment (self vs. other).

All measures were captured on a 1–7 Likert scale, with labels ranging from 1=strongly disagree to 7=strongly agree. All measures were identical to Study 2. Reliability was high for principled motivation (Cronbach's *a*=.82), reputation signaling (Cronbach's *a*=.89) and adequate for normative expectations (Cronbach's *a*=.68).

Results

Condition Differences

Following our pre-registered analytical plan, we first estimated multilevel regression models including the three-way interaction between observability (public/private), actor (self/other), and observer (self/other), along with all two-way interactions and main effects. No significant three-way interactions emerged for moral goodness, trust, principled motivation, or reputation signaling. While a statistically significant three-way interaction effect was found for normative expectations (*b*=-0.49, *SE*=0.25, *p*=.049), none of the corresponding two-way interactions was significant, and only three of the 15 possible interaction terms reached

significance, each for a different outcome. Given these inconsistent findings, and in line with our pre-registered analysis plan, we proceeded as if no significant three-way interactions had been detected. As such, we collapsed across observer conditions (self vs. other) and examined the interaction between observability (public vs. private) and actor (self vs. other), following the analytical approach used in Study 2. The results below (see Table 7 and Figure 5) reflect this simplified model, while full analyses, including the three-way interaction terms, are reported in the SOM.

Moral goodness. On average, private acts were rated as more morally good than public acts. No differences emerged for public acts, but private acts performed by participants were rated as less morally good on average than those performed by others.

Trust, principled motivation, and reputation signaling. Results fully replicated Study 2, showing the same patterns across these measures.

Normative expectations. Findings largely mirrored Study 2, with one difference: for private actions, participants, on average, rated their own behavior *and* expected their own behavior to be rated by others less normative than others' behavior.

Table 7

Multilevel regression model testing the effect of context (public vs. private) and agent (self vs other) collapsed across observers (self vs other) on the five outcomes across the 17 scenarios. Participants were only shown 5 randomly selected scenarios at a time.

	Moral Goodness	Trust	Principled Motivation	Reputation Signaling	Normative Expectation
	Estimate (*SE*)	Estimate (*SE*)	Estimate (*SE*)	Estimate (*SE*)	Estimate (*SE*)
Intercept (Private, Other)	5.89*** (.06)	5.38*** (.06)	5.65*** (.05)	2.85*** (.07)	4.85*** (.06)
Setting: Public=1 (Private=0)	-.50*** (.08)	-.52*** (.09)	-.53*** (.07)	1.38*** (.10)	.15 (.09)
Actor: Self=1 (Other=0)	-.22** (.08)	.15 (.09)	-.14 (.07)	-.05 (.10)	-.21* (.09)

Interaction	.23*	.44**	.28**	-.53***	-.29*
(Public, Self)	(.11)	(.13)	(.11)	(.14)	(.13)

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

Figure 5

Ratings for all outcomes in Study 3 by observability and agent (collapsed across observer)

Condition › ■ Private-Self ■ Private-Other ■ Public-Self ■ Public-Other
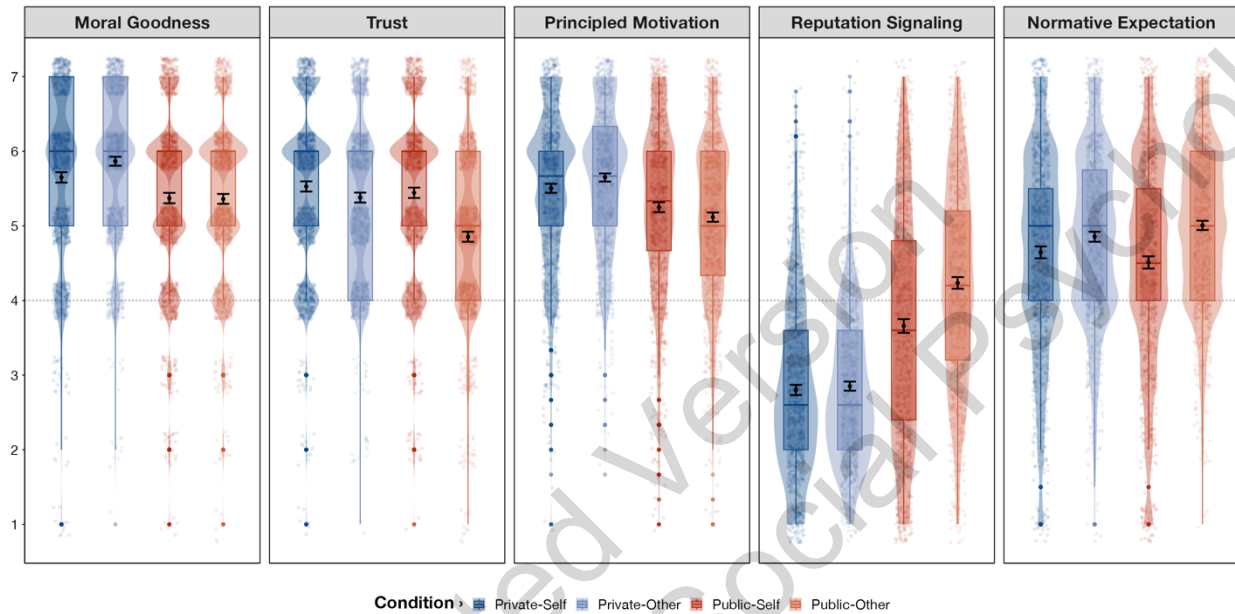
Indirect Effects

We estimated a similar model to that of Study 2, collapsing across observer conditions. Overall, significant differences were observed for all mediators for the public versus private comparison when the agent was somebody else. When the agent was the self, no significant difference was observed for trust. Regardless of the actor, public generous actions were seen as less morally good, even after accounting for the association of the four mediators with ratings of moral goodness. In turn, all four mediators significantly correlated with ratings of moral goodness, such that trust, principled motivation, and normative expectation correlated positively with ratings of moral goodness, while reputation signaling correlated negatively (see Figure 6). Importantly, all indirect effects except for those for trust when the agent was the participant significant (see Table 8).

Figure 6

Multilevel mediation (2-1-1) testing the indirect effect of condition (public=1, private=0) for ratings of others, on moral goodness, via the four proposed mechanisms, collapsed across observers.
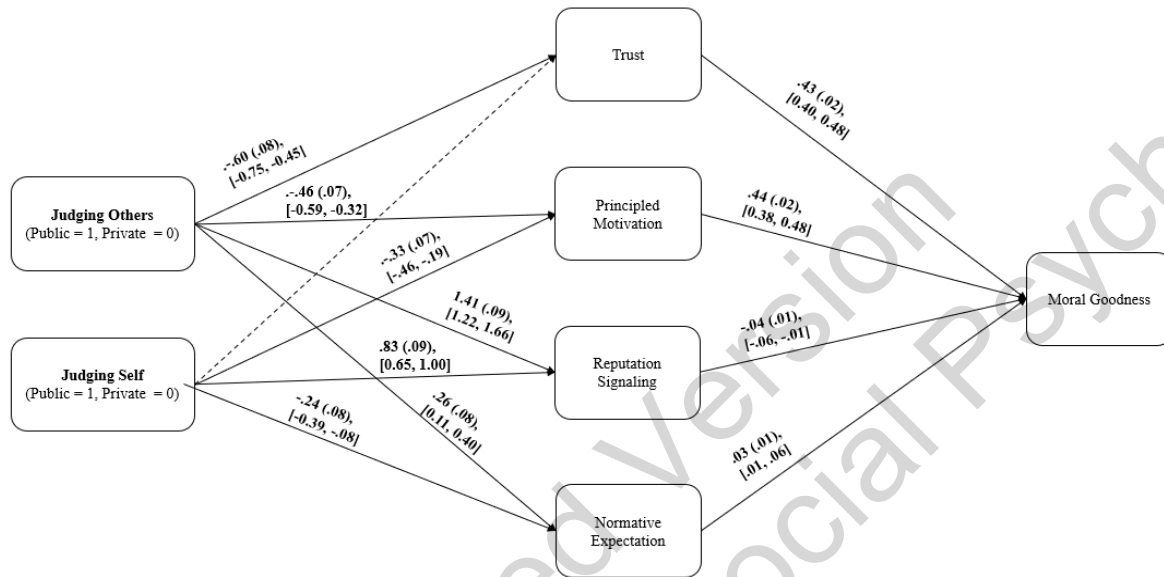
Note. The Bayesian Estimator was used, and findings are displayed as *b*, (*SE*), [95% Credible Interval]. Values that do not include 0 are significant, and non-significant effects are not displayed. The main effects of the two dummy-coded predictors on ratings of moral goodness were significant: judging others for public actions (compared to private): *b*=-.39, *SE*=.07, 95% Cr.I. [-.53, -.26]; judging the self for public actions (compared to private): *b*=-.39, *SE*=.07, 95% Cr.I. [-.52, -.25].

Table 8

Indirect effects tested in Figure 5.

	b	SE	95% Cr. I.	
			Lower	Upper
Judging the actions of self	*b*	*SE*	Lower	Upper
Public (= 1) → Trust → Moral Goodness	-0.01	0.03	-0.07	0.06
Public (= 1) → Principled Motivation → Moral Goodness	**-0.14**	**0.03**	**-0.20**	**-0.08**
Public (= 1) → Reputation Signaling → Moral Goodness	**-0.03**	**0.01**	**-0.06**	**-0.01**
Public (= 1) → Normative Expectation → Moral Goodness	**-0.01**	**0.004**	**-0.02**	**-0.001**
Judging the actions for others	*b*	*SE*	Lower	Upper
Public (= 1) → Trust → Moral Goodness	**-0.26**	**0.03**	**-0.32**	**-0.19**
Public (= 1) → Principled Motivation → Moral Goodness	**-0.20**	**0.03**	**-0.26**	**-0.14**
Public (= 1) → Reputation Signaling → Moral Goodness	**-0.06**	**0.02**	**-0.09**	**-0.02**
Public (= 1) → Normative Expectation → Moral Goodness	**0.01**	**0.004**	**0.002**	**0.02**

Note. Cr. I.=Credible Interval. Bolded effects are significant (Cr. I. does not include 0).

Discussion

Our third study replicated the findings of Study 2. First, the phenomenon of virtue discounting was replicated, as the moral goodness of actions themselves, and the perceived motives behind them, were rated significantly higher, more principled, and less reputation-driven in the private conditions. In turn, judging the actor's motives as more principled, more consistent with normative expectations, and the actor as more trustworthy, predicted higher moral goodness ratings. Meanwhile, attributing the action to reputational motives predicted lower moral goodness. Again, instances of a self-serving asymmetry were noted, but particularly for generous actions conducted in public, suggesting attenuation of the virtue discounting phenomenon consistent with self-enhancement. This attenuation emerged for principled motivation, reputation signaling, and trust, but not for moral goodness. In line with results from Study 2, participants in conditions where the actor was oneself expected their own public virtuous behaviors to be judged as less normative than participants judged others in the conditions where the actor was someone else. Here, however, the opposite asymmetry in the context of *private* behaviors was *not* noted.

Study 4

In our final pre-registered study (https://aspredicted.org/wfzx-rj2z.pdf), we tested for self-serving asymmetries in perceptions of virtue by asking participants to simultaneously evaluate the moral goodness and underlying motives of a generous action (i.e., donating to charity) from either a meta-perception perspective (how they believed others would judge *them*) or a third-party perspective (how they judged someone else), within a joint evaluation context (Hsee et al., 1996). We also sought to replicate the broader virtue discounting phenomenon in a fully within-subjects design. Whereas our earlier studies used between-subjects "separate evaluation" designs, which mirror everyday contexts in which people encounter a single moral target in isolation (e.g., scrolling past a charitable post online), Study 4 employed a within-subjects "joint

evaluation" design, which more closely resembles contexts where direct comparisons are salient (e.g., weighing one's own behavior against another's side by side). This distinction is both methodological and psychological: separate evaluation highlights whichever cues are most salient in the moment (like publicity), while joint evaluation encourages contrast and calibration across targets. Thus, we hypothesized that generous behavior performed in public (i.e., observable by others) would be judged as less morally good, and the motives behind it as less principled and more reputation-driven, compared to the same behavior performed in private, regardless of whether participants were evaluating someone else's actions or imagining how others would evaluate their own. Moreover, we predicted a negative relationship between moral goodness judgments of generous behavior and perceptions of reputation motivation, but a positive relationship between moral goodness judgments of generous behavior and perceptions of principled motivation.

Finally, based on the between-subjects findings from Studies 2 and 3, we hypothesized that in public contexts, participants would expect their own generous behavior to be judged by others as more morally good, more motivated by virtue, and less reputation-driven than participants judge others' generous behavior and motivations. In private contexts, we tested whether a similar self-serving asymmetry would emerge. Although Studies 2 and 3 found markedly stronger self-enhancement effects in public conditions, prior research has often demonstrated self-enhancing biases in within-subjects designs, particularly when people evaluate themselves and others side by side (e.g., Tappin & McKay, 2017; Sedikides et al., 2014) Thus, we included parallel predictions for private behavior to test whether these asymmetries would also appear in private contexts when both targets were evaluated jointly.

Methods

Participant

We recruited 250 participants via Prolific. Demographic information for this sample is provided in Table 1. Our sample size was determined based on an *a priori* power analysis for a within-subject design with four measurements specifying 80% power to detect an effect size of F=.10 in a repeated-measures ANOVA.

Materials and Procedure

Participants read a fictional scenario involving a generous act (donating to a charity), which varied along two dimensions: whether the act occurred in public or private, and whether the actor was the participant or another person. Each participant viewed four vignettes in total, one for each combination of context and actor. The order of public vs. private vignettes was counterbalanced, and the order of self vs. other was randomized within each context. For example, a participant might see the two public vignettes first (one with themselves and one with another person as the agent), followed by the two private vignettes, or vice versa. All measures were rated on 1–7 Likert scales (1 = strongly disagree, 7 = strongly agree). To reduce survey length, cost, and participant fatigue, we limited the number of constructs and items included (see Table 9).

Table 9

List of items used in Study 4

Outcome	Other Conditions	Self Conditions
	What do you think of the person's behavior?	*What do you think others would think of your behavior?*
Moral Goodness	1. This person's behavior is morally good	1. My behavior is morally good
	To what extent do you think that this person engaged in this behavior because...	*Now consider the degree to which others would think that you engaged in this behavior because...*
Principled Motivation (($a_{Public\ Self}$ = 0.83; $a_{Public\ Other}$ = 0.84; $a_{Private\ Self}$ = 0.74; $a_{Private\ Other}$ = 0.80)	1. They truly wanted to help others 2. They are a virtuous person 3. This behavior reflects their values	1. You truly wanted to help others 2. You are a virtuous person 3. This behavior reflects your values

Reputation Signaling ($a_{Public\ Self}$ = 0.94; $a_{Public\ Other}$ = 0.94; $a_{Private\ Self}$ = 0.92; $a_{Private\ Other}$ = 0.94)	1. They wanted to gain status 2. They wanted to enhance their reputation 3. They wanted to attract attention	1. You wanted to gain status 2. You wanted to enhance your reputation 3. You wanted to attract attention

Results

As indicated in Table 10 significant differences across all conditions were noted. Visual inspection of the results indicates that this difference was driven primarily by differences between public and private behaviors rather than the actors (self vs. other).

Table 10

Results for one-way repeated measures ANOVAs for each outcome

Outcome	ANOVA	Public Other		Public Self		Private Other		Private Self	
		M	SD	M	SD	M	SD	M	SD
Moral Goodness	$F(3,741) = 116.97^{***}$, $\eta^2_p = .321$	5.16	1.34	5.16	1.40	6.25	0.86	6.22	0.90
Principled Motivation	$F(3,747) = 114.32^{***}$, $\eta^2_p = .315$	5.09	1.16	5.15	1.12	6.10	0.87	6.08	0.91
Reputation Motivation	$F(3,747) = 497.96^{***}$, $\eta^2_p = .666$	4.98	1.47	4.99	1.46	2.04	1.31	2.07	1.35

Pre-registered planned contrasts supported these observations. Specifically, significant differences emerged in the expected direction for each public vs. private comparison. However, none of the predicted self vs. other differences reached significance (see Table 11). These results suggest that the self-serving asymmetry in perceived virtue motives observed in Studies 2 and 3 under between-subjects designs disappears when participants make explicit comparisons between themselves and another person in a joint-evaluation context. We further interpret this finding in the discussion of this study and in the General Discussion.

Table 11

Pre-registered planned contrasts

Hypothesis	Outcome	Expected Direction of Planned Contrast	Result
H1	Moral Goodness	Public Other > Private Other	$F(1,249) = 137.74$, $p < .001$, $\eta2p = .356$
H1	Moral Goodness	Public Self > Private Self	$F(1,249) = 129.96$, $p < .001$, $\eta2p = .343$
H3a	Principled Motivation	Public Other < Private Other	$F(1,249) = 139.31$, $p < .001$, $\eta2p = .359$
H3a	Principled Motivation	Public Self < Private Self	$F(1,249) = 136.49$, $p < .001$, $\eta2p = .354$
H3b	Reputation Motivation	Public Other > Private Other	$F(1,249) = 589.91$, $p < .001$, $\eta2p = .703$
H3b	Reputation Motivation	Public Self > Private Self	$F(1,249) = 561.97$, $p < .001$, $\eta2p = .693$

H4a	Moral Goodness	Public Self > Public Other	$F(1,249) = 0.00$, p = .999, $\eta2p < .001$
H4a	Moral Goodness	Private Self > Private Other	$F(1,249) = 0.00$, p = .999, $\eta2p < .001$
H4b	Principled Motivation	Public Self > Public Other	$F(1,249) = 1.06$, p = .304, $\eta2p = .004$
H4b	Principled Motivation	Private Self > Private Other	$F(1,249) = 0.24$, p = .623, $\eta2p = .001$
H4c	Reputation Motivation	Public Self <Public Other	$F(1,249) = 0.46$, p = .844, $\eta2p = .002$
H4c	Reputation Motivation	Private Self < Private Other	$F(1,249) = 0.20$, p = .655, $\eta2p = .001$

Finally, as predicted, regardless of the actor or setting, perceived principled (positively)

and reputational motives (negatively) correlated with action ratings of moral goodness (see Table

12). Interestingly, principled motives were significantly more predictive of moral goodness in

meta-perceptions than in third-party judgments (public context: $Z = 2.48$, $p = .013$; private

context: $Z = 2.03$, $p = .042$), suggesting that participants believed others would place greater

moral weight on their principled motivations than they themselves placed on others'. In contrast,

the predictive strength of reputational motives did not differ significantly between meta-

perceptions and third-party judgments (public: $Z = 0.26$, $p = .795$; private: $Z = –0.15$, $p = .881$).

This analysis was exploratory.

Table 12

Correlations and linear regressions (with both measures as simultaneous predictors) for the two

motives and moral goodness (outcome)›

	Principled Motivation		Reputation Motivation	
Vignette	r	β	r	β
Public Other ($R^2 = .49$)	.69***	0.65*** (.06)	-.35***	-.11* (.04)
Public Self ($R^2 = .64$)	.79***	0.75*** (.05)	-.37***	-.12** (.04)
Private Other ($R^2 = .42$)	.62***	.49*** (.05)	-.49***	-.25*** (.04)
Private Self ($R^2 = .54$)	.72***	.63*** (.05)	-.48***	-.19*** (.03)

Discussion

Our final study re-evaluated self-serving asymmetry and virtue discounting using a fully

within-subjects joint evaluation design. The results were clear: we found consistent support for

the virtue discounting effect (for the first time in a within-subjects design), but no evidence of

self-enhancement in meta-perceptions. One possibility is that when people consider only how

others might judge *them*, they assume more favorable evaluations than others actually give, but when given the opportunity to compare these expectations directly to how they themselves judge others, this comparative framing may dampen bias. At the same time, comparing public and private virtue within the same context may heighten sensitivity to reputational cues, making virtue signaling feel more apparent and virtue discounting more pronounced. Still, our findings show that even when imagining how others would evaluate their own behavior, participants expected their public virtue to be seen as less principled, more reputation-driven, and ultimately less morally good than the same behavior performed in private, and that self-serving attenuation of this effect is less likely in the joint evaluation context.

General Discussion

Across four primary and one supplementary study, we examined how virtue discounting (Kraft-Todd et al., 2023) and self-serving asymmetries (Alicke et al., 1995; Brown, 1986; Dunning et al., 1989) interact to shape judgments and attributional motivations of public versus private acts of virtue. We found that public acts of generosity were generally subject to virtue discounting in both judgments of moral value—where generous actions performed in public versus private were rated as less morally good (Studies 1–4)—and in motivational attributions, with public (vs. private) acts perceived as more reputation-driven and less principled (Studies 2–4). Similar virtue discounting effects were observed in judgments of trust, with public (vs. private) generous actors rated as less trustworthy (Studies 2–3).

Intriguingly, however, virtue discounting effects on motivational attributions and trust were consistently attenuated by a self-serving bias in meta-perceptions in between subjects-separate evaluation contexts. That is, participants, on average, expected their own public acts of virtue to be judged more favorably (i.e., as more principled, less reputation-driven, and more trustworthy) than participants tended to judge others' public acts of virtue (Studies 2-3).

Moreover, Study 3 demonstrated that this pattern of self-serving attenuation in virtue discounting extended beyond meta-perceptions to self-perceptions (i.e., how participants judged their own public acts of virtue) and was mirrored by a complementary effect in second-order perceptions, where participants expected others to discount the virtue of third-party actors just as they did. Together, these findings suggest an asymmetry in how individuals evaluate public virtue when the actor is oneself versus someone else—applying and anticipating leniency for their own public moral actions while applying, and expecting others to apply, greater skepticism toward others' (see Table 1 for a summary of all asymmetries found in the context of public virtue across Studies and Figure 1 for a visualization of each judgment context).

While participants in separate evaluation contexts expected greater leniency for their own public virtue in terms of motivational attributions and trust, their action-based judgments of moral goodness did not consistently reflect the same self-enhancing pattern (such a self-serving attenuation effect was only realized in Study 2). This inconsistency suggests that, in the context of moral judgments, virtue discounting tendencies may at times override self-serving biases. In other words, individuals may recognize the broader norm of skepticism toward public virtue and apply it even when morally evaluating their own behavior (Aristotle, 1998; Cokelet & Fowers, 2019). Unlike attributions of principled motivation and trust which may allow for more subjective interpretation, moral goodness judgments appear to be more rigidly shaped by the dominant tendency to discount public acts of generosity. Future research should explore whether self-serving biases in moral evaluation emerge under conditions where virtue discounting pressures are weaker, such as when acts of generosity are seen as particularly costly (e.g., extraordinary acts of altruism like living organ donation or extreme instances of philanthropy; Law et al., 2023; Marsh, 2019).

Another intriguing pattern of results was observed for normative expectations, beliefs about whether a virtuous behavior is more or less statistically common. Our findings suggest that normative expectations positively predict moral goodness, but their indirect effects varied in Study 3 depending on whether the participant was the actor. Participants tended to expect their own public virtue to be judged as morally good, yet they also tended to see their actions as less normative than others', complicating a straightforward self-enhancement—at least with respect to perceptions of normativity/uniqueness in particular. Rather than simply believing they are more moral than others in a general sense, participants perceived their own moral behaviors as outstanding in a distinct and meaningful way, a pattern consistent with optimal distinctiveness theory (Brewer, 1991). Kodipady et al. (2022) found that public virtue is less likely to be perceived as reputation-driven when it is seen as common, but here, participants rated their own public virtue as less normative while still expecting it to be judged favorably. This suggests that people may enhance their self-concept not only by assuming greater morality overall, but by viewing their own virtuous acts as exceptional rather than typical.

Study 4 offered insight into a boundary condition, suggesting a circumstance under which self-serving asymmetries might fail to emerge. Unlike the separate evaluation contexts used in prior studies, Study 4 employed a joint evaluation design in which participants simultaneously reported how they believed others would judge their own public behavior and how they themselves judged the identical behavior in others. In this setting, the self-serving asymmetries observed in Studies 2 and 3 disappeared entirely. Participants expected their own public acts to be judged no differently than others': They anticipated just as much skepticism toward their own public virtue as they applied to others. These expectations also replicated the broader virtue discounting effect: public acts, whether their own or others', were seen as less morally good, less principled, and more reputation-driven than private acts.

As discussed in the introduction, joint evaluation contexts (Hsee et al., 1996; see also McManus et al., 2023) tend to heighten contrast and bring evaluative differences into sharper focus. In other domains, this shift in evaluability has been shown to reverse or eliminate otherwise robust effects: for instance, people often prefer to help one identifiable victim over many in separate evaluation, but become sensitive to scope and choose to help more people when both options are presented side by side (Kogut & Ritov, 2005). Similarly, in Study 4, directly comparing public and private virtue may have made reputational cues more salient, or made participants more aware of their own tendency to apply evaluative standards unevenly, amplifying virtue discounting across both targets. However, without a neutral or baseline control condition, it remains unclear whether the disappearance of asymmetry in Study 4 reflects more charitable judgments of others, more critical expectations of how one will be judged, or a diminished motivation to self-enhance. Nonetheless, the findings suggest that when comparison is made explicit, rather than unfolding intuitively as it often does in everyday moral cognition, people may extend virtue discounting to the self as readily as to others.

Implications and Future Directions

The present findings make multiple contributions to existing theoretical knowledge at the intersection of social judgment, moral psychology, and ethical philosophy. While prior work (Kraft-Todd et al., 2023) has established that public virtue is often met with moral skepticism, we demonstrate that this effect is not uniformly applied but rather applied in a generally self-serving manner, at least when participants make judgments in naturalistic separate evaluation contexts. This suggests that virtue discounting can operate alongside a self-enhancement bias, shaping how individuals judge virtue in others, how they anticipate their own virtue to be judged, how they judge their own virtue, and how they expect others to judge virtue in others. Moreover, our findings replicate prior inquiry into self-serving asymmetries (Sedikides et al., 2014; Tappin

& McKay, 2017; Taylor et al., 2021) and extend knowledge in this domain beyond self-perceptions of moral character to meta-perceptions and judgments of moral behavior. In particular, we find that people not only believe they possess more elevated moral character traits than others, but assume their public virtuous actions are and appear to be more principled than those of others'. Further, while prior work has primarily evaluated self-enhancement in joint evaluation contexts, our studies, which combine insights from the self-enhancement literature with the design structure of virtue discounting, show that self-serving biases emerge even under separate evaluation, where participants assess only one target in isolation, mirroring how moral judgments are often made in everyday life. Together, these findings clarify when and for whom virtue discounting is most pronounced while underscoring the role of self-serving biases in shaping perceptions of public moral displays.

Additionally, our results align with prior work that has made initial strides in uncovering the psychological architecture underlying virtue discounting (Kraft-Todd et al., 2023). Our findings, like earlier ones, suggest that tendencies to discount the moral value of public virtuous acts are largely explained by shifts in motivational attributions. We extend these findings to show that this pattern replicates regardless of who the actor or observer may be. This reinforces the idea that moral judgments align closely with Aristotelian virtue ethics (Aristotle, 1998; Cokelet & Fowers, 2019), where behaviors are not valued merely for their outcomes or impacts but for the motivation that underlies them and how this motivation allows insight into individuals' moral character. Rather than evaluating a generous action as virtuous solely by its prosocial impact, people may view true virtue as autotelic—where the act itself should be its own reward. When a seemingly virtuous behavior appears driven by external incentives, such as reputational gain, its sincerity is questioned, leading to moral discounting. Future research could build on this framework to better understand moral judgments in other domains. Much of moral

psychology has focused on how people apply rigid moral rules (e.g., deontological ethics) or assess tangible outcomes (e.g., consequentialist reasoning) in moral decision-making (Bostyn et al., 2018; Caviola et al., 2021; Greene, 2009; Swann et al., 2010). Yet, irrespective of these frameworks, the present findings suggest that moral judgments also hinge on what actions and their motivations signal about an individual's underlying character—a domain of social judgment perhaps better explained through an Aristotelian lens.

Beyond theoretical advancements, the present investigation also provides practical insights into the persistence of public moral behavior despite widespread skepticism. If individuals routinely anticipated their public virtue being judged as critically as they judge others', they might be discouraged from engaging in conspicuous moral acts. However, public acts of virtue are somewhat common in contemporary society (Miller, 2019). Our results *suggest* that people's expectations for observers to evaluate them in a more charitable light than they tend to afford to other actors *may* help to perpetuate public acts of good. Yet, while individuals may feel motivated to engage in visible acts of good, continued skepticism toward others' moral signaling could contribute to broader societal cynicism regarding public virtuosity. Indeed, mounting evidence suggests that public virtuosity (or "virtue signaling," in common parlance) is increasingly subject to moral vitriol (Westra, 2021), at times in a politically charged manner (Malic, 2020). This is worthy of attention as considerable research provides evidence for the collective benefits outward displays of virtue afford as tools for cultivating cultures of virtuous norms that perpetuate and inspire acts of good within cooperative societies (Cole et al., 2022; Gintis, 2003; Kraft-Todd et al., 2015; Levy, 2021). As such, further research should investigate directly whether tendencies toward and awareness of virtue discounting might constrain actual prosocial and societally beneficial behaviors, especially if such tensions may influence how people engage with, for instance, activism, charitable giving, and ethical business practices. In

the meantime, the current findings offer a glimmer of hope that self-enhancement may insulate people from such effects, at least when people consider perceptions of their own public virtue in isolation.

Finally, Study 4 also offers a practical insight. Joint evaluation may help people calibrate their moral judgments more consistently across targets. While the mechanism remains unclear (e.g., whether people become more critical of themselves or more generous toward others), this symmetry suggests that making comparative judgments explicit may interrupt self-serving asymmetries. Whether that alignment is normatively desirable depends on the broader aim. If the goal is to foster fairness and consistency in moral appraisal, joint evaluation may offer a promising tool. But if the goal is to protect sincere public virtue from undue skepticism, such formats could risk reinforcing discounting tendencies. Future work should investigate how judgment contexts (i.e., separate vs. joint) shape not only evaluative consistency but also moral tone, and ultimately, how those shifts influence behavior in domains where public virtue is often on display, such as activism, philanthropy, and ethical leadership.

Limitations

Despite the strengths of this research—including pre-registered studies, large samples, and built-in replications—several limitations open avenues for future work. First, our research focused on hypothetical behaviors and generosity. Examining real behaviors or virtues less prone to virtue discounting (e.g., fairness; Kraft-Todd et al., 2023) could provide stronger evidence for self-enhancement biases in public virtue judgments. Future research could also explore motivational attributions beyond those studied here, such as norm-signaling, which may be more susceptible to self-enhancement within the virtue discounting framework. Since different virtues are tied to distinct motivations (Narvaez & Snow, 2019), further investigation is needed to understand how these dynamics vary across a broader range of virtues and motivations. Further,

all research on virtue discounting to date has been conducted in WEIRD samples, leaving open
the question of cross-cultural variation. Given that Western societies emphasize individualism
(Hofstede, 2011), authenticity (Sherman et al., 2012), and self-enhancement more than Eastern
cultures (Gebauer et al., 2013), asymmetries between self and other distinctions, as well as public
and private contexts, may be less pronounced outside WEIRD populations.

Finally, while Studies 2-4 revealed consistent relationships between perceived
motivations and judgments of moral goodness, these data were correlational. As such, we cannot
draw firm conclusions about the causal direction of these associations. It is possible that
perceptions of virtue or reputation motives influenced judgments of moral goodness, as
suggested by virtue ethics frameworks (e.g., Kraft-Todd et al., 2023), but it is equally plausible
that participants' global impressions of moral worthiness biased their subsequent attribution of
motivations. Moral goodness was measured first in all studies, consistent with prior research and
our goal of minimizing the risk of contaminating self-serving asymmetries (e.g., prompting
participants to reflect on their own or others' underlying motives before making a moral
judgment could amplify differences in favor of the self). However, this measurement order
complicates the interpretation of our mediation models, which assume that motive attributions
temporally precede moral judgment. Without temporal separation or experimental manipulation
of these dimensions, such inferences remain speculative. Future research could more precisely
identify causal pathways by manipulating perceived motives and examining their downstream
effects on moral judgment (or vice versa). Such work would help clarify whether specific
motivations are taken as evidence of moral worth, or whether moral worth guides attributional
reasoning in public virtue contexts.

Conclusion

This research provides new insights into how people judge public acts of virtue depending on whether they are imagining themselves or someone else as the actor. By integrating the concept of virtue discounting with broader theories of self-enhancement, we find that public moral behavior is not judged consistently across targets. In separate evaluation contexts, participants generally anticipated that their own public virtue would be viewed more favorably (i.e., more principled, less reputation-driven, and more morally good) than comparable public acts performed by others. However, this self-serving asymmetry disappeared in a joint evaluation context, where participants assessed perceptions of self and other side by side. These findings clarify when people are more likely to apply moral skepticism and point to the role of self-enhancement in shaping expectations around public moral behavior.

References

Alicke, M. D., & Govorun, O. (2005). The Better-Than-Average Effect. In *The Self in Social Judgment*. Psychology Press.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, *68*(5), 804–825. https://doi.org/10.1037/0022-3514.68.5.804

Allison, P. D. (1992). The Cultural Evolution of Beneficent Norms. *Social Forces*, *71*(2), 279–301. https://doi.org/10.1093/sf/71.2.279

Appelbaum, S. H., Deguire, K. J., & Lay, M. (2005). The relationship of ethical climate to deviant workplace behaviour. *Corporate Governance: The International Journal of Business in Society*, *5*(4), 43–55. https://doi.org/10.1108/14720700510616587

Aquino, K., & Reed II, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, *83*(6), 1423–1440. https://doi.org/10.1037/0022-3514.83.6.1423

Aristotle. (1998). *Nicomachean ethics: Books VIII and IX* (M. Pakaluk, Ed. & Trans.). Clarendon Press.

Axelrod, R. (1986). An evolutionary approach to norms. *Amer. Polit. Sci. Rev.*, *80*(4), 1095–1111.

Berkowitz, L. (1972). Social norms, feelings, and other factors affecting helping and altruism. In *Advances in experimental social psychology* (Vol. 6, pp. 63–108). Elsevier.

Berntsen, D., & Rubin, D. C. (2024). Collectives closer to the self are anticipated to have a brighter future: Self-enhancement in collective cognition. *Journal of Experimental Psychology: General*, *153*(5), 1226–1235. https://doi.org/10.1037/xge0001550

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of Mice, Men, and Trolleys: Hypothetical

    Judgment Versus Real-Life Behavior in Trolley-Style Moral Dilemmas. *Psychological*

    Science, *29*(7), 1084–1093.

Brewer, M. B. (1991). The social self: On being the same and different at the same time.

    Personality and Social Psychology Bulletin, *17*(5), 475–482.

    https://doi.org/10.1177/0146167291175001

Brookes, J. (2015). The effect of overt and covert narcissism on self-esteem and self-efficacy

    beyond self-esteem. *Personality and Individual Differences*, *85*, 172–175.

    https://doi.org/10.1016/j.paid.2015.05.013

Brown, J. D. (1986). Evaluations of Self and Others: Self-Enhancement Biases in Social

    Judgments. *Social Cognition*, *4*(4), 353–376. https://doi.org/10.1521/soco.1986.4.4.353

Brown, J. D., Collins, R. L., & Schmidt, G. W. (1988). Self-esteem and direct versus indirect

    forms of self-enhancement. *Journal of Personality and Social Psychology*, *55*(3), 445–

    453. https://doi.org/10.1037/0022-3514.55.3.445

Caviola, L., Schubert, S., & Greene, J. D. (2021). The Psychology of (In)Effective Altruism.

    Trends in Cognitive Sciences. https://doi.org/10.1016/j.tics.2021.03.015

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and

    compliance. In S. T. F. D. T. Gilbert &. G. Lindzey (Ed.), *The handbook of social*

    psychology (pp. 151–192). McGraw-Hill.

Cokelet, B., & Fowers, B. J. (2019). Realistic virtues and how to study them: Introducing the

    STRIVE-4 model. *Journal of Moral Education*, *48*(1), 7–26.

    https://doi.org/10.1080/03057240.2018.1528971

Cole, J. C., Ehret, P. J., Sherman, D. K., & Van Boven, L. (2022). Social norms explain

   prioritization of climate policy. *Climatic Change*, *173*(1), 10.

   https://doi.org/10.1007/s10584-022-03396-x

Critcher, C. R., Dunning, D., & Rom, S. C. (2015). Causal trait theories: A new form of person

   knowledge that explains egocentric pattern projection. *Journal of Personality and Social

   Psychology*, *108*(3), Article 3. https://doi.org/10.1037/pspa0000019

Curry, O., Mullins, D., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of

   morality-as-cooperation in 60 societies. *Current Anthropology*, *60*(1).

   https://ora.ox.ac.uk/objects/uuid:8dd8d82d-3829-4857-bcf4-eebf196d11be

Day, M. V., Fiske, S. T., Downing, E. L., & Trail, T. E. (2014). Shifting liberal and conservative

   attitudes using moral foundations theory. *Personality & Social Psychology Bulletin*,

   40(12), 1559–1573. https://doi.org/10.1177/0146167214551152

Ding, Y., & Sugiura, M. (2024). The psychological mechanisms of the better-than-average effect

   in the moral and competence domains under self-enhancement and self-protection

   motives among young Japanese adults. *Frontiers in Psychology*, *15*.

   https://doi.org/10.3389/fpsyg.2024.1367568

Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated Moral Reasoning. In B. H.

   Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 50, pp. 307–338). Academic

   Press. https://doi.org/10.1016/S0079-7421(08)00410-6

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The

   role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of

   Personality and Social Psychology*, *57*(6), 1082–1090. https://doi.org/10.1037/0022-

   3514.57.6.1082

Ellemers, N., & van den Bos, K. (2012). Morality in Groups: On the Social-Regulatory

Functions of Right and Wrong. *Social and Personality Psychology Compass*, *6*(12), 878–

889. https://doi.org/10.1111/spc3.12001

Evans, A. M., & van de Calseyde, P. P. F. M. (2018). The reputational consequences of

generalized trust. *Personality and Social Psychology Bulletin*, *44*(4), 492–507.

https://doi.org/10.1177/0146167217742886

Gebauer, J. E., Wagner, J., Sedikides, C., & Neberich, W. (2013). Agency-Communion and Self-

Esteem Relations Are Moderated by Culture, Religiosity, Age, and Sex: Evidence for the

"Self-Centrality Breeds Self-Enhancement" Principle. *Journal of Personality*, *81*(3),

261–275. https://doi.org/10.1111/j.1467-6494.2012.00807.x

Gerber, A. S., & Rogers, T. (2009). Descriptive Social Norms and Motivation to Vote:

Everybody's Voting and so Should You. *The Journal of Politics*, *71*(01), 178–191.

https://doi.org/10.1017/S0022381608090117

Gintis, H. (2003). The hitchhiker's guide to altruism: Gene-culture coevolution, and the

internalization of norms. *Journal of Theoretical Biology*, *220*(4), 407–418.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person

perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148–

168. https://doi.org/10.1037/a0034726

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the

Moral Domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385.

https://doi.org/10.1037/a0021847

Greene, J. D. (2009). The cognitive neuroscience of moral judgment. *The Cognitive

Neurosciences IV*.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*(12), 517–523.

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture*, *2*(1). https://doi.org/10.9707/2307-0919.1014

Hsee, C. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, *67*(3), 247–257. https://doi.org/10.1006/obhd.1996.0077

Johnson, S. G., & Park, S. Y. (2021). Moral signaling through donations of money and time. *Organizational Behavior and Human Decision Processes*, *165*, 183–196.

Jordan, J. J., & Rand, D. G. (2019). Signaling when nobody is watching: A reputation heuristics account of outrage and punishment in one-shot anonymous interactions. *Journal of Personality and Social Psychology*, *118*(1), 57–88.

Jung, M. H., Nelson, L. D., Gneezy, U., & Gneezy, A. (2017). Signaling Virtue: Charitable Behavior Under Consumer Elective Pricing. *Marketing Science*, *36*(2), 187–194. https://doi.org/10.1287/mksc.2016.1018

Kim, M. Y., & Han, K. (2023). For me or for others? The better-than-average effect and negative feelings toward average others during the COVID-19 pandemic. *Current Psychology*, *42*(15), 13173–13181. https://doi.org/10.1007/s12144-021-02548-z

Klar, Y., & Giladi, E. E. (1997). No one in my group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology*, *73*(5), 885–901. https://doi.org/10.1037/0022-3514.73.5.885

Klar, Y., & Giladi, E. E. (1999). Are Most People Happier than their Peers, or Are They Just Happy? *Personality and Social Psychology Bulletin*, *25*(5), 586–595. https://doi.org/10.1177/0146167299025005004

Kodipady, A., Kraft-Todd, G., Sparkman, G., Hu, B., & Young, L. (2023). Beyond virtue

    signaling: Perceived motivations for pronoun sharing. *Journal of Applied Social*

    Psychology, *53*(7), 582–599. https://doi.org/10.1111/jasp.12937

Kogut, T., & Ritov, I. (2005). The singularity effect of identified victims in separate and joint

    evaluations. *Organizational Behavior and Human Decision Processes*, *97*(2), 106–116.

    https://doi.org/10.1016/j.obhdp.2005.02.003

Konuk, F. A., & Otterbring, T. (2024). The dark side of going green: Dark triad traits predict

    organic consumption through virtue signaling, status signaling, and praise from others.

    Journal of Retailing and Consumer Services, *76*, 103531.

    https://doi.org/10.1016/j.jretconser.2023.103531

Kraft-Todd, G. T., Kleiman-Weiner, M., & Young, L. (2023). Virtue Discounting: Observability

    Reduces Moral Actors' Perceived Virtue. *Open Mind*, *7*, 460–482.

    https://doi.org/10.1162/opmi_a_00085

Kraft-Todd, G., Yoeli, E., Bhanot, S., & Rand, D. (2015). Promoting cooperation in the field.

    Current Opinion in Behavioral Sciences, *3*, 96–101.

    https://doi.org/10.1016/j.cobeha.2015.02.006

Law, K. F., Amormino, P., Marsh, A. A., & O'Connor, B. B. (2023). Ethical reasoning versus

    empathic bias: A false dichotomy? *Trends in Cognitive Sciences*, *0*(0).

    https://doi.org/10.1016/j.tics.2023.10.006

Law, K. F., Gittle, E., Young, L., & Syropoulos, S. (2024). *Mapping and Increasing Americans'*

    Actual and Perceived Support for Initiatives Protecting Future Generations. OSF.

    https://osf.io/9my5g

Levy, N. (2021). Virtue signalling is virtuous. *Synthese*, *198*(10), 9545–9562.

    https://doi.org/10.1007/s11229-020-02653-9

Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime.com: A versatile crowdsourcing

 data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 1–10.

Malic, N. (2020). RT's The Wokies: America's politicians had a busy 2020 virtue-signaling.

 Here are the top 10. *RT International*. https://www.rt.com/op-ed/510675-wokies-america-

 politicians-virtue-signaling/

Marsh, A. A. (2019). The Caring Continuum: Evolved Hormonal and Proximal Mechanisms

 Explain Prosocial and Antisocial Extremes. *Annual Review of Psychology*, *70*(1), 347–

 371. https://doi.org/10.1146/annurev-psych-010418-103010

McManus, R. M., Young, L., & Sweetman, J. (2023). Psychology Is a Property of Persons, Not

 Averages or Distributions: Confronting the Group-to-Person Generalizability Problem in

 Experimental Psychology. *Advances in Methods and Practices in Psychological Science*,

 6(3), 25152459231186615. https://doi.org/10.1177/25152459231186615

Miller, G. F. (2019). *Virtue Signaling*. Cambrian Moon.

Mladenović, D., Jirásek, M., Ondráček, T., Opatrná, Z., & Štangová, R. (2023). The influence of

 social conformity on mask-wearing behavior during the COVID-19 pandemic. *Heliyon*,

 9(3), e14496. https://doi.org/10.1016/j.heliyon.2023.e14496

Narvaez, D., & Snow, N. E. (2019). Introduction to self, motivation and virtue studies. *Journal

 of Moral Education*, *48*(1), 1–6. https://doi.org/10.1080/03057240.2019.1556939

Nook, E. C., Ong, D. C., Morelli, S. A., Mitchell, J. P., & Zaki, J. (2016). Prosocial Conformity:

 Prosocial Norms Generalize Across Behavior and Empathy. *Personality and Social

 Psychology Bulletin*, *42*(8), 1045–1062. https://doi.org/10.1177/0146167216649932

Ostermaier, A., & Uhl, M. (2017). Spot on for liars! How public scrutiny influences ethical

 behavior. *PLOS ONE*, *12*(7), e0181682. https://doi.org/10.1371/journal.pone.0181682

Pizarro, D. A., & Bloom, P. (2003). The intelligence of the moral intuitions: A comment on

Haidt (2001). *Psychological Review*, *110*(1), 193–196. https://doi.org/10.1037/0033-

295x.110.1.193

Sedikides, C., & Gregg, A. P. (2008). Self-Enhancement: Food for Thought. *Perspectives on

Psychological Science*, *3*(2), 102–116. https://doi.org/10.1111/j.1745-6916.2008.00068.x

Sedikides, C., Meek, R., Alicke, M. D., & Taylor, S. (2014). Behind bars but above the bar:

Prisoners consider themselves more prosocial than non-prisoners. *British Journal of

Social Psychology*, *53*(2), 396–403. https://doi.org/10.1111/bjso.12060

Sherman, R. A., Nave, C. S., & Funder, D. C. (2012). Properties of persons and situations related

to overall and distinctive personality-behavior congruence. *Journal of Research in

Personality*, *46*(1), 87–101. https://doi.org/10.1016/j.jrp.2011.12.006

Sparkman, G., Geiger, N., & Weber, E. U. (2022). Americans experience a false social reality by

underestimating popular climate policy support by nearly half. *Nature Communications*,

13(1), Article 1. https://doi.org/10.1038/s41467-022-32412-y

Swann, W. B., Gómez, A., Dovidio, J. F., Hart, S., & Jetten, J. (2010). Dying and killing for

one's group: Identity fusion moderates responses to intergroup versions of the trolley

problem. *Psychological Science*, *21*(8), 1176–1183.

https://doi.org/10.1177/0956797610376656

Tappin, B. M., & McKay, R. T. (2017). The Illusion of Moral Superiority. *Social Psychological

and Personality Science*, *8*(6), 623–631. https://doi.org/10.1177/1948550616673878

Taylor, S. G., Eisenbarth, H., Sedikides, C., & Alicke, M. D. (2021). Explaining the better-than-

average effect among prisoners. *Journal of Applied Social Psychology*, *51*(2), 90–106.

https://doi.org/10.1111/jasp.12719

Westra, E. (2021). Virtue Signaling and Moral Progress. *Philosophy & Public Affairs*, *49*(2), 156–178. https://doi.org/10.1111/papa.12187

Wojciszke, B., & Białobrzeska, O. (2014). Agency versus Communion as Predictors of Self-esteem: Searching for the Role of Culture and Self-construal. *Polish Psychological Bulletin*. https://doi.org/10.2478/ppb-2014-0057

Zaki, J. (2020). *The War for Kindness: Building Empathy in a Fractured World*. Crown.

Zell, E., Strickhouser, J. E., Sedikides, C., & Alicke, M. D. (2020). The better-than-average effect in comparative self-evaluation: A comprehensive review and meta-analysis. *Psychological Bulletin*, *146*(2), 118–149. https://doi.org/10.1037/bul0000218

People Are More Skeptical of Others' Public Virtue Motivations than Their Own in Separate

(but not Joint) Evaluations

Supplementary Online Material

Table of Contents

Supplementary Study S1

Study S1 served as a proof of concept/pilot study, replicating and extending existing research demonstrating that a self-serving asymmetry effect manifests in moral contexts. Specifically, we assessed the extent to which individuals, on average, believe they are seen as more moral than others—broadly and in a characterological sense—and whether these meta-perceptions are linked to how they judge the moral character of others. While prior research has established self-other asymmetries in specific moral traits (e.g., Tappin & McKay, 2017), less is known about whether these asymmetries extend to broader moral character, particularly in the context of meta-perceptions. Since people often use behaviors to infer moral character (Hartman et al., 2022), understanding how individuals believe they are morally perceived served as a first step prior to testing whether self-enhancement biases extend beyond traits to public moral actions.

Method

Participants

A total of 163 participants recruited via Prolific were included in our analyses.

Materials and Procedure

Participants completed a battery of measures in a randomized order. Of relevance to the current investigation were two items, each focusing on perceptions of moral character. Both were captured on scales ranging from -100 (morally bad) to 100 (morally good). The items were: (a) "Please use the scale provided to rate how moral you think other people see you as..." and (b) "Please use the scale provided to rate how moral a normal person is."

Results

Participants believed that others saw them as significantly more moral ($M = 71.48$, $SD = 23.17$) than the average person ($M = 49.07$, $SD = 34.98$), $t(162) = 8.87$, $p < .001$, $d = 0.69$.

Moreover, meta-perceptions of one's own moral character were positively associated with perceptions of others' moral character ($b = 0.67$, 95% CI [0.46, 0.88], $\beta = 0.44$, $p < .001$, $R^2 =$ .20), suggesting that individuals who believed they were viewed as moral also tended to see others as more moral. Sensitivity analyses indicated that with 80% power, we could detect effect sizes as small as $R^2 = .06$ and $d = 0.22$ for two-tailed tests.

Discussion

We replicate the better than average effect in the moral domain finding that participants believe others see them as more moral than the average person.

Study 1 – Supplementary Information

Table S1.

Correlations for ratings of moral goodness across the 17 vignettes.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Lending money to a friend	--															
Paying someone a compliment	0.50	--														
Holding the door for someone	0.38	0.53	--													
Helping someone fix a flat tire	0.32	0.37	0.50	--												
Helping an elderly person cross the street	0.29	0.36	0.52	0.62	--											
Giving a waiter a large tip	0.42	0.47	0.51	0.45	0.40	--										
Buying someone groceries	0.39	0.31	0.45	0.52	0.67	0.37	--									
Buying a homeless person food	0.41	0.34	0.47	0.58	0.63	0.40	0.72	--								
Donating your car to charity	0.49	0.40	0.39	0.40	0.42	0.45	0.46	0.48	--							
Helping someone mow their lawn	0.50	0.51	0.53	0.56	0.57	0.50	0.55	0.57	0.48	--						
Walk someone's dog	0.47	0.48	0.54	0.49	0.42	0.47	0.42	0.50	0.36	0.61	--					
Donating to artists or content creators	0.50	0.36	0.37	0.43	0.38	0.52	0.37	0.36	0.52	0.48	0.41	--				
Picking up trash in a park	0.41	0.42	0.52	0.59	0.64	0.46	0.58	0.67	0.49	0.59	0.48	0.39	--			
Donating money to charity	0.43	0.35	0.37	0.42	0.49	0.40	0.42	0.55	0.70	0.43	0.38	0.54	0.50	--		
Buying a friend an expensive gift	0.54	0.58	0.38	0.25	0.32	0.52	0.33	0.35	0.49	0.50	0.46	0.44	0.39	0.41	--	
Letting someone ahead of you in line	0.37	0.45	0.48	0.34	0.40	0.42	0.39	0.38	0.34	0.45	0.48	0.36	0.43	0.36	0.41	--
Staying late to help a coworker	0.41	0.44	0.45	0.56	0.56	0.44	0.56	0.55	0.38	0.59	0.47	0.34	0.58	0.37	0.36	0.44

Note. All *p*s < .001.

Re-Analysis of Study 1 as a Mixed Model

A linear mixed-effects model was conducted to re-examine the effects of agent (Self vs. Other) context (Public vs. Private), and their interaction, on moral ratings across 17 vignettes. Random intercepts were included for both participants and vignettes to account for variability across individuals and stimuli. The dependent variable was ratings of moral goodness for each vignette. Denominator degrees of freedom were estimated using the Kenward–Roger method.

There was significant variability in baseline ratings across participants ($\sigma^2 = 0.575$) and across vignettes ($\sigma^2 = 0.131$). The residual variance was 0.722. The main effect of agent was not significant, $F(1, 453) = 0.07$, $p = .789$, indicating no difference between the Self and Other conditions. The main effect of context was significant, $F(1, 453) = 79.61$, $p < .001$, with higher moral ratings in the public compared to the private condition. The interaction between term was not significant, $F(1, 453) = 1.46$, $p = .227$. Thus, moral ratings differed depending on whether the behavior was observed publicly or privately, but there was no evidence that having the participant or somebody else be the agent influenced ratings, nor was there an interaction between the two factors. Participants and vignettes varied in their baseline levels, justifying the inclusion of crossed random intercepts. The estimated marginal means indicated that participants in the public condition ($M = 4.10$, $SE = 0.08$, 95% CI [3.95, 4.25]) rated vignettes significantly higher than participants in the private condition ($M = 3.20$, $SE = 0.08$, 95% CI [3.05, 3.35]). This corresponds to a Cohen's $d \approx 0.79$, indicating a large effect size.

Study 2 – Supplementary Information

Overall, when we examine bivariate correlations collapsed across conditions, we find that

perceiving an action as genuinely motivated by virtue (principled motivation) is related to

increased trust in the agent, and increased perceptions of moral goodness. Perceiving an action as

an attempt to benefit one's reputation had an opposite pattern of associations. Trusting the agent

positively and strongly related to perceived moral goodness of the agent. Further, believing that

an action is normative (i.e., that it should be conducted) related to greater perceptions of moral

goodness, trust and principled motivation.

Table S2.

Bivariate correlation across all vignettes, conditions and participants.

	1	2	3	4	5
1. Moral Goodness	--				
2. Trust	0.70*	--			
3. Principled Motivation	0.68*	0.65*	--		
4. Reputation Signaling	-0.32*	-0.27*	-0.36*	--	
5. Normative Expectation	0.28*	0.27*	0.36*	0.08*	--

Note. $*p < .001$.

Table S3.

Descriptive statistics for moral goodness by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	6.04 (0.92)	5.96 (0.89)	5.05 (1.10)	4.46 (1.24)
Pay colleague a compliment	5.33 (1.19)	5.48 (0.88)	5.17 (1.01)	5.22 (1.19)
Hold door for someone	5.76 (1.13)	5.61 (0.92)	5.68 (0.94)	5.20 (1.16)
Fix someone's flat tire	6.17 (0.93)	6.29 (1.05)	6.36 (0.80)	6.00 (0.87)
Walk elderly person across street	6.17 (0.91)	6.26 (0.82)	5.97 (0.88)	5.80 (1.15)
Tip waiter generously	5.73 (1.06)	5.82 (0.95)	5.35 (1.12)	5.10 (0.99)
Buy groceries for someone in need	5.92 (1.12)	6.14 (0.85)	6.03 (0.98)	5.81 (1.14)
Buy food for homeless person	6.14 (0.86)	6.41 (0.82)	6.11 (0.88)	5.58 (1.38)
Donate car to charity	6.19 (0.81)	6.18 (0.97)	5.23 (1.16)	5.11 (1.09)
Mow friend's lawn	5.94 (0.99)	5.94 (1.03)	5.89 (0.92)	5.49 (1.26)
Walk friend's dog	5.58 (1.01)	5.70 (1.00)	5.31 (1.06)	5.14 (1.20)
Donate to an artist	5.25 (1.31)	5.75 (0.92)	4.98 (1.18)	4.84 (1.10)
Pick up trash	6.07 (0.96)	6.13 (0.85)	5.78 (0.95)	5.69 (1.34)
Donate to charity	5.95 (1.16)	6.24 (0.90)	5.51 (1.05)	5.08 (1.26)
Buy friend an expensive gift	5.54 (1.28)	5.35 (1.07)	4.65 (1.37)	4.37 (1.07)
Let someone in a rush skip the line	5.68 (1.05)	5.72 (1.00)	5.19 (1.23)	5.34 (1.13)
Help coworker after hours	5.89 (0.99)	6.06 (0.93)	5.73 (0.85)	5.90 (0.84)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S4.

Descriptive statistics for trust by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	5.73 (1.31)	5.63 (0.99)	4.78 (1.08)	4.24 (1.29)
Pay colleague a compliment	5.25 (1.22)	5.02 (1.00)	4.94 (1.04)	4.88 (1.02)
Hold door for someone	5.50 (1.10)	5.24 (0.93)	5.24 (1.09)	4.64 (1.17)
Fix someone's flat tire	5.90 (1.07)	5.65 (1.19)	5.91 (0.98)	5.38 (1.03)
Walk elderly person across street	5.90 (1.10)	5.94 (1.01)	5.59 (0.96)	5.14 (1.27)
Tip waiter generously	5.22 (1.17)	5.37 (1.11)	4.70 (1.18)	4.52 (1.16)
Buy groceries for someone in need	5.58 (1.21)	5.55 (1.16)	5.56 (1.02)	5.16 (1.12)
Buy food for homeless person	5.60 (1.21)	5.88 (1.03)	5.61 (0.97)	4.90 (1.42)
Donate car to charity	5.60 (1.26)	5.60 (1.18)	4.72 (1.16)	4.53 (1.08)
Mow friend's lawn	5.67 (1.08)	5.51 (1.08)	5.33 (1.01)	5.05 (1.08)
Walk friend's dog	5.46 (0.99)	5.42 (1.16)	5.11 (1.10)	4.84 (1.24)
Donate to an artist	4.99 (1.40)	5.19 (1.12)	4.79 (1.25)	4.48 (1.17)
Pick up trash	5.37 (1.42)	5.40 (1.11)	5.06 (1.26)	4.96 (1.33)
Donate to charity	5.41 (1.19)	5.64 (1.09)	5.04 (1.08)	4.51 (1.08)
Buy friend an expensive gift	5.49 (1.22)	5.10 (1.14)	4.53 (1.37)	4.12 (1.19)
Let someone in a rush skip the line	5.29 (1.26)	5.16 (1.16)	4.91 (1.19)	5.09 (1.16)
Help coworker after hours	5.80 (1.15)	5.63 (1.13)	5.67 (1.00)	5.52 (1.00)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S5.

Descriptive statistics for principled motivation by vignette split by observability and target

agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	5.66 (0.92)	5.81 (0.79)	5.14 (1.07)	4.52 (1.12)
Pay colleague a compliment	5.14 (1.05)	5.23 (0.77)	4.82 (1.03)	4.95 (1.01)
Hold door for someone	5.65 (0.88)	5.64 (0.82)	5.43 (0.89)	5.08 (1.03)
Fix someone's flat tire	5.98 (0.86)	6.02 (0.99)	6.06 (0.86)	5.63 (1.06)
Walk elderly person across street	6.03 (0.72)	6.13 (0.78)	5.57 (1.02)	5.61 (1.01)
Tip waiter generously	5.47 (0.91)	5.59 (1.02)	4.96 (1.11)	4.79 (1.03)
Buy groceries for someone in need	5.92 (0.78)	5.71 (0.97)	5.67 (0.89)	5.52 (1.12)
Buy food for homeless person	5.75 (0.81)	6.18 (0.77)	5.59 (0.86)	5.28 (1.29)
Donate car to charity	5.74 (1.08)	5.88 (0.84)	4.93 (1.02)	4.77 (1.17)
Mow friend's lawn	5.77 (0.87)	5.92 (0.81)	5.56 (0.79)	5.33 (0.91)
Walk friend's dog	5.21 (1.04)	5.69 (0.77)	5.24 (0.99)	4.99 (0.90)
Donate to an artist	5.50 (1.04)	5.69 (0.75)	4.93 (1.01)	4.81 (1.01)
Pick up trash	5.59 (1.08)	5.72 (0.80)	5.30 (1.09)	5.23 (1.25)
Donate to charity	5.72 (1.20)	6.03 (0.83)	5.25 (0.96)	4.74 (0.96)
Buy friend an expensive gift	5.24 (1.13)	4.96 (0.82)	4.63 (1.22)	4.22 (0.70)
Let someone in a rush skip the line	5.64 (0.90)	5.56 (0.90)	5.10 (1.08)	5.27 (0.97)
Help coworker after hours	5.65 (0.83)	5.81 (0.79)	5.36 (0.93)	5.58 (0.86)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S6.

Descriptive statistics for reputation signaling by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	3.14 (1.53)	2.72 (1.12)	4.66 (1.27)	4.99 (1.24)
Pay colleague a compliment	3.26 (1.24)	3.18 (1.11)	4.09 (1.17)	4.27 (1.40)
Hold door for someone	2.87 (1.28)	2.62 (0.99)	3.40 (1.34)	3.72 (1.23)
Fix someone's flat tire	3.19 (1.26)	2.80 (1.24)	3.43 (1.36)	3.40 (1.22)
Walk elderly person across street	3.09 (1.29)	2.54 (1.08)	3.55 (1.28)	3.86 (1.49)
Tip waiter generously	3.24 (1.06)	3.30 (1.31)	4.69 (1.27)	4.98 (1.13)
Buy groceries for someone in need	3.02 (1.19)	2.79 (0.92)	3.95 (1.40)	4.18 (1.22)
Buy food for homeless person	3.06 (1.22)	2.68 (1.08)	3.99 (1.38)	4.29 (1.50)
Donate car to charity	3.16 (1.41)	2.90 (1.30)	4.78 (1.35)	5.39 (0.91)
Mow friend's lawn	2.98 (1.21)	2.82 (1.01)	3.64 (1.42)	4.01 (1.20)
Walk friend's dog	3.09 (1.44)	2.83 (1.25)	3.35 (1.43)	4.21 (1.33)
Donate to an artist	3.19 (1.42)	2.77 (1.14)	4.80 (1.25)	5.11 (1.08)
Pick up trash	3.28 (1.43)	2.81 (1.01)	3.75 (1.41)	4.21 (1.47)
Donate to charity	3.12 (1.35)	2.66 (1.02)	4.90 (1.16)	5.36 (1.00)
Buy friend an expensive gift	3.14 (1.38)	3.55 (1.24)	4.85 (1.34)	5.05 (1.01)
Let someone in a rush skip the line	3.13 (1.31)	3.05 (1.22)	3.59 (1.20)	3.71 (1.31)
Help coworker after hours	3.42 (1.35)	3.42 (1.17)	3.86 (1.26)	4.17 (1.18)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S7.

Descriptive statistics for normative expectation by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	4.59 (1.41)	4.54 (1.16)	4.52 (1.07)	4.57 (1.06)
Pay colleague a compliment	4.98 (1.20)	4.81 (1.02)	4.85 (1.04)	5.13 (1.15)
Hold door for someone	5.59 (1.13)	5.44 (1.03)	5.43 (1.06)	5.34 (1.09)
Fix someone's flat tire	5.23 (1.12)	5.23 (1.09)	5.14 (1.05)	5.19 (1.16)
Walk elderly person across street	5.48 (1.11)	5.14 (0.96)	5.06 (1.02)	5.18 (0.91)
Tip waiter generously	4.87 (1.24)	4.41 (1.43)	4.64 (1.06)	4.63 (1.19)
Buy groceries for someone in need	4.52 (1.14)	4.45 (1.19)	4.25 (1.31)	4.37 (1.21)
Buy food for homeless person	4.89 (1.26)	4.81 (1.23)	4.71 (1.00)	5.03 (0.96)
Donate car to charity	4.73 (1.40)	4.10 (1.34)	4.34 (1.12)	4.70 (1.23)
Mow friend's lawn	4.89 (1.31)	4.98 (1.10)	4.99 (1.08)	4.89 (1.09)
Walk friend's dog	4.75 (1.12)	4.76 (1.36)	4.89 (1.28)	4.90 (1.03)
Donate to an artist	4.54 (1.17)	4.29 (1.20)	4.64 (1.13)	4.87 (1.07)
Pick up trash	5.01 (1.35)	4.53 (1.06)	4.70 (1.12)	4.97 (1.22)
Donate to charity	4.82 (1.25)	4.52 (1.26)	4.71 (0.92)	4.93 (1.00)
Buy friend an expensive gift	4.69 (1.36)	4.35 (1.13)	4.35 (1.25)	4.40 (0.91)
Let someone in a rush skip the line	4.98 (1.30)	5.00 (1.12)	4.51 (1.16)	4.82 (1.11)
Help coworker after hours	4.95 (1.23)	4.88 (0.98)	4.77 (1.13)	5.18 (0.88)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Study 3 – Supplementary Information

The observed pattern of correlations replicated the findings of Study 2 (see Table S8).

Table S8.

Bivariate correlation across all vignettes, conditions and participants.

	1	2	3	4	5
1. Moral Goodness	--				
2. Trust	0.65*	--			
3. Principled Motivation	0.67*	0.60*	--		
4. Reputation Signaling	-0.26*	-0.25*	-0.34*	--	
5. Normative Expectation	0.28*	0.21*	0.35*	0.07*	--

Note. $*p < .001$.

Table S9.

Pre-registered multilevel regression

	Moral Goodness	Trust	Principled Motivation	Reputation Signaling	Normative Expectation
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
Intercept	5.76*** (.08)	5.38*** (.08)	5.65*** (.07)	2.82*** (.10)	4.96*** (.09)
Public = 1 (Private = 0)	-.28** (.11)	-.39 ** (.12)	-.43*** (.11)	1.44*** (.14)	.07 (.13)
Self-Agent = 1 (Other- Agent = 0)	-.32** (.11)	-.32* (.12)	-.21 (.11)	-.01 (.14)	-.34** (.13)
Self-Rater = 1 (Other-Rater = 0)	.22* (.11)	-.001 (.12)	.002 (.11)	.07 (.14)	-.20 (.12)
Observability*Rater	-.44** (.15)	-.27 (.17)	-.21 (.15)	-.12 (.19)	.16 (.18)
Rater*Agent	.19 (.15)	.90*** (.17)	-.13 (.15)	-.08 (.19)	.26 (.18)
Observability*Agent	.23 (.15)	.39* (.17)	.23 (.15)	-.44* (.19)	-.04 (.18)
Observability* Rater*Agent	.00 (.22)	.13 (.24)	.11 (.21)	-.17 (.27)	-.49* (.25)

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

Table S10

Descriptive statistics for moral goodness by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	5.57 (1.16)	5.72 (0.90)	4.92 (1.26)	4.81 (1.19)
Pay colleague a compliment	5.24 (1.11)	5.11 (1.24)	5.27 (1.08)	5.03 (1.17)
Hold door for someone	5.51 (1.18)	5.52 (1.03)	5.77 (0.89)	5.61 (0.99)
Fix someone's flat tire	6.00 (1.01)	6.14 (0.94)	5.90 (0.98)	6.06 (0.80)
Walk elderly person across street	6.12 (1.10)	6.39 (0.91)	5.85 (0.97)	5.91 (1.07)
Tip waiter generously	5.61 (1.09)	5.84 (1.25)	5.24 (1.07)	5.29 (0.96)
Buy groceries for someone in need	6.07 (1.08)	6.31 (0.81)	5.83 (1.24)	5.98 (0.74)
Buy food for homeless person	6.04 (1.05)	6.32 (0.94)	5.55 (1.17)	5.49 (1.15)
Donate car to charity	5.78 (1.18)	6.10 (0.88)	4.90 (1.24)	5.22 (1.04)
Mow friend's lawn	5.86 (1.04)	6.00 (1.03)	5.52 (1.02)	5.67 (0.93)
Walk friend's dog	5.25 (1.27)	5.53 (1.00)	5.19 (1.13)	4.85 (1.20)
Donate to an artist	5.07 (1.21)	5.46 (0.95)	4.83 (1.14)	4.90 (1.11)
Pick up trash	6.06 (1.16)	6.32 (0.70)	5.86 (1.03)	5.71 (0.96)
Donate to charity	5.62 (1.32)	6.11 (0.83)	5.28 (1.07)	5.28 (0.99)
Buy friend an expensive gift	5.27 (1.33)	5.52 (0.99)	4.26 (1.02)	4.34 (0.99)
Let someone in a rush skip the line	5.44 (1.18)	5.47 (1.06)	4.96 (1.41)	4.97 (1.42)
Help coworker after hours	5.76 (0.95)	5.88 (0.98)	5.90 (1.04)	5.67 (0.82)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S11

Descriptive statistics for trust by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	5.71 (1.07)	5.34 (1.18)	5.17 (1.29)	4.40 (1.27)
Pay colleague a compliment	5.35 (0.98)	4.84 (1.21)	5.53 (1.01)	4.50 (1.10)
Hold door for someone	5.47 (1.13)	5.02 (1.14)	5.73 (1.09)	4.93 (1.08)
Fix someone's flat tire	5.53 (1.22)	5.69 (1.18)	5.85 (1.08)	5.35 (0.98)
Walk elderly person across street	5.92 (0.93)	5.87 (1.17)	5.96 (0.94)	5.50 (1.05)
Tip waiter generously	5.56 (1.12)	5.33 (1.30)	5.40 (1.21)	4.58 (1.13)
Buy groceries for someone in need	5.87 (1.06)	5.64 (1.13)	5.69 (1.20)	5.31 (1.00)
Buy food for homeless person	5.51 (1.29)	5.70 (1.17)	5.53 (1.03)	4.78 (1.18)
Donate car to charity	5.38 (1.11)	5.55 (0.97)	4.93 (1.27)	4.68 (0.99)
Mow friend's lawn	5.61 (1.13)	5.51 (1.12)	5.65 (0.96)	5.13 (1.12)
Walk friend's dog	5.49 (1.24)	5.14 (1.11)	5.53 (1.20)	4.94 (1.11)
Donate to an artist	5.28 (1.09)	4.95 (1.07)	4.75 (1.24)	4.44 (1.06)
Pick up trash	5.57 (1.33)	5.46 (0.99)	5.69 (1.06)	4.87 (1.11)
Donate to charity	5.44 (1.32)	5.38 (1.05)	5.23 (1.10)	4.66 (0.97)
Buy friend an expensive gift	5.29 (1.33)	5.45 (1.22)	4.91 (1.16)	4.11 (1.12)
Let someone in a rush skip the line	5.37 (1.32)	5.05 (0.96)	5.01 (1.30)	4.65 (1.35)
Help coworker after hours	5.69 (1.02)	5.55 (1.05)	5.88 (1.09)	5.53 (0.90)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S12

Descriptive statistics for principled motivation by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	5.46 (1.12)	5.55 (0.86)	5.03 (1.21)	4.75 (1.19)
Pay colleague a compliment	5.08 (0.98)	4.86 (0.96)	5.12 (1.02)	4.79 (1.06)
Hold door for someone	5.42 (1.00)	5.37 (1.07)	5.46 (1.12)	5.34 (1.07)
Fix someone's flat tire	5.58 (1.06)	6.09 (0.81)	5.80 (0.80)	5.70 (0.85)
Walk elderly person across street	6.03 (0.78)	6.02 (0.93)	5.70 (1.05)	5.58 (0.88)
Tip waiter generously	5.39 (0.86)	5.53 (1.03)	5.00 (1.22)	4.70 (1.13)
Buy groceries for someone in need	5.79 (0.95)	6.06 (0.78)	5.55 (1.00)	5.53 (0.85)
Buy food for homeless person	5.72 (0.97)	6.16 (0.68)	5.36 (1.05)	5.13 (1.11)
Donate car to charity	5.60 (1.05)	5.71 (0.80)	4.70 (1.38)	4.93 (1.02)
Mow friend's lawn	5.72 (0.96)	5.76 (0.91)	5.63 (0.82)	5.37 (0.99)
Walk friend's dog	5.26 (1.13)	5.36 (0.90)	5.27 (1.12)	4.99 (1.03)
Donate to an artist	5.36 (0.95)	5.43 (0.92)	4.83 (0.97)	4.85 (1.04)
Pick up trash	5.77 (0.93)	5.89 (0.69)	5.53 (1.00)	5.20 (1.09)
Donate to charity	5.65 (1.04)	5.98 (0.76)	5.05 (1.12)	5.15 (0.88)
Buy friend an expensive gift	4.97 (1.04)	5.26 (0.92)	4.35 (1.05)	4.21 (0.97)
Let someone in a rush skip the line	5.22 (1.23)	5.33 (0.94)	5.01 (1.23)	5.09 (1.09)
Help coworker after hours	5.63 (0.91)	5.69 (0.83)	5.68 (0.93)	5.40 (0.78)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S13

Descriptive statistics for reputation signaling by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	2.72 (1.17)	2.82 (1.00)	4.05 (1.68)	4.68 (1.49)
Pay colleague a compliment	2.92 (1.35)	3.48 (1.07)	3.45 (1.38)	4.09 (1.42)
Hold door for someone	3.12 (1.26)	2.82 (1.14)	2.85 (1.37)	3.29 (1.11)
Fix someone's flat tire	2.69 (1.05)	2.60 (0.95)	3.23 (1.31)	3.63 (1.26)
Walk elderly person across street	2.66 (0.99)	2.63 (0.96)	3.04 (1.31)	3.87 (1.31)
Tip waiter generously	2.88 (1.16)	3.20 (1.01)	4.22 (1.63)	4.77 (1.18)
Buy groceries for someone in need	3.00 (1.41)	2.65 (0.91)	3.43 (1.55)	4.24 (1.13)
Buy food for homeless person	2.75 (1.07)	2.68 (0.87)	3.61 (1.35)	4.26 (1.19)
Donate car to charity	2.94 (1.20)	2.77 (1.10)	4.77 (1.60)	5.18 (0.99)
Mow friend's lawn	2.67 (1.07)	2.89 (0.94)	3.03 (1.21)	3.83 (1.26)
Walk friend's dog	2.77 (1.18)	2.88 (0.92)	3.09 (1.46)	3.90 (1.29)
Donate to an artist	2.55 (1.00)	2.63 (0.88)	4.34 (1.49)	4.93 (1.03)
Pick up trash	2.57 (1.01)	2.70 (0.98)	3.41 (1.42)	3.90 (1.33)
Donate to charity	2.70 (1.11)	2.50 (1.06)	4.47 (1.68)	5.10 (1.12)
Buy friend an expensive gift	2.90 (1.37)	3.16 (1.21)	4.31 (1.55)	5.15 (0.96)
Let someone in a rush skip the line	2.77 (1.12)	2.93 (1.06)	3.18 (1.46)	3.43 (1.30)
Help coworker after hours	2.98 (1.19)	3.12 (1.16)	3.68 (1.43)	4.19 (1.02)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S14

Descriptive statistics for normative expectation by vignette split by observability and target agent.

Vignette	Private		Public	
	Self	Other	Self	Other
Lend money to a friend	4.40 (1.34)	4.60 (1.21)	4.16 (1.32)	4.55 (1.21)
Pay colleague a compliment	4.84 (1.13)	4.89 (0.88)	4.73 (1.46)	5.17 (0.86)
Hold door for someone	5.38 (1.25)	5.53 (1.00)	5.27 (1.31)	5.70 (0.87)
Fix someone's flat tire	4.88 (1.27)	5.30 (1.04)	4.89 (1.22)	5.32 (1.05)
Walk elderly person across street	5.19 (1.29)	5.15 (1.07)	4.76 (1.49)	5.29 (1.05)
Tip waiter generously	4.29 (1.32)	5.02 (1.22)	4.30 (1.58)	4.70 (1.18)
Buy groceries for someone in need	4.43 (1.45)	4.87 (0.99)	4.23 (1.34)	4.71 (1.08)
Buy food for homeless person	4.58 (1.38)	4.95 (1.02)	4.25 (1.39)	4.92 (0.99)
Donate car to charity	4.13 (1.38)	4.38 (1.14)	4.08 (1.37)	4.72 (1.00)
Mow friend's lawn	4.81 (1.44)	4.65 (1.28)	4.71 (1.41)	5.01 (0.95)
Walk friend's dog	4.75 (1.33)	4.72 (1.05)	4.44 (1.39)	5.17 (1.04)
Donate to an artist	4.17 (1.18)	4.45 (1.10)	4.31 (1.16)	4.96 (1.01)
Pick up trash	4.87 (1.35)	5.09 (1.17)	4.77 (1.37)	5.15 (1.04)
Donate to charity	4.57 (1.45)	4.80 (1.27)	4.48 (1.32)	5.19 (1.03)
Buy friend an expensive gift	4.42 (1.54)	4.45 (1.25)	3.99 (1.37)	4.53 (0.89)
Let someone in a rush skip the line	4.55 (1.18)	4.90 (1.08)	4.34 (1.47)	4.84 (1.10)
Help coworker after hours	4.71 (1.25)	4.93 (1.06)	4.71 (1.19)	4.99 (0.86)

Note. Values indicate means and standard deviations in the format: *M* (*SD*).

Table S15

*Comparisons of direct and indirect association for the Path Models of Study 2 and Study 3.*

| Comparison	Study 2				Study 3			
	b	*SE*	*Lower 95% Cr.I.*	*Upper 95% Cr.I.*	*b*	*SE*	*Lower 95% Cr.I.*	*Upper 95% Cr.I.*
Trust vs. RS (correlation with outcome)	0.44	0.03	0.39	0.49	0.48	0.02	0.43	0.53
Trust vs. NE (correlation with outcome)	0.38	0.02	0.33	0.42	0.40	0.03	0.36	0.45
PM vs. RS (correlation with outcome)	0.40	0.03	0.35	0.45	0.48	0.02	0.43	0.52
PM vs. NE (correlation with outcome)	0.33	0.03	0.28	0.38	0.40	0.03	0.35	0.45
PM vs. RS (indirect effect for self on outcome)	-0.11	0.03	-0.18	-0.05	-0.09	0.04	-0.16	-0.01
PM vs. NE (indirect effect for self on outcome)	-0.15	0.02	-0.20	-0.11	-0.15	0.03	-0.21	-0.09
Trust vs. RS (indirect effect for other on outcome)	-0.24	0.04	-0.32	-0.17	-0.20	0.04	-0.28	-0.12
Trust vs. NE (indirect effect for other on outcome)	-0.28	0.03	-0.35	-0.22	-0.27	0.04	-0.33	-0.20
PM vs. RS (indirect effect for other on outcome)	-0.18	0.03	-0.25	-0.12	-0.14	0.04	-0.22	-0.07
PM vs. NE (indirect effect for other on outcome)	-0.22	0.03	-0.28	-0.17	-0.21	0.03	-0.27	-0.15

Note. RS = Reputation Signaling, NE = Normative Expectation, PM = Principled Motivation, Cr.I. = Credible Interval.

Supplementary Figures – Study 2

Figures S2A-S2E.

Ratings of moral goodness (lower scores denote more virtue discounting), normative expectation, reputation si

signaling, and trust (for the agent) for the 17 actions, when the action was conducted by oneself or other people,
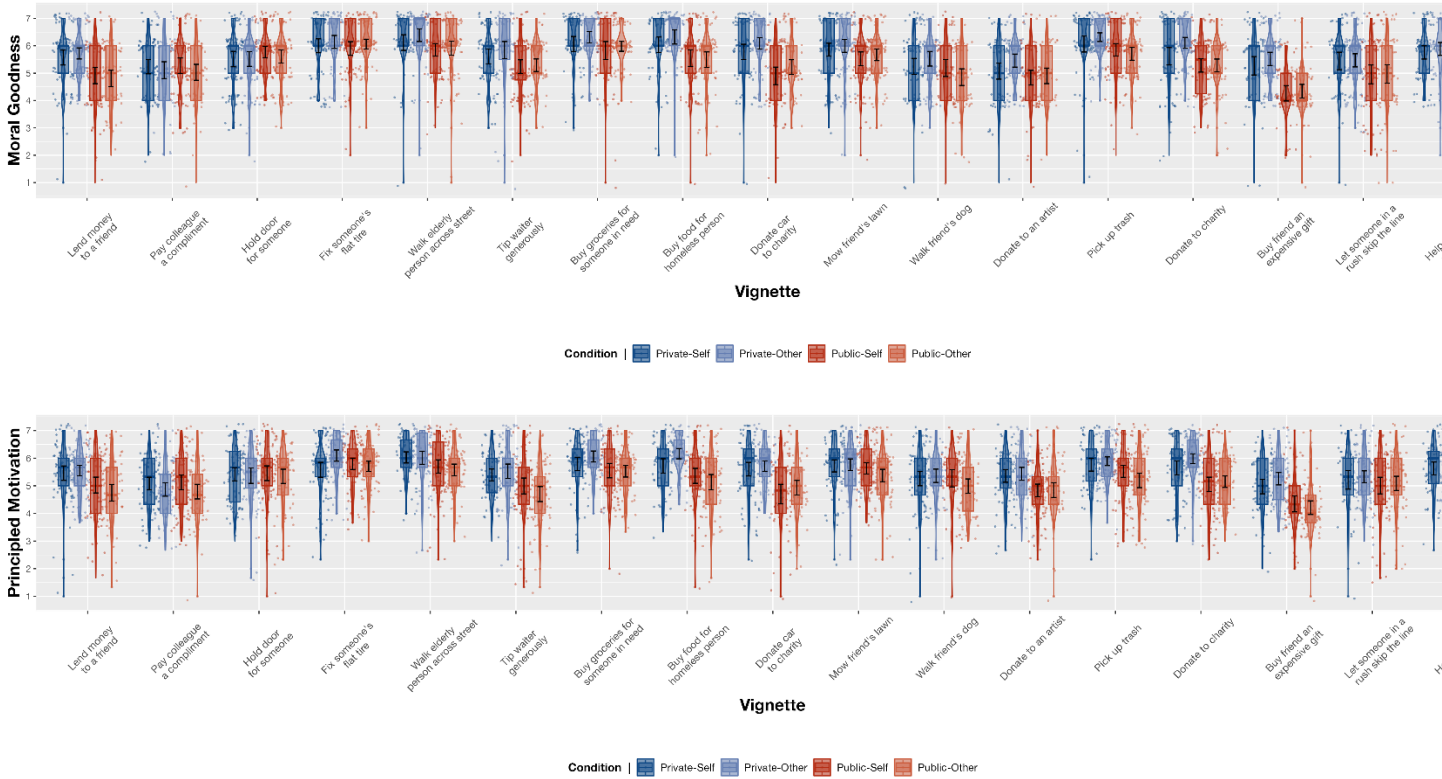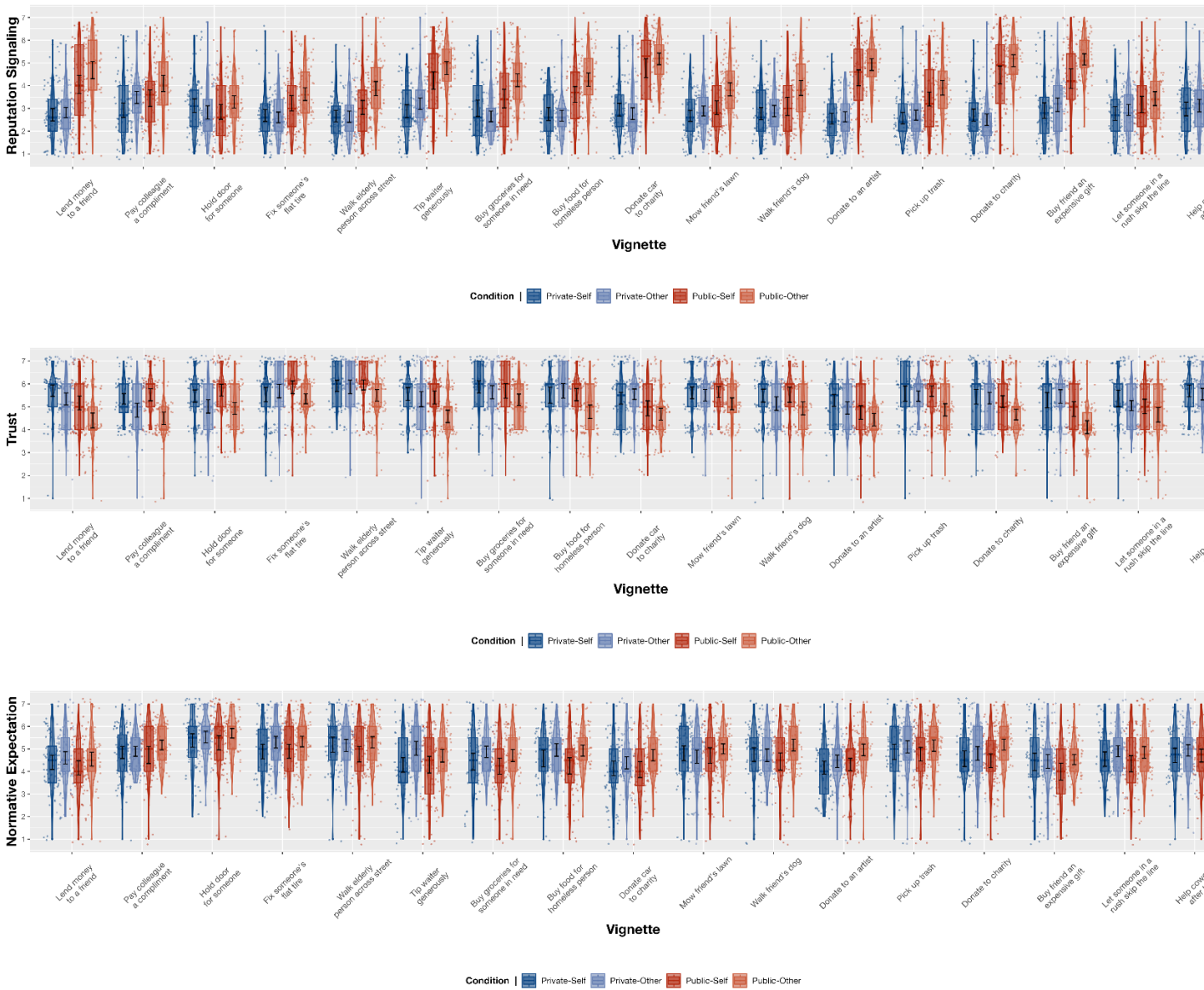
observability for Study 2.
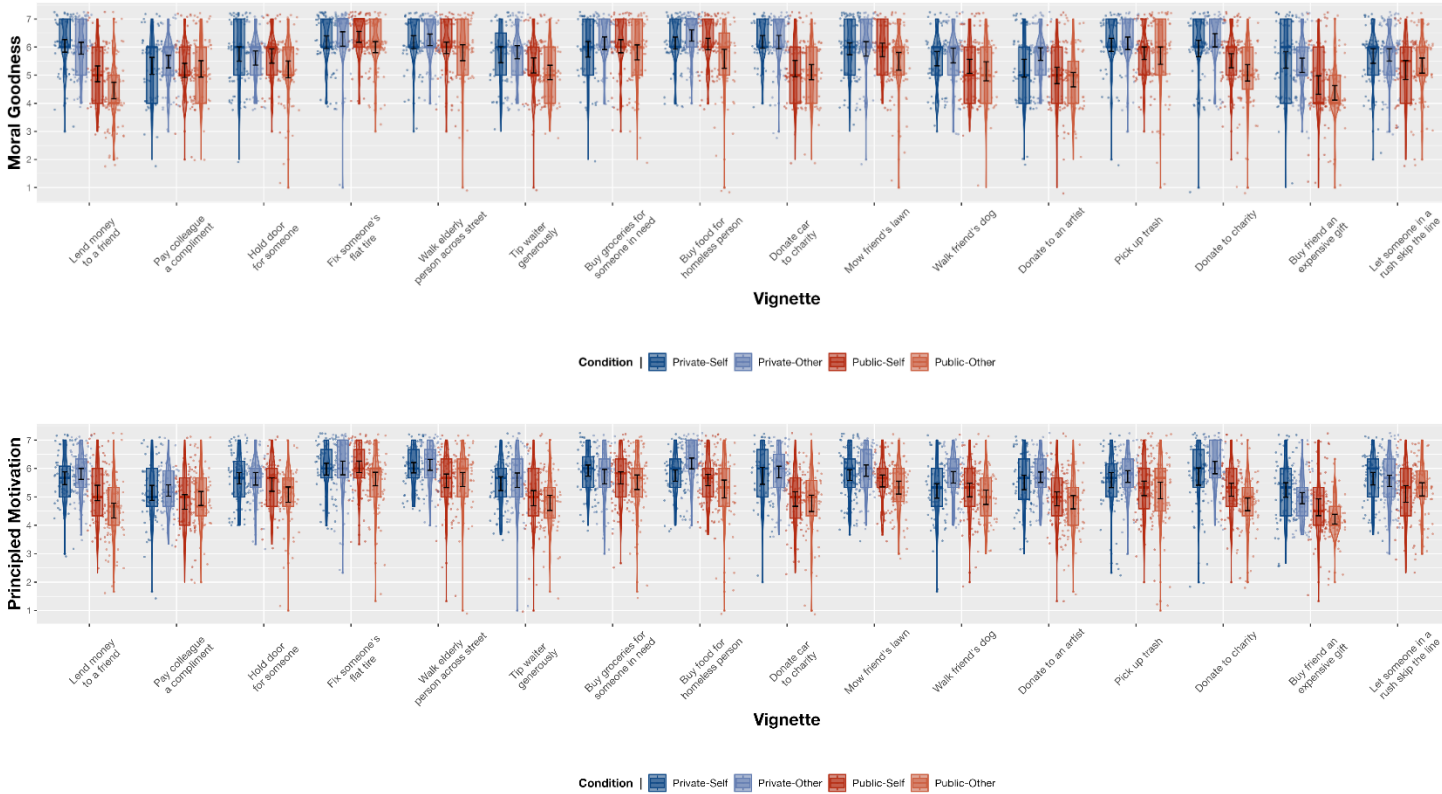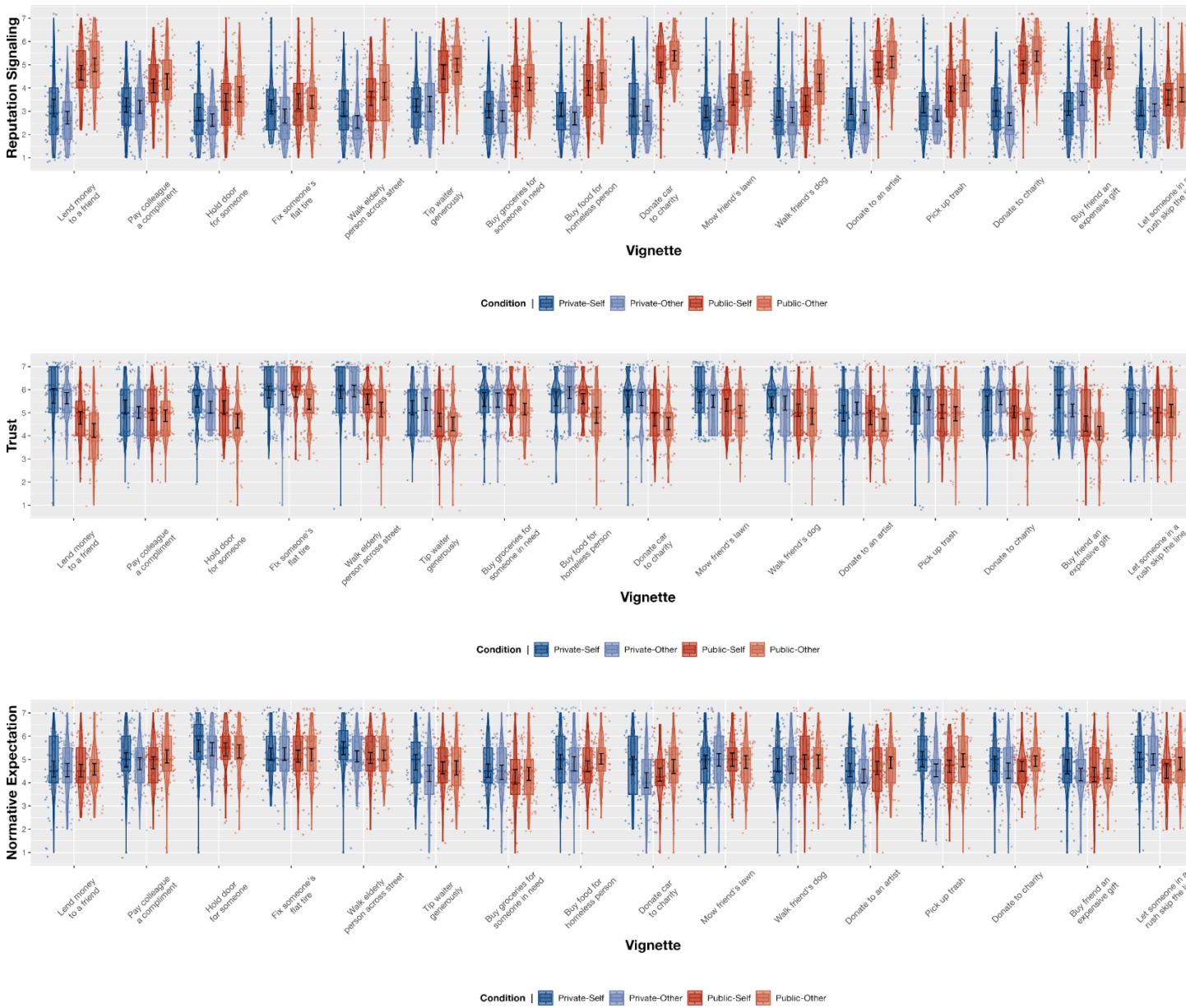
Supplementary Figures – Study 3

Figures S3A-S3E.

Ratings of moral goodness (lower scores denote more virtue discounting), normative expectation, reputation sig

signaling, and trust (for the agent) for the 17 actions, when the action was conducted by oneself or other people,

observability for Study 3.

Supplementary References

Hartman, R., Blakey, W., & Gray, K. (2022). Deconstructing moral character judgments.

Current Opinion in Psychology, *43*, 205–212.

https://doi.org/10.1016/j.copsyc.2021.07.008

Tappin, B. M., & McKay, R. T. (2017). The Illusion of Moral Superiority. *Social Psychological*

and Personality Science, *8*(6), 623–631. https://doi.org/10.1177/1948550616673878