Abstract

Self-control refers to the ability to choose options with greater long-term benefits over more immediately tempting options. For personal choices that do not affect others, self-control is often conceptualized as morally irrelevant. However, four focal experiments and five supplemental experiments demonstrate that self-control success in apparently non-moral domains enhances evaluations of moral character, but self-control failure is not regarded as evidence of moral corruption. This asymmetry supports our *moral-ability* hypothesis: self-control is regarded as the ability to bring about intended outcomes, which is believed necessary for moral goodness but not moral badness.

Keywords: self-control, moral character, social inference, impression formation

Willpower as Moral Ability

Self-control refers to the ability to choose options that are beneficial in the long term over those that are immediately tempting but costlier in the long term (Inzlicht et al., 2014). Self-control dilemmas may have social as well as personal consequences (e.g., procrastination on group projects versus individual tasks). While self-control is deemed morally relevant when it affects group welfare (Mooijman et al., 2018), open questions remain about when *personal* self-control dilemmas attain moral relevance.

In fact, such personal dilemmas are often described as demanding "non-moral" self-control because they involve no direct social consequences (Berman & Small 2018; Hofmann et al., 2018). Nonetheless, these personal choices (e.g., between healthy versus unhealthy foods) are routinely described in terms of "virtues versus vices" or "shoulds versus wants" (Milkman et al., 2008; Wertenbroch, 1998). Do people indeed regard willpower in explicitly personal, non-moral choices as diagnostic of moral character? If so, when, and why? We investigate these questions by examining naïve theories about the relationship between non-moral self-control and moral character.

Past research suggests two competing hypotheses. One is the *moral-neutrality hypothesis*, which holds that willpower is a morally neutral ability and naïve perceptions of self-control and moral character are largely independent (Blasi, 2005). In line with this view, self-control traits (e.g., being hardworking) are often perceived as peripheral to prototypes of moral character (Lapsley & Lasky, 2001; Smith et al., 2007), and are both distinct from and dominated by moral traits in social impressions (Cottrell et al., 2007; Goodwin et al., 2014; Skowronski & Carlston, 1987). A competing perspective is the *master-virtue hypothesis*, whereby willpower is essential to moral character and enables the development of other virtues and adherence to social norms (Baumeister & Exline, 2000; McCullough & Carter, 2011). For instance, Righetti and Finkenauer (2011) suggest that people invest greater trust in

partners high in personal self-control because they see it as necessary for generating good social outcomes. Therefore, while the moral-neutrality hypothesis predicts no clear relation between perceptions of self-control and moral character, the master-virtue hypothesis predicts that self-control success is viewed as evidence of moral goodness and failure as evidence of moral badness.

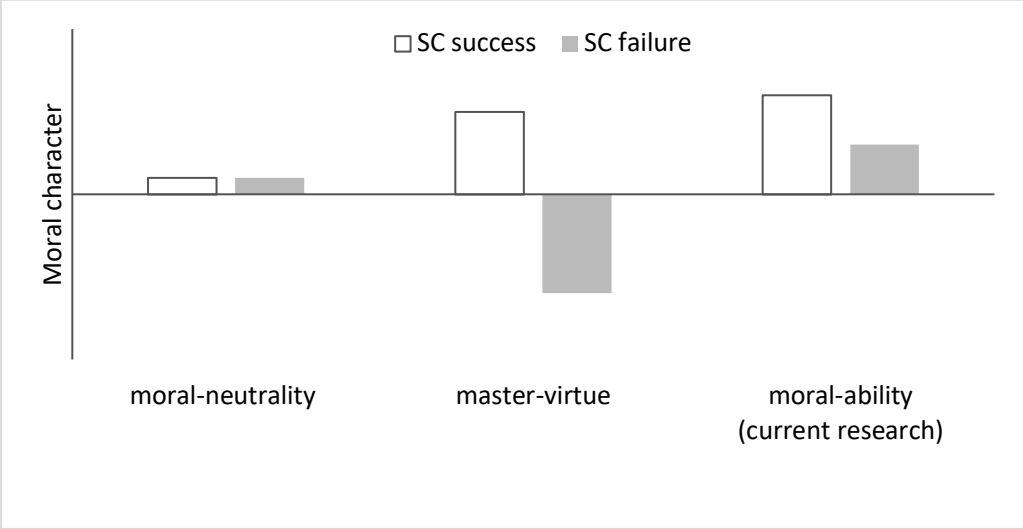
This research advances a third possibility that we call the *moral-ability hypothesis,* whereby judgments of moral character are jointly determined by perceived intentions *and* the perceived ability to enact them. We predict that non-moral self-control primarily informs the latter. Our account also predicts an inferential asymmetry, whereby self-control is perceived as particularly diagnostic of moral goodness, based on two possibilities. First, absent any information about an individual's intentions, people often assume good intentions (Cacioppo et al., 1997), and might thus view self-control as the ability to specifically realize *good* outcomes and be a good person. Second, judgments of bad moral character are primarily determined by evil desires, even without any ability to realize harm (Inbar et al., 2012). We expect that enhancing good moral character is instead thought to require *both* good intentions *and* the ability to generate good outcomes (cf. Stellar & Willer, 2018). Supporting this premise, when asked to evaluate prototypical moral agents, people readily saw good moral character as evidence of strong self-control, but viewed bad moral character as much poorer evidence of weak self-control (see Pilot in SOM).

We thereby predict that while self-control success is thought to demonstrate the ability to realize good moral intentions, and thus enhance good moral character, self-control failure is not thought to imply bad intentions or signal badness. Figure 1 summarizes these predictions against existing perspectives. Four pre-registered experiments and five supplemental experiments in the SOM test the moral-ability hypothesis and distinguish it from alternative explanations. Table 1 provides an overview of the studies and the observed

asymmetry in moral character inferences from each one. Unless specified otherwise, we recruited US-based participants from Amazon Mechanical Turk. All data and materials are publicly available at osf.io/f9bm5 (Gai 2021).

Figure 1

Competing predictions about moral character inferences from personal self-control.



Note. Values above the x-axis represent moral goodness and values below the x-axis represent moral badness.

Table 1

Overview of studies and observed moral character inferences by self-control outcome.

Exp	Participants	Design	$d_{success}$	$d_{failure}$
1	N = 200 43 females, M_{age} = 34, SD_{age} = 9	Self-control success or failure, between subjects	1.81	0.94
2	N = 247 in China from Credamo, 130 females, M_{age} = 28, SD_{age} = 5	Choice (wants or shoulds) × self-control conflict (present or absent), between subjects	2.49	0.38
3	N = 201 in US or UK from Prolific, 123 females, M_{age} = 41, SD_{age} = 14	Self-control success or failure, between subjects	1.96	0.64
4	N = 246 172 males, M_{age} = 36, SD_{age} = 10	High vs. low self-control, within subjects	1.46	0.30

S1a	N = 300 108 females, M_{age} = 37, SD_{age} = 10	Self-control success or failure, between subjects	2.08	1.27
S1b	N =202 61 females, M_{age} = 36, SD_{age} = 10	Self-control success or failure, between subjects	1.56	1.19
S2	N = 623 from various sources 260 females, M_{age} = 36, SD_{age} = 11	Self-control success, failure, or no-conflict, between subjects	1.73	0.82
S3	N = 432, 173 females, M_{age} = 36, SD_{age} = 11	Self-control success or failure, between subjects	1.70	0.69
S4	N = 204 85 females M_{age} = 34, SD_{age} = 10	Self-control (success or failure) × measure (moral or non-moral traits), between subjects	1.57	0.23

Note. The effects in the two rightmost columns are Cohen's *d*s of moral character ratings against the scale midpoint (which represent "can't tell"). In Experiment 3, the difference between moral and immoral trait selections was compared to zero (i.e., equal selection). Positive *d* values indicate perceived moral goodness and negative values moral badness. Only effect sizes from the self-control success and self-control failure conditions are shown, though some experiments included other conditions.

Experiment 1

Experiment 1 (preregistered at bit.ly/3hBH6xh) served as an initial test of the predicted asymmetry in moral character inferences from personal self-control. To further assess the moral-ability hypothesis, we also examined how self-control affects perceived moral intentions and moral abilities, and how each contributes to moral character judgments.

Method

We randomly assigned participants to evaluate either personal self-control successes or failures. They each read four different within-subjects vignettes about two females and two males facing different self-control dilemmas (health, procrastination, budgeting, patience), with order randomized. Each vignette described a target individual who experienced self-control conflict and chose the "should" versus the "want" option, as shown in the example below (the *self-control success* condition of the Procrastination vignette, adapted from Berman & Small, 2018):

> **Mike** is a college student. He is in his dorm room and is thinking about writing an important paper for class. He notices that there are a bunch of football games on TV. He could slack off and watch TV or turn off the TV and focus on writing his paper.
>
> Mike feels **very conflicted** about this decision. He feels a strong desire to watch football, and he is very tempted to do so now. However, even though he is tempted to watch football, he decides to turn off the television and focus on his paper.

Failures had no direct social consequences, though some successes involved small social costs (e.g., rejecting a friend's invitation to save money; see SOM for all stimuli).

Following each vignette, participants evaluated the target individual's moral character on two items (adapted from Inbar et al., 2012; Pearson's *r*s between .43-.59): "Do you think that ___ is mainly a good person or a bad person?" (1 = mainly a bad person, 4 = can't tell, 7 = mainly a good person), and "Do you think that ___ has good or bad moral standards?" (1 = definitely bad moral standards, 4 = can't tell, 7 = definitely good moral standards).

Finally, participants completed two items concerning each target's moral intentions (1 = definitely bad, 4 = can't tell, 7 = definitely good) and ability to achieve good or bad outcomes (1 = definitely weak, 4 = can't tell, 7 = definitely strong), with order randomized.

Results

For ease of interpretation, we recoded the scale to range from -3 to +3, with zero representing "can't tell." As shown in Figure 2, self-control success increased moral character evaluations relative to failure across vignettes (M = 1.51 95% CI = [1.32, 1.70] vs. 1.02 [0.83, 1.21], $F(1, 198) = 13.05$, $p < .001$, $\eta_p^2 = .06$). Supporting the moral-ability hypothesis, ratings of moral character fell significantly above the scale midpoint in *both* conditions (ts > 6.97, ps < .001, ds > 0.69), suggesting that self-control failure was not thought to indicate moral badness.

Further in line with the moral-ability hypothesis, self-control success increased perceived moral abilities relative to failure (M = 1.65 [1.40, 1.90] vs. 0.87 [0.62, 1.11], F(1,

198) = 19.57, $p < .001$, $\eta_p^2 = 0.09$). Self-control success also increased perceived good intentions ($M = 1.63$ [1.43, 1.84] vs. 1.19 [0.98, 1.39], $F(1, 198) = 9.47$, $p = .002$, $\eta_p^2 = .05$), but the effect was significantly weaker (interaction: $F(1, 198) = 9.10$, $p = .003$, $\eta_p^2 = .04$). A linear regression (with random intercepts nested within individuals) showed that perceived moral abilities contributed independently to evaluations of moral character ($b = 0.25$, $SE = 0.03$, $t(598) = 9.81$, $p < .001$) alongside moral intentions ($b = 0.31$, $SE = 0.03$, $t(598) = 10.94$, $p < .001$).

These findings were replicated using between-subjects and categorical measurement approaches (see Experiments S1a-S1b in SOM), suggesting they are robust and cannot be explained by measurement artifacts or scaling choices.

Figure 2

Results of Experiment 1.

Note. Bar plots represent the means and the error bars the 95% confidence intervals.

Experiment 2

Experiment 2 (pre-registered at bit.ly/3dZqJJb) compared moral character inferences from self-control (i.e., moral ability) versus mere *preferences* for "should" or "want" options.

Thus, beyond choice outcomes, we also manipulated whether choices involved self-control conflict and required willpower.

To explore this distinction, we also measured individual beliefs in free will (BFW), the belief that moral actions are caused by intentions rather than predetermined dispositions (Everett et al., 2021).[1] Accordingly, BFW should increase the impact of choices involving willpower (i.e., the ability to realize intentions) on perceived moral character, but reduce the impact of choices based on dispositional preferences.

Method

We used a 2 (choice: "should" vs. "want") × 2 (self-control conflict: present vs. absent) between-subjects design and recruited native Chinese participants from Credamo.com. Participants read the four vignettes from Experiment 1 (translated into Mandarin) about target individuals choosing "shoulds" or "wants." This time, we also manipulated whether self-control conflict was present or absent (e.g., as in the *preference for wants* condition of the Health vignette: "M feels **no conflict** at all about this decision. She doesn't like working out and she really likes KFC. She decides to join her friend to have KFC"). Participants then rated each individual's moral character on the same items used in Experiment 1 (*r*s between .60-.70). Afterwards, participants completed the free-will subscale of the FAD-plus (Cronbach's α = .89; Paulhus & Carey, 2011).

Results

Replicating Experiment 1, self-control success (vs. failure) increased ratings of moral character ($F(1, 241) = 67.08$, $p < .001$, $\eta_p^2 = .36$). Mere preferences for shoulds (vs. wants) also increased moral character ratings, though to a slightly lesser extent ($F(1, 241) = 51.41$, $p < .001$, $\eta_p^2 = .18$). See Experiment S2 for additional comparisons of these effect sizes.

---

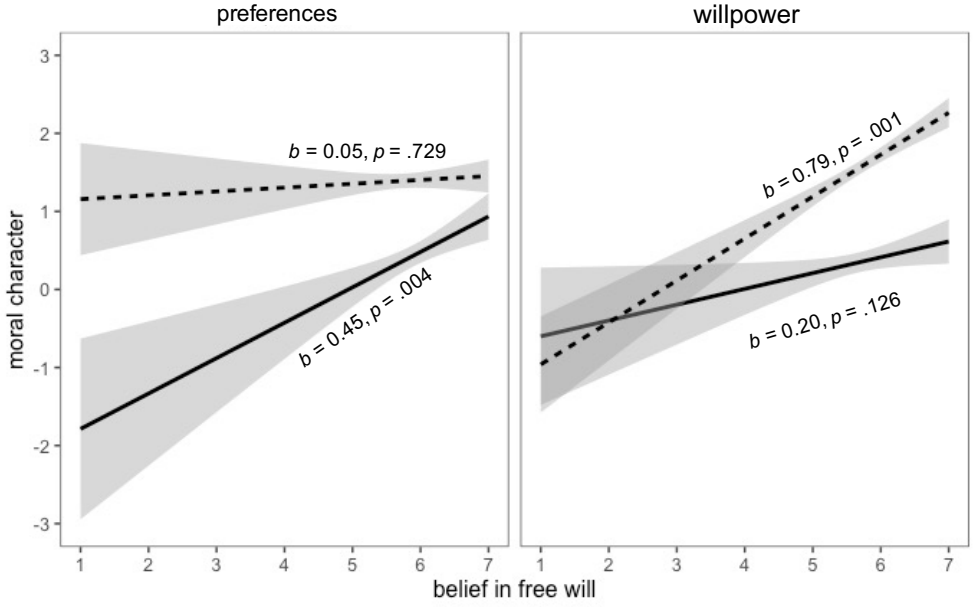
[1] Because the BFW measurement and moderation analyses were not pre-registered, we treat them as exploratory.

A linear regression with random intercepts revealed a significant three-way interaction among choice, self-control conflict, and BFW ($t(237) = 2.56$, $p = .011$; see SOM for full results), with starkly different patterns as a function of self-control conflict (see Figure 3). Specifically, the asymmetry in moral character inferences from choices involving self-control conflict strengthened as belief in free will increased (Johnson-Neyman cutoff at 4.22; see Experiment S3 for replications). In contrast, the impact of choices based on preference strengthened as people believed *less* in free will.

These findings suggest that self-control shapes inferences of moral character by altering one's perceived ability to act on their intentions, not via moralization of their preferences (Mooijman et al. 2020). They also imply that people who believe strongly in free will tend to view moral character as contingent on the ability to overcome conflict, while free will skeptics see unconflicted expressions of preference as clearer moral signals.

Figure 3

Fitted results of Experiment 2.



Wants —	M = 0.43 ± 0.19	M = 0.34 ± 0.19
Should - -	M = 1.39 ± 0.19	M = 1.57 ± 0.19
$d_{success\text{-}failure}$	1.30	1.67

Note. Shaded bands and ± are ranges of 95% confidence intervals.

Experiment 3

Experiment 3 (pre-registered at bit.ly/2T6Ikb2) had three goals. First, we examined moral inferences from personal self-control in an incentive-compatible game with externally valid stimuli. Another goal was to rule out halo effects (i.e., overgeneralization of positive traits) as an alternative explanation (Nisbett & Wilson, 1977). Because people often view competent individuals as less warm (Fiske et al., 2007), we predicted that self-control success (versus failure) would reduce warmth perceptions, contrary to its effect on moral character perceptions. Finally, because individual choices might seem unrepresentative of "true" abilities, we manipulated *patterns* of behavioral self-control outcomes to assess the robustness of our prior findings.

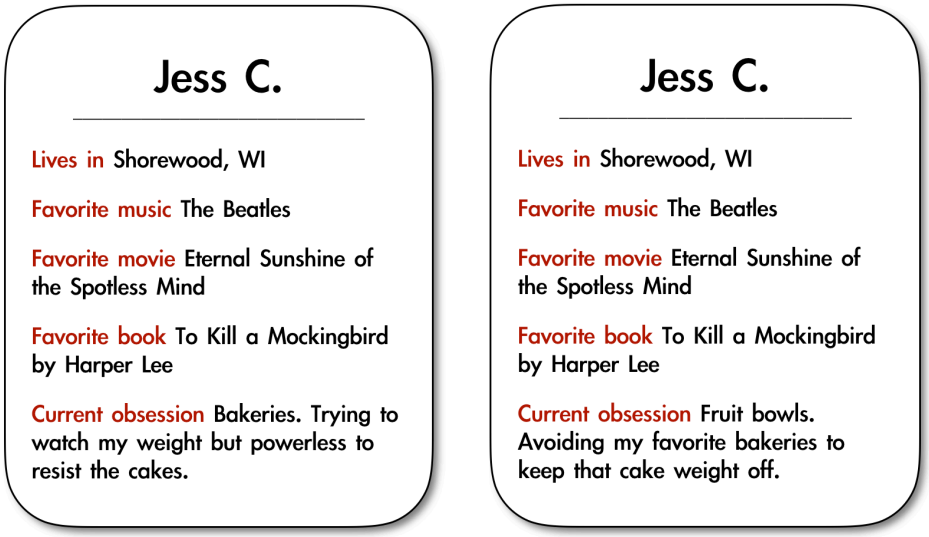
Method

Participants were invited to play "Guess Their Personality," a game modeled on real-world social media apps. They were randomly assigned to view the profiles of four fictitious individuals (two females and two males) whose "current obsessions" described patterns of either self-control success or failure (see Figure 4).

For each individual, participants then guessed the top 3 traits voted by others in that individual's network. Participants chose from sixteen traits (adapted from Goodwin et al., 2014): four moral (courageous, principled, just, trustworthy), four immoral (faithless, unfair, irresponsible, and dishonest), four warm (easy-going, warm, sociable, happy) and four cold traits (serious-minded, disagreeable, calm, and humorless). To incentivize accuracy, we promised a 5-cent bonus for each correct guess. All participants received a $0.10 bonus.

Figure 4

Sample profiles from self-control failure (left) and success (right) condition in Experiment 3.



Results

Non-parametric tests of trait selection frequencies by condition are reported in the SOM. To facilitate comparison across studies, we report ANOVA results below, which are entirely consistent.

Relative to self-control failure, self-control success increased moral trait selections (M = 0.99 [0.92, 1.06] vs. 0.50 [0.42, 0.58], $F(1,199) = 89.73$, $p < .001$, $\eta_p^2 = .31$) and decreased immoral trait selections (M = 0.06 [0.02, 0.10] vs. 0.20 [0.16, 0.24], $F(1,199) = 19.70$, $p < .001$, $\eta_p^2 = .09$). Consistent with the moral-ability hypothesis, moral traits were selected far more often than immoral traits across conditions ($F(1,199) = 342.74$, $p < .001$, $\eta_p^2 = .63$), and self-control outcome had a more pronounced effect on the selection of moral versus immoral traits (interaction: $F(1,199) = 43.78$, $p < .001$, $\eta_p^2 = .18$).

Importantly, self-control success *decreased* warm trait selections relative to failure (M = 1.20 [1.19, 1.21] vs. 1.60 [1.50, 1.61], $F(1,199) = 32.90$, $p < .001$, $\eta_p^2 = .14$), though it did not significantly affect cold trait frequencies (M = 0.76 [0.69, 0.83] vs. 0.71 [0.64, 0.78], $F(1,199) = 0.90$, $p = .343$, $\eta_p^2 = .01$). These results support the theoretical distinction between

morality and warmth, and suggest that our findings cannot be explained by halo effects (see Experiment S4 for a replication), unrepresentative individual choices, or specific stimuli.

Experiment 4

Finally, Experiment 4 (pre-registered at bit.ly/2Qw7ZcC) tested whether moral character inferences from self-control affect consequential behaviors. Participants played an incentive-compatible trust game adapted from Everett, Pizarro, and Crockett (2016) with partners of varying personal self-control ability.

Method

Participants received $0.60 in base pay and were promised a bonus between $0.00-$1.20. All had to pass a comprehension check about the trust game payoffs to participate.

The study was divided into two parts. First, participants completed a survey (ostensibly drawn from a pool of possibilities) adapted from the Brief Self-Control Scale (Tangney et al., 2004). It included eight items about personal (i.e., "non-moral") self-control (e.g., "I am bad at resisting temptations"; $\alpha = .88$), and two about moral self-control outcomes that affect others ("I say inappropriate things," and "I lose my temper too easily."), on a five-point scale (1 = not at all, 5 = very much; $\alpha = .90$ for all ten).

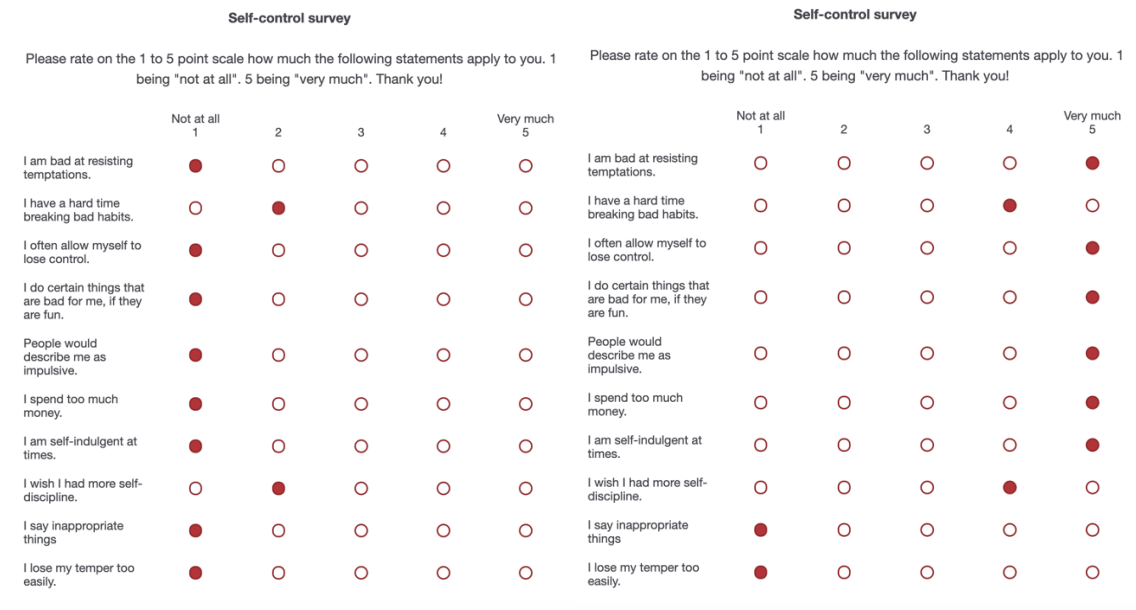
Next, participants learned that we had paired them with two other workers who completed the same survey, person A and person B, and would play the trust game sequentially with each. One ostensibly reported strong personal self-control ability by answering 1 or 2 on every item, while the other reported weak self-control by answering 4 or 5 on each, with the person A or B label randomly assigned (see Figure 5). Notably, to ensure that these individuals varied only in *personal* self-control ability, their ratings on the two moral self-control items was held constant at 1 across conditions.

Participants read each person's survey responses in turn and completed the following measurements from Everett et al. (2016): 1) the person's moral character ("What do you

think about this person?" -3 = extremely immoral/bad, +3 = extremely moral/good; -3 = extremely untrustworthy, +3 = extremely trustworthy; participants were told to skip the question if they could not tell), 2) a dollar amount to pass to the person, and 3) the dollar amount they expected the person to return. After completing these questions for both persons A and B, participants indicated who they would prefer to play with if they could choose, then reported their demographics. All participants received a $0.40 bonus.

Figure 5

Strong (left panel) versus weak (right panel) self-control stimuli from Experiment 4.



Results

As in our prior studies, the partner with strong self-control was thought to have better moral character than the partner with weak self-control (M = 1.58 [1.45, 1.71] vs. 0.48 [0.28, 0.68], $t(245) = 9.32$, $p < .001$, $d = 0.59$). No participants skipped these items due to uncertainty.

More importantly, participants passed more money to the partner with strong self-control than the partner with weak self-control (M = $16.76 [15.66, 17.86] vs. $13.89 [12.71, 14.77], $t(245) = 5.31$, $p < .001$, $d = 0.34$). They also expected the partner with strong (vs. weak) self-

control to return a higher percentage of that money (M = 39% [36%, 42%] vs. 33% [30%, 36%], $t(190) = 3.64$, $p < .001$, $d = 0.26$, excluding those who passed zero). All results were robust to nonparametric tests (see SOM).

Lastly, participants preferred to play the trust game with the individual with strong self-control (169/246, 69%, $Z = 32.69$, $p < .001$), regardless of their own reported personal self-control (8-item: $Z = 0.39$, $p = .531$) or overall self-control (10-item: $Z = 0.54$, $p = .462$).

Exploratory within-subjects mediation analyses with 10,000 bootstrapped samples (Montoya & Hayes, 2017) showed a significant indirect effect via moral character perceptions on monetary allocations (95% CI = [1.17, 2.75]), which fully accounted for the direct effect of partner self-control ($b = 0.96$, $SE = 0.59$, $t(243) = 1.63$, $p = .104$, 95% CI = [-0.20, 2.11]). Moral character perceptions also mediated the effect of partner self-control on participants' return percentage expectations (95% CI = [0.70, 2.75]; direct effect [0.04, 2.73]), as well as on their partner preferences (with a logit link, 95% CI = [0.98, 1.06]; direct effect [0.04, 0.35]), supporting the moral-ability hypothesis.

General Discussion

Across various contexts and measures, we find a consistent inferential asymmetry: personal self-control success is thought to signal moral goodness, but self-control failure is not seen as evidence of moral corruption. This asymmetry contradicts both the moral-neutrality hypothesis, which predicts no systematic relation between self-control and moral character perceptions, and the master-virtue hypothesis, which predicts that self-control failures should signal bad moral character. It instead supports our moral-ability hypothesis, whereby people assume good intentions and regard personal self-control as the ability to realize them. While strong ability can thus enhance good moral character, weakness cannot establish bad intentions or bad moral character on its own.

Some work suggests that moral inferences from self-control might be confined to cultures steeped in Protestant ethics who especially value self-restraint (Uhlmann et al., 2009), or depend on people's own self-control abilities (e.g., their motivation to denigrate or praise others' strong abilities; Minson & Monin, 2012). However, our results hold among American and Chinese adults alike (Experiment 2), and appear unaffected by participants' endorsement of Protestant ethics or self-reported self-control abilities (Experiment S3, 4). We also present preliminary evidence that moral inferences from willpower are conceptually distinct from the moralization of preferences (Feinberg et al., 2019), though further investigation seems warranted (Experiments 2, S2). Finally, our findings validate the distinction between morality and warmth (Goodwin, 2015; Goodwin et al., 2014), as well as the primacy of moral character in consequential contexts involving interpersonal trust (Experiments 3-4).

This research supports the notion that the contribution of some qualities to moral character judgments is conditional on intentions (Landy et al. 2016; Piazza et al., 2014), but suggests that this also applies to apparently non-moral abilities. Thus, we demonstrate that naïve judgments may diverge from scholarly assumptions that morality and competence are orthogonal. Though research has frequently examined morality versus competence separately (Brycz & Wojciszke, 1992; Phalet & Poppe, 1997; Skowronski & Carlston, 1987), little work has examined how they are perceived to relate. We suggest that exemplary moral character (i.e., beyond basic decency) may depend on self-control and many other non-moral competencies (e.g., social intelligence; Stellar & Willer, 2018).

Accordingly, such abilities may be morally relevant in a broader range of situations than previously assumed (Mooijman et al., 2018, 2020). The moral impact of their manifestations (i.e., good or bad realized outcomes) may also warrant further study, in light of established evaluative asymmetries (Inbar et al., 2012; Knobe, 2003). Finally, the moral

relevance of these factors may depend on individual differences that are not yet fully understood (e.g., between free will pessimists vs. optimists; Experiment 2). Our work highlights the many open questions about naïve moral inferences under realistic conditions in which outcomes are observable but intentions are not.

Navigating the social world entails constantly evaluating the character and trustworthiness of others. The current research indicates that people may regard apparently non-moral competencies as morally diagnostic. Inasmuch as good moral character is defined by achieving good moral outcomes, good intentions may not be enough.

References

Baumeister, R. F., & Exline, J. J. (2000). Self-control, morality, and human strength. *Journal of Social and Clinical Psychology*, *19*(1), 29–42. https://doi.org/10.1521/jscp.2000.19.1.29

Berman, J. Z. & Small, D. A. (2018). Discipline and desire: on the relative importance of willpower and purity in signaling virtues. *Journal of Experimental Social Psychology*.

Blasi, A. (2005). Moral character: a psychological approach. In D. K. Lapsley & F. C. Power (Eds.), *Character psychology and character education* (pp. 67-100). Notre Dame, IN: University of Notre Dame Press.

Brycz, H., & Wojciszke, B. (1992). Personality impressions on ability and morality trait dimensions. *Polish Psychological Bulletin, 23*(30), 223-236.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, *1*(1), 3–25. https://doi.org/10.1207/s15327957pspr0101

Cottrell, C. A., Neuberg, S. L., & Li, N. P. (2007). What do people desire in others? A sociofunctional perspective on the importance of different valued characteristics. *Journal of Personality and Social Psychology*, *92*(2), 208–231. https://doi.org/10.1037/0022-3514.92.2.208

Everett, J. A. C., Pizarro, D., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, *145*(6), 772–787. https://doi.org/10.1037/xge0000165

Everett, J. A. C., Clark, C. J., Meindl, P., Luguri, J. B., Earp, B. D., Graham, J., Ditto, P. H., & Shariff, A. F. (2021). Political differences in free will belief are associated with

differences in moralization. *Journal of Personality and Social Psychology, 120*(2), 461–483. https://doi.org/10.1037/pspp0000286

Gai, P. J. (2021, November 2). non-moral self-control and moral character. https://doi.org/10.17605/OSF.IO/F9BM5

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

Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, 24 (1), 38–44. https://doi.org/10.1177/0963721414550709

Feinberg, M., Kovacheff, C., Teper, R., & Inbar, Y. (2019). Understanding the process of moralization: How eating meat becomes a moral issue. *Journal of Personality and Social Psychology, 117*(1), 50–72. https://doi.org/10.1037/pspa0000149

Hofmann, W., Meindl, P., Mooijman, M., & Graham, J. (2018). Morality and self-control: how they are intertwined and where they differ. *Current Directions in Psychological Science*, *27*(4), 286–291. https://doi.org/10.1177/0963721418759317

Inbar, Y., Pizarro, D. A., & Cushman, F. (2012). Benefiting from misfortune: When harmless actions are judged to be morally blameworthy. *Personality and Social Psychology Bulletin*, *38*(1), 52–62. https://doi.org/10.1177/0146167211430232

Inzlicht, M., Legault, L., & Teper, R. (2014). Exploring the mechanisms of self-control improvement. *Current Directions in Psychological Science*, *23*(4), 302–307. https://doi.org/10.1177/0963721414534256

Knobe, J. (2003). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology, 16*(2), 309-324. https://doi.org/10.1080/09515080307771

Landy, J. F., Piazza, J., & Goodwin, G. P. (2016). When it's bad to be friendly and smart: The desirability of sociability and competence depends on morality. *Personality and*

Social Psychology Bulletin, *42*(9), 1272–1290. https://doi.org/10.1177/0146167216655984

Lapsley, D. K., & Lasky, B. (2001). Prototypic moral character. *Identity: An International Journal of Theory and Practice*, *1*(4), 345–363. https://doi.org/10.1207/S1532706XID0104_03

McCullough, M. E., & Carter, E. C. (2011). Waiting, tolerating, and cooperating: Did religion evolve to prop up humans' self-control abilities? In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation* (pp. 422–437). New York: The Guilford Press.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2008). Harnessing our inner angels and demons: what we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making. *Perspectives on Psychological Science*, *3*(4), 324–338. https://doi.org/10.1111/j.1745-6924.2008.00083.x

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*(1), 6–27. http://dx.doi.org/10.1037/met0000086

Mooijman, M., Meindl, P., Oyserman, D., Monterosso, J., Dehghani, M., Doris, J. M., & Graham, J. (2018). Resisting temptation for the good of the group: Binding moral values and the moralization of self-control. *Journal of Personality and Social Psychology*, *115*(3), 585–599. https://doi.org/10.1037/pspp0000149

Mooijman, M., Meindl, P., & Graham, J. (2020). Moralizing Self-Control. In A. R. Mele (Ed.), *Surrounding Self-Control* (pp. 259–274). Oxford Unviersity Press.

Minson, J. A., & Monin, B. (2012). Do-gooder derogation: Disparaging morally motivated minorities to defuse anticipated reproach. *Social Psychological and Personality Science, 3*(2), 200–207. https://doi.org/10.1177/1948550611415695

Paulhus, D. L., & Carey, J. M. (2011). The FAD-Plus: measuring lay beliefs regarding free will and related constructs. *Journal of Personality Assessment*, *93*(1), 96–104. https://doi.org/10.1080/00223891.2010.528483

Phalet, K., & Poppe, E. (1997). Competence and morality dimensions of national and ethnic stereotypes: A study in six eastern-European countries. *European Journal of Social Psychology*, *27*(6), 703-723. https://doi.org/10.1002/(SICI)1099-0992(199711/12)27:6<703::AID-EJSP841>3.0.CO;2-K

Piazza, J., Goodwin, G. P., Rozin, P., & Royzman, E. B. (2014). When a virtue is not a virtue: Conditional virtues in moral evaluation. *Social Cognition*, *32*(6), 528–558. https://doi.org/10.1521/soco.2014.32.6.528

Righetti, F., & Finkenauer, C. (2011). If you are able to control yourself, I will trust you: The role of perceived self-control in interpersonal trust. *Journal of Personality and Social Psychology*, *100*(5), 874–886. https://doi.org/10.1037/a0021827

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology, 52*(4), 689–699. https://doi.org/10.1037/0033-2909.105.1.131

Smith, K. D., Smith, S. T., & Christopher, J. C. (2007). What defines the good person? cross-cultural comparisons of experts' models with lay prototypes. *Journal of Cross-Cultural Psychology*, *38*(3), 333–360. https://doi.org/10.1177/0022022107300279

Stellar, J. E., & Willer, R. (2018). Unethical and inept? The influence of moral information on perceptions of competence. *Journal of Personality and Social Psychology, 114*(2), 195–210. https://doi.org/10.1037/pspa0000097

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271–324. https://doi.org/10.1111/j.0022-3506.2004.00263.x

Uhlmann, E. L., Poehlman, T. A., & Bargh, J. A. (2009). American moral exceptionalism. In J. T. Jost, A. C. Kay, & H. Thorisdottir (Eds.), *Social and psychological bases of ideology and system justification* (pp. 27–52). New York, NY: Oxford University Press.

Wertenbroch, K. (1998). Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing Science*, *17*(4), 317–337. https://doi.org/10.1287/mksc.17.4.317