

The Moralization of Effort

Jared B. Celniker¹, Andrew Gregory¹, Hyunjin J. Koo¹, Paul K. Piff¹, Peter H. Ditto¹, and Azim F. Shariff²

¹Department of Psychological Science, University of California, Irvine

²Department of Psychology, University of British Columbia

People believe that effort is valuable, but what kind of value does it confer? We find that displays of effort signal moral character. Eight studies ($N = 5,502$) demonstrate the nature of these effects in the domains of paid employment, personal fitness, and charitable fundraising. The exertion of effort is deemed morally admirable (Studies 1–6) and is monetarily rewarded (Studies 2–6), even in situations where effort does not directly generate additional product, quality, or economic value. Convergent patterns of results emerged in South Korean and French cross-cultural replications (Studies 2b and 2c). We contend that the seeming irrationality of valuing effort for its own sake, such as in situations where one's efforts do not directly increase economic output (Studies 3–6), reveals a “deeply rational” social heuristic for evaluating potential cooperation partners. Specifically, effort cues engender broad moral trait ascriptions, and this moralization of effort influences donation behaviors (Study 5) and cooperative partner choice decision-making (Studies 4 and 6). In situating our account of effort moralization into past research and theorizing, we also consider the implications of these effects for social welfare policy and the future of work.

Keywords: character judgment, effort, morality, partner choice, work

Supplemental materials: <https://doi.org/10.1037/xge0001259.supp>

How have so many humans reached the point where they accept that even miserable, unnecessary work is actually morally superior to no work at all? —David Graeber (2018), *Bullshit Jobs*

Is effort inherently valuable? From an economic perspective, the answer should be *no*: effort—active, goal-directed physical or

mental activity—should only be valuable insofar as it produces something of value. If the same outcome can be produced with less effort, it is rational to do so. This logic has motivated surges of automation throughout history, including current waves that may threaten nearly half of all jobs worldwide in the coming decade (Nedelkoska & Quintini, 2018).

Yet people's reasoning about the world does not always follow economic rationality (Boyer & Petersen, 2017; Kahneman, 2003). For example, imagine two office workers with identical jobs, but one employee clearly works harder than the other. Most people likely prefer the worker who exerts more effort, and may even want to pay that person more, because that individual seems like a more dedicated employee whose efforts could produce more for the company. Now imagine both workers were equally productive: Would you still prefer the hard worker even if their work did not result in greater production?

Several bodies of psychological research suggest the answer may be yes: Effort is linked to broad conceptions of value. For example, humans and other animals, such as pigeons and rats, place greater value on rewards generated through increased effort (Aronson & Mills, 1959; Clement et al., 2000; Lydall et al., 2010; for a review, see Inzlicht et al., 2018). In evaluating objects like art or clothing, people use the amount of labor required to produce the item as a heuristic for its quality and value (Kruger et al., 2004). These valuations of effort also extend to interpersonal judgments and behavior. Individuals perceived as lazy or who otherwise exhibit reduced effort (e.g., the poor or unemployed) are regularly dehumanized, devalued, and deemed less deserving of assistance (Harris & Fiske, 2011; Petersen et al., 2012). We

Jared B. Celniker  <https://orcid.org/0000-0001-5123-8455>

This material is based on work supported by the National Science Foundation Graduate Research Fellowship under Grant DGE-1839285 (to Jared B. Celniker) and was funded by a Canada 150 Research Chair from the Social Sciences and Humanities Research Council of Canada (to Azim F. Shariff). Prior versions of this article have been available on PsyArXiv since January 14, 2020. We thank members of the Hot Cognition Lab, Center for Applied Moral Psychology, and Melody M. Moore for their constructive feedback on this work. All preregistrations, data, and materials are available on our OSF page: <https://osf.io/zwqbe/>.

Jared B. Celniker, Paul K. Piff, Peter H. Ditto, and Azim F. Shariff developed and designed Studies 1–6. Andrew Gregory contributed to the development and design of Studies 4–6. Hyunjin J. Koo contributed to the development and design of Studies 2b, 2c, 4, and 6. Data collection and analyses were performed by Jared B. Celniker (Studies 1–6), Hyunjin J. Koo (Studies 2b and 2c), and Andrew Gregory (Study 5). Jared B. Celniker drafted the manuscript in conjunction with Azim F. Shariff, and all authors provided critical revisions and approved the final version of the manuscript for submission.

Correspondence concerning this article should be addressed to Jared B. Celniker, Department of Psychological Science, University of California, Irvine, 4201 Social and Behavioral Sciences Gateway, Irvine, CA 92697, United States. Email: jcelnike@uci.edu

propose that the valuation of effort in interpersonal settings is primarily driven by a moralization of effort; displays of effort, even those that produce little or no material value, are ascribed with moral value.

Effort has been directly tied to interpersonal moral valuations in research on Protestant Work Ethic (PWE) beliefs (Weber, 1904/1958). The PWE is a suite of cultural beliefs primarily defined by the endorsement of traditional morality and individual achievement (Uhlmann & Sanchez-Burks, 2014). One recognized feature of PWE beliefs is the appreciation of economic productivity as a moral end unto itself. These beliefs have been invoked to explain individual and societal differences in the valuing of hard work (Furnham et al., 1993; van Hoorn & Maseland, 2013). Recent research suggests that PWE beliefs in American culture may cause individuals to heuristically associate work ethic with being a moral person—for example, hardworking laborers are perceived by Americans as more honest than lazier counterparts (Amos et al., 2019).

These findings are suggestive of a culturally bound explanation of effort moralization; nevertheless, several limitations preclude strong interpretations. First, the role of PWE has been assumed but rarely measured or tested (e.g., Amos et al., 2019; Uhlmann & Sanchez-Burks, 2014), so there is little direct evidence linking PWE to effort moralization. Second, these moral judgments have not been disentangled from other facets of person perception (e.g., warmth and competence; Goodwin, 2015), so effort cues may engender broad halo effects rather than moralization specifically. Third, and most importantly, productivity has not been held constant in these studies; working harder is frequently associated with increased productivity, so participants may have moralized implied production rather than effort in past research.

Preferring a more industrious and productive worker is sensible and unsurprising; more curious, though, are indications that people value hard work even when those efforts do not produce direct economic benefits. For instance, the phenomenon of “bullshit jobs” (Graeber, 2018)—societally useless or redundant work undertaken merely out of monetary or social obligations—suggests that economically inefficient effort is maintained in the workplace. Although difficult to quantify, initial evidence suggests bullshit jobs (e.g., unnecessary middle management) may account for roughly 20% to 40% of work in some Western economies (Amárah Research, 2019; Dahlgreen, 2015; Schouten & Nelissen, 2017). Furthermore, bullshit labor, such as superfluous administrative work that could be easily eliminated or automated, is a reportedly common feature of otherwise productive jobs (including within academia). Such observations are perplexing from a Western economic and cultural perspective, in which costly but inefficient labor should be trimmed by market and social forces. Indeed, a core tenet of PWE beliefs, in addition to valuing hard work, is valuing efficiency and frugality; the very beliefs argued to cause the valorization of economically productive effort should simultaneously promote the denigration of economically frivolous effort. But this possibility is undermined by select research. Namely, people positively evaluate those who engage in “needless work,” such as continuing to work after winning the lottery (Tierney et al., 2020). Research on the “martyrdom effect” similarly finds that people report greater willingness to donate to fundraisers that involve greater effort, even when one’s efforts are untethered

from the fundraiser’s cause (Olivola & Shafir, 2013). Running a race does not directly contribute more to finding a cure than simply requesting donations, yet people find advocacy paired with economically unnecessary effort—effort that does not directly increase economic output or personal compensation—more valuable than advocacy alone. These findings underscore the possibility that effort is valued for reasons other than the economic value it directly produces.

The inability of economic reasoning or PWE beliefs to explain these phenomena suggests that the perceived value of effort may be better explained by a more fundamental psychological process. Displays of altruism once presented a similar puzzle. However, theoretical and empirical work on “costly signaling” and “competitive altruism” provided a fruitful explanation: although costly, these displays signal the moral traits that people find most important when selecting social (Goodwin, 2015) and cooperative partners (Gintis et al., 2001). People compete to be chosen as cooperation partners, so they engage in altruistic displays to distinguish themselves from their rivals and be seen as a more attractive option within the market of available cooperation partners (Barclay, 2013). We contend that displays of effort function in similar ways. Although effort is an inconsistent and “noisy” cue of ability and productivity (Markovits, 2019; Shepperd et al., 1994; Stibbard-Hawkes et al., 2018), effort may be a reliable signal of one’s cooperative intent. Whereas prior research has found that effort cues amplify how much praise or blame individuals receive for prosocial and antisocial actions, respectively (Bigman & Tamir, 2016), our claim is slightly different. We argue that effort itself is perceived as a costly signal of moral character, even when one’s efforts are devoted to tasks that do not have direct moral or economic consequences. Someone who invests effort in one task may be seen as a preferable partner for future cooperative tasks—as one who will not take a “free ride” on the backs of others’ collective efforts (Cosmides & Tooby, 2013). Thinking about effort moralization as a “deeply rational” heuristic process, one that facilitates social decision-making irrespective of its economic irrationality (Kenrick et al., 2009), may provide stronger theoretical foundations for understanding why displays of effort increase perceptions of one’s moral character—even in situations where one’s efforts are not materially beneficial or necessary. Observing such effects across relevant individual, situational, and cultural differences would further suggest that effort moralization is more widespread than can be accounted for by cultural explanations alone.

These conjectures offer a set of testable hypotheses. First, if displays of effort are used to convey one’s value as a cooperation partner, then effort cues should specifically and reliably relate to judgments of moral character. Second, if effort serves as a signal of moral character, people should preferentially select those who display higher levels of effort as cooperation partners. Third, if the moralization of effort is a more universal phenomenon than previously recognized, then it should be observed across individual and cultural differences in work ethic beliefs, as well as across domains of behavior.

Overview of Studies

We report eight studies that examine the relationships between perceptions of effort and judgments of moral and monetary value.

Our seven experiments (Studies 2–6) test whether effort is morally valued, even in situations where one’s efforts do not directly increase material value, and whether such moral character judgments guide subsequent decisions regarding scarce monetary allocations (wages and donations) and cooperative partner choice. Furthermore, we assess whether these effects are moderated by individual differences in PWE beliefs and extend beyond the United States to other cultural contexts.

Study 1 examines whether perceived effort is a specific predictor of people’s moral evaluations of different types of laborers. Studies 2a–2c test whether people judge a high-effort worker to be more moral and deserving of greater pay than a similar low-effort worker in the United States (Study 2a), South Korea (Study 2b), and France (Study 2c). Study 3 investigates whether an individual engaging in economically unnecessary work is deemed more moral than an individual who does not engage in such work. Study 4 evaluates whether individuals who put more effort into a personal activity (running) are seen as more moral and, consequently, selected more frequently as a cooperation partner for a trust-based task. Study 5 extends our investigation of effort moralization to prosocial behavior, testing whether an individual who runs a marathon for a charitable cause is seen as more moral and, in turn, accrues more donations than an individual who runs a shorter race. Last, Study 6 addresses limitations of the prior studies by explicitly controlling for the economic value of targets’ efforts and by showing that effort cues influence two distinguishable sets of moral characteristics: traits directly related to one’s cooperative intent and traits related to one’s capacity to enact those cooperative intentions. By refining these moral trait constructs and testing different models of the experimental effects, Study 6 also provides more detail about the causal relationships between perceptions of effort, moral character evaluations, and cooperative partner choice. All our studies received ethics committee approval at our respective institutions, seven studies (Studies 2–6) were preregistered in advance of data collection, and all our preregistrations, data, and materials are available on our OSF page (<https://osf.io/zwqbe/>; Celniker et al., 2020).

Study 1

In an initial investigation, we explored whether people whose jobs are perceived to require greater effort are judged to be more moral than those employed in less effortful work, over and above factors like the job’s societal contribution.

Method

We recruited 755 U.S. adults (376 female, five other; age: $M = 37.2$, $SD = 11.8$) through Amazon Mechanical Turk (MTurk) to give their general impressions of various jobs and the people who do them. Each participant was presented with 10 jobs randomly drawn from a bank of 46 jobs (see materials on our OSF page) selected to be a roughly representative sampling of 702 occupations analyzed in Frey and Osborne (2013).

For each job, participants responded to the following questions on 7-point scales: “How much effort does this job require?,” “How difficult is the work in this job?,” “How much do others respect people who do this job?,” “How much does this job

contribute to society?,” “How financially compensated are the people who do this job?,” and “How moral are the people who do this job?” Participants reported their opinions of all 10 jobs in one domain (such as effort) before proceeding to the question in another domain (such as morality). The order in which the measures were presented was randomized.¹

After the job characteristic measures for all 10 jobs, participants completed the Protestant Work Ethic (PWE) Scale (Mirels & Garrett, 1971), six items on Work Ethos tapping perceptions of work as good from the 2000–2004 wave of the World Values Survey, and demographic measures (age, sex, social political orientation, economic political orientation, subjective socioeconomic status, and objective socioeconomic status; the full materials are available on our OSF page).

Results

We analyzed these data using linear mixed multilevel models with the GAMLj package in jamovi. The models were specified to predict moral judgments from the fixed effects of perceived effort, difficulty, respect, contribution to society, and financial compensation. The models included random intercepts for participants and jobs to account for the nonindependence of responses and to allow us to generalize our findings to the broader populations of individuals and jobs from which we sampled (Judd et al., 2012). Random slopes for effort by participant and job were also included to account for the variability in effort moralization elicited across subjects and professions (Barr et al., 2013). Mean-centered fixed effects for the demographic variables were entered into the models as well. Lastly, the models included mean-centered fixed effects for PWE and Work Ethos scores and their interactions with the effort variable. Separate models were specified using PWE and Work Ethos scores. Only the model including PWE scores is presented in the main text; the models using Work Ethos scores are presented in the online supplemental materials and produce substantively similar results.

The fixed effect estimates from the full model are presented in Table 1. When controlling for the other fixed effects as well as participant- and job-level² variability, effort was still a significant positive predictor of moral evaluations, $b = .09$, 95% CI [.05, .12], $p < .001$. Although there was a significant main effect of PWE, $b = .15$, 95% CI [.04, .26], $p = .007$, the interaction between PWE and effort was not significant, $b = .02$, 95% CI [–.02, .05], $p = .27$. This indicates that, although those endorsing a stronger Protestant Work Ethic generally rated jobs as more moral, individual differences in these beliefs did not moderate the relationship between effort and moral character judgments—participants

¹ As part of a separate research question, half the participants in Study 1 were assigned to answer the question, “How likely do you think it is for this job to be automated in the near future?” before completing the rest of the dependent measures. The manipulation did not substantively impact the results of this study, so we report a model collapsed across conditions in the main text. Additional analyses that model the effect of the manipulation are presented in the online supplemental materials. Ultimately, perceptions of automatability did not moderate the relationship between effort and moral character, our primary interest.

² While the inclusion of random intercepts and slopes significantly improved model fit, the association between effort and moral evaluations was positive for 44 of the 46 jobs. See the online supplemental materials for more information about the random effects.

Table 1
Fixed Effects Predicting Moral Character in Study 1

Predictor	Estimate (<i>b</i>)	95% confidence interval	<i>SE</i>	<i>p</i>
Intercept	4.59	[4.51, 4.68]	0.05	<.001
Job characteristics				
Effort	0.09	[0.05, 0.12]	0.02	<.001
Contribution	0.19	[0.17, 0.21]	0.01	<.001
Difficulty	0.03	[0.01, 0.05]	0.01	.007
Respect	0.10	[0.09, 0.12]	0.01	<.001
Compensation	−0.10	[−0.11, −0.08]	0.01	<.001
Individual differences				
Sex	−0.14	[−0.25, −0.03]	0.06	.017
Age	0.00	[−0.00, 0.01]	0.00	.12
Social politics	−0.05	[−0.10, 0.003]	0.03	.065
Economic politics	0.03	[−0.02, 0.08]	0.03	.26
Subjective SES	0.04	[0.01, 0.08]	0.02	.027
Objective SES	0.01	[−0.03, 0.05]	0.02	.72
PWE	0.15	[0.04, 0.26]	0.06	.007
PWE × Effort	0.02	[−0.02, 0.05]	0.02	.27

Note. All fixed effects are mean centered, except for sex. The sex variable was dummy coded (0 = *Male*, 1 = *Female*).

varying in PWE endorsement did not significantly differ in the extent to which their effort evaluations predicted moral judgments.

Discussion

Study 1 showed that across a variety of jobs, people who were seen as doing more effortful work were perceived as more moral, even when accounting for other relevant job characteristics and individual differences. This provides initial support for a relationship between effort evaluations and moral judgments, over and above variables like the perceived societal contributions of the jobs. By modeling participants and jobs as random effects, we can be more confident that this effect exists among the populations of people and professions from which we sampled (Yarkoni, 2022). Study 1 also shows that the moralization of effort emerges irrespective of individual differences in PWE beliefs, indicating this effect may not be dependent on that set of cultural values.

Study 2a

In Study 2a, we experimentally tested the relationship between perceived effort and morality using vignettes about two individuals doing the same work. We extend prior work on interpersonal evaluations of workers (e.g., Amos et al., 2019) by examining whether effort specifically signals moral virtues and, crucially, if it does so even when the amount of work output is held constant.

Method

Procedure

We aimed to recruit at least 480 total participants (240 per job condition) to have more than 80% power to detect small ($d = .20$) within-subjects effects (two-tailed, $\alpha = .05$) in each condition. Ultimately, 486 U.S. adults (256 female, three other; age: $M = 38.6$, $SD = 13.0$) were recruited through MTurk.

After consenting, participants were assigned to one of four experimental conditions as part of a 2 (job condition: factory worker or

accounting) × 2 (first target effort: low or high) mixed-factors design. A pilot study with a student sample, which only employed the factory worker condition, found materially identical results (see online supplemental materials). Here we included the accounting condition to test generalizability across types of jobs.

Each participant read two vignettes, one of a high-effort target and one of a low-effort target. The order in which target vignettes were presented varied by first target condition. If a low-effort prompt was presented first, participants started the experiment by reading the following:

Justin/Mark works at (an accounting firm auditing financial disclosure statements/in a factory making widgets). Justin/Mark is able to (audit approximately 10 statements per week, around two statements every day/produce approximately six widgets per hour, one widget around every 10 minutes). For Justin/Mark, (auditing financial disclosure statements/making widgets) requires minimal effort—while he works as quickly as possible, it is easy work.

Participants then evaluated this first target on a series of dependent variables in randomized order. After completing those items, participants were presented with the description of the second target, which varied from the first target only in the target's name (Justin/Mark) and effort exerted (low/high). The high-effort prompts were identical to the low-effort prompts besides the last sentence, which read, "For Mark/Justin, (auditing financial disclosure statements/making widgets) requires a lot of effort - while he works as quickly as possible, it is hard work." Participants subsequently completed an identical set of dependent variables for the second target. Targets' names were counterbalanced across conditions.

Measures

Participants indicated how well they thought each of 15 traits described the targets using 7-point scales (1 = *Does not describe Justin/Mark well*, 4 = *Describes Justin/Mark moderately well*, 7 = *Describes Justin/Mark extremely well*). The traits were selected to

capture evaluations of competence (“Competent,” “Talented,” “Logical,” “Organized,” and “Intelligent”), warmth (“Warm,” “Agreeable,” “Sociable,” “Funny,” and “Happy”), and morality (“Moral,” “Responsible,” “Dedicated,” “Honest,” and “Principled”) and were presented in randomized order. The characteristics were selected from previously established items shown to discriminate between these three dimensions of person perception (Goodwin, 2015).

Participants then responded to 15 face-valid questions on 7-point scales, such as, “How much effort do you think Justin/Mark puts into his work?” (1 = *No effort at all*, 4 = *An average amount of effort*, 7 = *A lot of effort*), “What quality of (widgets/audits) do you think Justin/Mark produces?” (1 = *Very low quality*, 4 = *Average quality*, 7 = *Very high quality*), “Compared with other jobs, how difficult is Justin/Mark’s job?” (1 = *Not at all difficult*, 4 = *Moderately difficult*, 7 = *Extremely difficult*), and “How much do you think Justin/Mark suffers?” (1 = *Does not suffer at all*, 4 = *Suffers an average amount*, 7 = *Suffers a lot*). The remaining items tapped into additional dimensions of person perception not captured in the trait evaluations (e.g., trustworthiness). Additionally, one item assessed how much participants believed each target deserved to be paid per hour.³ After completing all dependent measures, participants completed the PWE scale (Mirels & Garrett, 1971) and demographic measures. The wordings for all the items used in our studies are available in the materials section on our OSF page (<https://osf.io/zwqbe/>; Celniker et al., 2020).

Results

Within Subjects

Following our preregistration, we excluded participants who did not rate the high-effort target as exerting more effort than the low-effort target. This resulted in a final sample of 384 participants (the results are substantively identical when including all participants). Indices of competence (Cronbach’s $\alpha_{\text{high-effort}} = .92$, $\alpha_{\text{low-effort}} = .89$), warmth ($\alpha_{\text{high-effort}} = .90$, $\alpha_{\text{low-effort}} = .86$), and morality ($\alpha_{\text{high-effort}} = .92$, $\alpha_{\text{low-effort}} = .91$) were used as the dependent measures. Further following our preregistration, we collapsed responses across job conditions⁴ and entered participants’ evaluations of both targets into paired samples *t* tests for the primary analyses (for full descriptives and secondary analyses, see the online supplemental materials).

Participants rated the high-effort target as putting significantly more effort into his work ($M = 6.1$, $SD = 1.0$) than the low-effort target ($M = 3.3$, $SD = 1.3$), $t(383) = 39.69$, $p < .001$, $d = 2.03$, 95% CI [1.85, 2.20]. As predicted, the high-effort target was seen as significantly more moral ($M = 5.3$, $SD = 1.1$) than the low-effort target ($M = 4.6$, $SD = 1.0$), $t(383) = 11.71$, $p < .001$, $d = .60$, 95% CI [.49, .71]. In contrast, the high-effort target ($M = 4.3$, $SD = 1.0$) was seen as significantly *less* warm than the low-effort target ($M = 4.5$, $SD = 1.0$), $t(383) = -3.49$, $p < .001$, $d = -.18$, 95% CI [-.28, -.08]. There were no significant differences in perceived competence between the high-effort ($M = 4.9$, $SD = 1.2$) and low-effort targets ($M = 4.9$, $SD = 1.1$), $t(383) = -.47$, $p = .64$, $d = -.02$, 95% CI [-.12, .08]. Despite being perceived as less warm and generating the same amount of output—yet tracking with judgments of moral character—the high-effort target was seen as deserving a higher hourly wage ($M = \$1.0$, $SD = \$2.1$) than the low-effort

target ($M = \$.3$, $SD = \$2.1$), $t(383) = 6.64$, $p < .001$, $d = .34$, 95% CI [.24, .44].

Mediation Analysis. In exploratory analyses conducted using the jAMM package in jamovi, we examined whether perceived morality mediated the effect of perceived effort on pay deservingness. Difference scores were created for this analysis by subtracting the low-effort target rating from the high-effort target ratings for each variable. The difference scores for effort perceptions were entered as the independent variable, the difference scores for moral character were entered as the mediator, and the difference scores for perceived pay deservingness were entered as the dependent variable. Despite the correlational nature of these within-subjects analyses, our experimental design provides controlled conditions in which we can yield evidence consistent with a causal effect of differences in perceived effort on differences in perceived moral character and pay deservingness (Judd et al., 2012). Additionally, to control for potential alternative mechanisms, a model including differences in perceived work quality, job difficulty, and suffering as parallel mediators was also constructed. For these and all other mediation models reported in this paper, confidence intervals were calculated using 1,000 bootstrap replications.

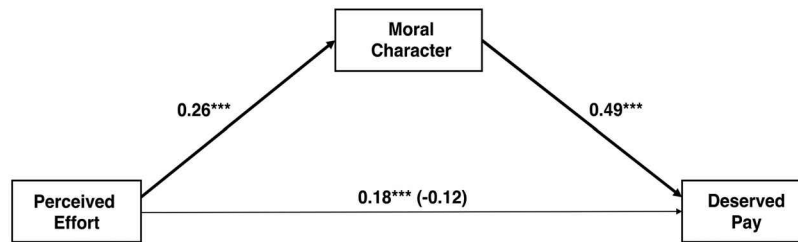
Critically, the indirect effect of perceived effort on pay deservingness through moral character judgment was significant. This was true when moral character was entered as the only mediator, $b = .22$, 95% CI [.11, .32], $p < .001$, and when controlling for other potential mechanisms, $b = .13$, 95% CI [.04, .21], $p = .003$ (presented in Figure 1). In the latter model, there were significant indirect paths through perceived work quality, $b = .11$, 95% CI [.04, .18], $p = .004$, and job difficulty, $b = .07$, 95% CI [.02, .12], $p = .009$, but not through suffering, $b = -.01$, 95% CI [-.04, .02], $p = .46$. While conceptually replicating the finding that effort is used as a heuristic of quality and value (Kruger et al., 2004), we documented a separate predictor—moral character judgment—through which effort cues influenced the perceived value of human labor.

Moderation Analyses. Using the medmod package in jamovi, we explored whether participants’ PWE scores moderated the effects of effort on moral character, effort on pay deservingness, and moral character on pay deservingness. As in our mediation analysis, we used difference score variables and conducted these analyses collapsed across job conditions (though see the online supplemental materials for some differences in results across job conditions). For these and all other moderation models reported in this paper, standard errors were calculated using 1,000 bootstrap replications. The key interaction terms between PWE and the predictor variables were not significant in any of these analyses: PWE scores did not moderate the effects of effort predicting moral character, $b = .09$, 95% CI [-.11, .31], $p = .44$, effort predicting pay deservingness, $b = -.27$, 95% CI [-.63, .18], $p = .18$, or moral

³ For the deserved pay measure in Study 2a, participants responded on a sliding scale ranging from \$6–\$18 anchored at the scale midpoint of \$12 the factory condition, whereas in the accounting condition the range of pay (\$16–\$28) and scale anchor (\$22) differed but spanned the same range. The deserved pay variable collapsed across conditions was analyzed on a scale of –\$6 to \$6.

⁴ Repeated measures ANOVAs indicated that there were no significant interactions with job condition for the main dependent measures ($ps = .15-.84$), so we report *t* test analyses for ease of effect size interpretation.

Figure 1
 Mediation Model for Study 2a Showing the Effect of Differences in Perceived Effort on Differences in the Deserved Pay of the Targets, as Mediated by Differences in Perceived Moral Character



Note. The presented path estimates control for differences in perceived work quality, work difficulty, and suffering. Unstandardized coefficients are displayed. On the center path, the coefficient outside the parentheses is the total effect, and the coefficient inside the parentheses is the direct effect. Asterisks indicate significant paths (***) $p < .001$.

character predicting pay deservingness, $b = -.15$, 95% CI $[-.45, .41]$, $p = .50$.

Between Subjects

Although we designed this study to test within-subjects differences, our experimental design also allowed for between-subjects analyses. Specifically, participants were randomly assigned to evaluate either the high- or low-effort target first, and these first-target judgments were entered into Welch's independent samples t tests to examine condition differences (Delacre et al., 2017).

The results of these secondary analyses converged with the within-subjects results. Participants rated the high-effort target as putting more effort into their work ($M = 6.1$, $SD = 1.1$) than the low-effort target ($M = 5.3$, $SD = 1.1$), $t(366.46) = 20.71$, $p < .001$, $d = 2.13$, 95% CI $[1.84, 2.42]$, and most importantly, the high-effort target was seen as significantly more moral ($M = 5.3$, $SD = 1.1$) than the low-effort target ($M = 4.7$, $SD = 1.0$), $t(381.71) = 5.54$, $p < .001$, $d = .56$, 95% CI $[.36, .77]$. The high-effort target was also seen as deserving greater pay ($M = \$1.2$, $SD = \$2.1$) than the low-effort target ($M = $.3$, $SD = $.2$), $t(376.15) = 4.35$, $p < .001$, $d = .45$, 95% CI $[.24, .65]$.

Additionally, between-subjects mediation and moderation analyses yielded results that mirrored those of the within-subjects analyses. The effect of condition on perceptions of moral character was significantly mediated by perceptions of effort, and the effect of condition on perceptions of pay deservingness was significantly mediated by perceptions of moral character, with and without controlling for perceived work quality, work difficulty, and suffering. However, it is important to note that power analyses for indirect effects (Schoemann et al., 2017) indicated that these mediational analyses were underpowered. The moderation analyses yielded nonsignificant results with one exception: there was a significant interaction between first-target moral characters judgments and PWE scores in predicting first-target pay deservingness, $b = .26$, 95% CI $[.04, .49]$, $p = .026$. The relationship between moral character and pay deservingness was stronger for those with stronger PWE beliefs. Nevertheless, the association between moral character judgment and pay deservingness was significant across levels of participants' PWE endorsement. Please see the online supplemental materials for

more details regarding these analyses and the full set of between-subjects results.

In sum, manipulating the effort exerted by the first target caused differences in the moral character and wage deservingness judgments of that target. High-effort workers were perceived as more moral and deserving of more money than low-effort workers. These secondary, between-subjects results largely aligned with those of the primary, within-subjects analyses.

Studies 2b and 2c

In Study 2a, U.S. participants moralized effort independent of their PWE endorsement, further calling into question the cultural basis of effort moralization. In Studies 2b and 2c, we employed the same experimental paradigm in two additional countries—South Korea (2b) and France (2c)—to more rigorously test the generalizability of effort moralization using the triangulation approach.

The triangulation approach involves researchers first examining a phenomenon in two cultures that differ on one theoretically important dimension before examining the same phenomenon in a third culture that differs from either of the first two on a separate theoretical attribute (Norenzayan & Heine, 2005). Through this process, one can gather convergent evidence about the universality of a psychological phenomenon. In this case, the United States and South Korea are considered Western and Eastern cultures, respectively, whereas the United States and France are both considered Western. At the same time, whereas American and South Korean citizens work longer hours than most countries in the Organization for Economic Cooperation and Development (OECD), the French are far below OECD averages in the number of hours worked per year (OECD, 2019) and embrace less stringent and moralized work ethic norms than Americans (Lamont, 2000).

Method

We aimed to recruit economically representative samples (approximately 100 participants from each of five income brackets) of at least 500 South Korean and French residents, respectively, to have over 90% power to detect small ($d = .15$) within-subjects effects (two-tailed, $\alpha = .05$) in each study. We ultimately recruited

532 South Korean participants via Dataspring, a Korean survey company (Study 2b), and 521 French participants via Qualtrics Panels (Study 2c). Participant demographics are presented in the online supplemental materials.

The procedures were identical to Study 2a, with four exceptions. First, since the pattern of results was identical across the two occupations in Study 2a, we only used the vignette involving accountants in Studies 2b and 2c. Second, all study instruments were translated from English to Korean and French using a standard back-translation method (Brislin, 1970), and the target names were changed to common Korean names (Kyoungsoo and Junho) in Study 2b and common French names (Michel and Nicolas) in Study 2c. Translated materials are available in the online supplemental materials. Third, participants in Study 2c responded to a truncated set of exploratory dependent measures (e.g., trustworthiness). Fourth, after responding to the key questions for Studies 2b and 2c, participants answered several other questions. These were preregistered as exploratory for a separate project not relevant to this article and are thus not mentioned further.

Results

Within Subjects

Following our preregistration, participants who completed the survey in less than three minutes, rated the low-effort target as exerting equal or greater effort than the high-effort target, or failed an attention check (“To ensure that the survey is working properly, please choose ‘7: strongly agree’”) were excluded from analyses.⁵ This resulted in final samples of 322 participants in Study 2b (South Koreans) and 350 participants in Study 2c (French). As in Study 2a (Americans), the results of Studies 2b and 2c are substantively identical when including all participants. The composite measures of morality, warmth, and competence were sufficiently reliable in these samples (Study 2b Cronbach’s α s = .81–.93; Study 2c Cronbach’s α s = .89–.92).

Study 2b: South Korean Sample. The high-effort target was perceived as putting in more effort ($M = 5.9$, $SD = .9$) than the low-effort target ($M = 3.9$, $SD = 1.1$), $t(321) = 33.01$, $p < .001$, $d = 1.84$, 95% CI [1.66, 2.02]. Crucially, the high-effort target was again perceived as more moral ($M = 4.7$, $SD = 1.1$) than the low-effort target ($M = 4.0$, $SD = 1.1$), $t(321) = 12.75$, $p < .001$, $d = .71$, 95% CI [.59, .83]. Unlike in the American sample, the high-effort target was not seen as significantly less warm ($M = 3.0$, $SD = 1.3$) than the low-effort target ($M = 3.1$, $SD = 1.3$), $t(321) = -1.13$, $p = .26$, $d = -.06$, 95% CI [-.17, .05], yet the high-effort target was perceived as significantly less competent ($M = 4.3$, $SD = 1.3$) than the low-effort target ($M = 4.7$, $SD = 1.2$), $t(321) = -5.59$, $p < .001$, $d = -.31$, 95% CI [-.42, -.20] (see Figure 2). Also mirroring Study 2a, the high-effort target was seen as deserving higher pay (in KRW: $M = 26,074.1$, $SD = 3,945.4$) than the low-effort target (in KRW: $M = 24,927.5$, $SD = 4,442.9$), $t(319) = 5.47$, $p < .001$, $d = .31$, 95% CI [.19, .42].⁶

Study 2c: French Sample. The high-effort target ($M = 6.0$, $SD = 1.0$) was again rated as putting in more effort than the low-effort target ($M = 3.4$, $SD = 1.3$), $t(349) = 34.17$, $p < .001$, $d = 1.83$, 95% CI [1.66, 2.00]. As predicted, and replicating the results of the previous studies, the high-effort target was perceived as

more moral ($M = 5.1$, $SD = 1.2$) than the low-effort target ($M = 4.7$, $SD = 1.3$), $t(349) = 7.18$, $p < .001$, $d = .38$, 95% CI [.28, .49]. Unlike in the U.S. or South Korea, the French sample rated the high-effort target as both significantly less warm ($M = 3.6$, $SD = 1.2$) than the low-effort target ($M = 3.9$, $SD = 1.3$), $t(349) = -5.85$, $p < .001$, $d = -.31$, 95% CI [-.42, -.21], and as significantly less competent ($M = 4.9$, $SD = 1.2$) than the low-effort target ($M = 5.3$, $SD = 1.2$), $t(349) = -5.33$, $p < .001$, $d = -.29$, 95% CI [-.39, -.18] (see Figure 2). Nevertheless, despite being seen as less warm and competent, the high-effort target was still seen as deserving higher pay ($M = €23.5$, $SD = €2.7$) than the low-effort target ($M = €23.00$, $SD = €2.9$), $t(349) = 4.70$, $p < .001$, $d = .25$, 95% CI [.15, .36].

Within-Subjects Mediation Analyses. We conducted identical mediation analyses to the ones specified in Study 2a to examine whether perceived morality mediated the effect of effort on pay deservingness in each sample. In models without control variables, the indirect effect through moral character was significant in both Study 2b, $b = 418.41$, 95% CI [189.80, 653.75], $p < .001$, and Study 2c, $b = .08$, 95% CI [.06, .16], $p = .04$. In models controlling for work quality, job difficulty, and suffering, the indirect effect through moral character was significant in Study 2b, $b = 227.66$, 95% CI [54.30, 399.22], $p = .01$, but not in Study 2c, $b = .04$, 95% CI [-.01, .08], $p = .13$. For Study 2c, the path from effort to moral character held when controlling for the same covariates, $b = .15$, 95% CI [.06, .25], $p = .002$, yet the path from moral character to pay deservingness was only marginally significant, $b = .24$, 95% CI [.00, .49], $p = .056$. However, post hoc power analyses for indirect effects (Schoemann et al., 2017) indicated that we were underpowered to detect the indirect effect through moral character when including the control variables in Study 2c.⁷ The full model statistics are presented in the online supplemental materials.

Within-Subjects Moderation Analyses. As in Study 2a, we explored whether participants’ PWE scores moderated the effects of effort on moral character, effort on pay deservingness, and moral character on pay deservingness. We conducted these analyses separately for each sample, and we used the same difference score variables that were calculated for our mediation analyses. In sum, PWE beliefs did not moderate any of these three effects in either Study 2b or 2c ($ps = .19$ –.95; full statistics are presented in the online supplemental materials). As in the United States, individual differences in PWE endorsement did not predict differences in effort moralization in South Korea or France.

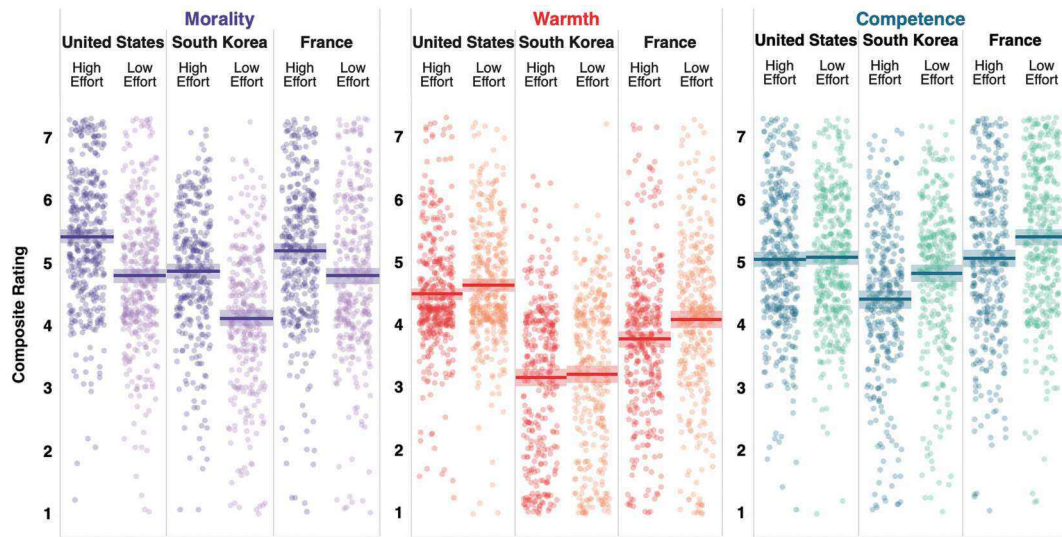
⁵ Our preregistration for Study 2b indicated that we would conduct a multilevel model as a secondary analysis to compare the results of Study 2a and 2b. However, this analysis was poorly conceived in terms of statistical power and was not conducted.

⁶ The test of pay deservingness follows directly from Study 2a but, owing to experimenter oversight, was not preregistered as a key question.

⁷ Supporting the notion that the mediation analysis in Study 2c was underpowered when including control variables, identical analyses using the full sample of 521 participants, rather than just the 350 who passed our inclusion criteria, resulted in a significant indirect effect, $b = 0.04$, 95% CI [0.01, 0.06], $p = .004$.

Figure 2

Mean Within-Subjects Ratings of Each Target on the Person–Perception Composite Measures From Studies 2a–2c



Note. Error bands represent standard errors of the mean. Results for the United States (Study 2a) are collapsed across job conditions. See the online article for the color version of this figure.

Between Subjects

Secondary between-subjects analyses were also conducted for Studies 2b and 2c. Like in the U.S. sample, these analyses were conducted using only the first target description the participants received, and the results largely converged with those of the within-subjects analyses. Most critically, the high-effort targets were deemed more moral than the low-effort targets in both studies, and the condition differences in moral character judgments were explained by condition differences in perceived effort (with and without controlling for perceived work quality, work difficulty, and suffering). Although there was some evidence of moral character judgments explaining the effect of first-target condition on perceptions of first-target pay deservingness in both studies, these results were generally inconclusive due to a lack of statistical power (Schoemann et al., 2017). Additionally, most of the between-subjects moderation analyses (four of six for Study 2b, five of six for Study 2c) yielded nonsignificant results, suggesting that PWE beliefs had a limited influence on participants' evaluations. Please see the online supplemental materials for more details regarding the between-subjects analyses for these studies.

Study 2a–2c Discussion

Studies 2a–2c found that participants perceived high-effort workers as more moral than low-effort workers. Our effort manipulation consistently increased moral trait ascriptions but not warmth or competence evaluations, and moral character judgments influenced subsequent judgments about how much money the targets deserved. These findings held when controlling for alternative mechanisms, across individual differences in PWE endorsement, and in samples from the United States, South Korea, and France.

Together, these studies suggest that effort moralization is a generalizable phenomenon and not reducible to PWE beliefs.

Study 3

Although we attempted to hold the economic productivity of the targets constant across Studies 2a–2c, participants perceived the more effortful target as producing work of higher quality. Consequently, participants may have seen the high-effort target as engaging in more valuable work than the low-effort target. If people infer greater morality from greater effort—even when economic productivity is held constant—then effort moralization should occur even when an individual's labor does not increase the amount or quality of output produced relative to an effortless alternative. In other words, individuals should perceive an individual who chooses to engage in labor as more moral than one who avoids labor, even when this labor is materially unnecessary and produces no relative economic value. Study 3 examined this prediction.

Method

Procedure

We aimed to recruit 800 participants to have at least 80% power to detect a small ($d = .2$) between-subjects effect (two-tailed, $\alpha = .05$). Ultimately, 801 U.S. adults (377 female, seven other; age: $M = 37.6$ years, $SD = 12.1$) were recruited through MTurk and randomly assigned to one of two experimental conditions.

In each condition, participants were asked to read a short vignette about Geoff, a medical scribe, and were presented with the same first three paragraphs of the vignette:

Geoff is a single man who works as a medical scribe. In this role, Geoff works with doctors by preparing and documenting patient information during medical visits. He signed a three-year contract with the hospital for his position and works 40 hr per week on average. Geoff is starting the second year of his contract with the hospital.

Owing to recent technological advances, the hospital acquired free scribe software that will perform all the tasks of a medical scribe with the same high quality. This software would replace Geoff's job. The hospital told Geoff that, because they are able to implement the software at no cost, they can continue to pay Geoff's salary for the final years of his contract regardless of how Geoff chooses to spend his time.

The hospital gave Geoff two options: (a) He can opt to retain his job and continue to do the work without using the automated software, or (b) he can opt to get paid without coming to work while the automated scribe software fulfills all the demands of the job. In either scenario, Geoff cannot work anywhere else due to a noncompete clause in his contract, and the hospital will not renew Geoff's contract when it expires.

The experimental conditions only differed in the final sentence of the vignette. In the "Stays on" condition, the vignette read, "Geoff decided on Option 1, to keep working at the hospital." The "Stops working" condition read, "Geoff decided on Option 2, to not keep working at the hospital."

Measures

An attention check, asking which option Geoff decided on, was presented on the page after participants read their assigned vignette. Approximately 95% of participants ($N = 763$) passed the attention check, and only their results were used in the analyses presented below, as specified in our preregistration. Consequently, 394 participants remained in the "Stays on" condition, and 369 remained in the "Stops working" condition (the results are substantively identical when those who failed the attention check are included in the analyses).

Following Studies 2a–2c, participants then completed the same 15 trait items and a similar set of 14 face-valid measures, including items measuring perceived effort, job difficulty, suffering, and deserved pay that were nearly identical to those used in Study 2a. Of note, items that captured perceptions of perceived meaning in life, job enjoyment, and loyalty read as follows: "How much meaning do you think Geoff has in his life?" (1 = *No meaning at all*, 4 = *Moderate amount of meaning*, 7 = *Lots of meaning*) "How much do you think Geoff enjoys his job?" (1 = *Does not enjoy it at all*, 4 = *Moderately enjoys it*, 7 = *Enjoys it a lot*), and "How loyal do you think Geoff is?" (1 = *Not loyal at all*, 4 = *Moderately loyal*, 7 = *Extremely loyal*). After finishing these measures, participants completed the PWE scale, Work Ethos items, and demographic measures before concluding the survey (please see our OSF page for the study materials and the online supplemental materials for the full set of descriptives and results).

Results

Following the previous study and our preregistration, the trait composites of competence ($\alpha = .86$), warmth ($\alpha = .83$), and morality ($\alpha = .91$) were used in the primary analyses. To conduct these analyses, participants' evaluations were entered into Welch's independent samples t tests (Delacre et al., 2017).

The character in the "Stays on" condition was seen as engaging in significantly more effort ($M = 6.0$, $SD = 1.1$) than the character

in the "Stops working" condition ($M = 5.0$, $SD = 1.3$), $t(733) = 11.73$, $p < .001$, $d = .85$, 95% CI [.70, 1.01]. Supporting our hypotheses, the character who chose to continue working even though his job could be done with automation was seen as significantly more moral ($M = 5.9$, $SD = 1.0$) than the character who stopped working ($M = 4.8$, $SD = 1.1$), $t(749) = 14.32$, $p < .001$, $d = 1.04$, 95% CI [.88, 1.20]. The character who stayed on was also seen as significantly warmer ($M = 4.9$, $SD = 1.0$) than the character who stopped working ($M = 4.6$, $SD = 1.0$), $t(758) = 4.28$, $p < .001$, $d = .31$, 95% CI [.17, .45], yet the target who stayed on was deemed significantly *less* competent ($M = 4.4$, $SD = 1.0$) than the target who stopped working ($M = 4.8$, $SD = .8$), $t(759) = -5.92$, $p < .001$, $d = -.43$, 95% CI [-.57, -.28].

Mediation Analysis

Using the jAMM package in jamovi, we explored whether the effect of our manipulation on moral character judgments was mediated by perceived effort. The condition variable was dummy coded (0 = "Stops working," 1 = "Stays on"). The indirect effect through effort was significant, $b = .53$, 95% CI [.42, .65], $p < .001$, indicating that evaluations of effort were a driver of participants' moral judgments of an individual engaging in economically unnecessary effort. This indirect effect remained significant when controlling for potential alternative mechanisms (perceived job difficulty, suffering, meaning in life, job enjoyment, and loyalty⁸), $b = .26$, 95% CI [.17, .35], $p < .001$. Full model statistics are presented in the online supplemental materials.

Moderation Analyses

Using the medmod package in jamovi, we tested whether participants' PWE or Work Ethos scores moderated the effect of perceived effort on moral evaluations. As in our prior studies, the interaction between effort and PWE scores, $b = .01$, 95% CI [-.09, .11], $p = .88$, and between effort and Work Ethos scores, $b = -.03$, 95% CI [-.09, .03], $p = .40$, were not significant, suggesting that individuals with varying work ethic beliefs moralized effort equivalently.

Discussion

Even when holding economic productivity constant, an individual conducting economically unnecessary labor was deemed more moral than an individual abstaining from such labor, and perceptions of effort partially drove this effect. As in our previous studies, this effect was not moderated by individual differences in PWE endorsement. These results further demonstrate the moral value placed on effort in the domain of work.

Study 4

We have posited that displays of effort, like other costly displays, signal a specific set of qualities and motivations that make one an attractive cooperation partner (Barclay, 2013; Gintis et al., 2001). Supporting this theoretical account, we have found that

⁸ Loyalty is typically thought of as a moral trait (e.g., Goodwin, 2015) and can be considered an outcome of effort moralization rather than a competing explanation. Nevertheless, controlling for loyalty in this analysis indicates that effort cues engendered a range of moral trait evaluations above and beyond perceptions of loyalty.

effort cues reliably engender moral character evaluations. Prior research has found that moral character judgments drive partner choice decision-making (Everett et al., 2016), and we hypothesized that effort-induced moral character judgments should also serve that function.

Indeed, if effort moralization is a pervasive social heuristic that subverts cooperative partner choice, then this process should occur even when potential partners are evaluated in domains of behavior where increased effort cannot directly increase economic productivity or prosocial outcomes, such as running for personal fitness. Although the domains differ considerably, someone who engages in more effort for their physical fitness may be signaling traits that indicate they would work similarly hard as a social or cooperation partner. In Study 4, we experimentally tested this prediction by examining whether the moralization of effort—in a context where effort provides no direct economic utility—helps explain cooperative partner choice.

Method

Procedure

The results of a pilot study and power analyses for indirect effects (Schoemann et al., 2017) indicated we would need at least 728 participants to have 90% power ($\alpha = .05$) to detect the hypothesized indirect effect. Ultimately, because we expected to exclude several participants (exclusion criteria were preregistered and are explicated below), 1000 U.S. adults (523 female, 12 other; age: $M = 31.2$, $SD = 10.9$) were recruited through Prolific Academic.

After consenting, each participant was presented with prompts about a high-effort and a low-effort runner in randomized order. Participants were presented with both target prompts before completing any dependent measures (target names were counterbalanced across participants). The prompts described two characters, each of whom runs five kilometers in 30 minutes twice a week; thus, both the duration and distance of the runs were constant across vignettes. The only difference between the two characters was described in the last sentence of each prompt, which read “For Justin (Mark), running a 5k is not very hard work (very hard work)—although the run is moderately difficult and he maintains a consistent pace, it takes minimal effort for him (takes a great deal of effort for him).” After reading both prompts, participants responded to dependent measures about each target.

Next, to test which target would be preferred as a cooperative partner, we followed Everett et al. (2016) in asking participants to select a partner for a trust game (TG). Participants first read a description of a TG and were given three comprehension check questions at the end of the description (these materials are available on our OSF page). Participants had three attempts to correctly answer these comprehension questions, after which all participants (including those who failed to correctly answer the questions) advanced to the next set of dependent measures.

On the subsequent page, participants completed a hypothetical partner choice measure, which read:

Imagine that you were playing this game as Person A. You have \$1.00 and Person B has \$1.00. Any money you send will be doubled and delivered to Person B. They will then decide how much money they would like to send back to you. If you had a choice and could select one of the other people from earlier in this study (Justin or Mark),

which one would you rather have in this game with you? Would you rather play with Justin or Mark?

The dichotomous response options read “Rather play with Justin (runs 5ks in 30 minutes with minimal effort)” and “Rather play with Mark (runs 5ks in 30 minutes with a great deal of effort),” with target names and descriptions matching the stimuli presented to each participant. We hypothesized that significantly more than 50% of participants would select the high-effort target to be their TG partner.

Measures

We used the same 15 trait items to measure morality, warmth, and competence, and a similar set of face-valid measures to those used in our previous studies were also deployed (e.g., effort, difficulty, suffering; see our OSF page for the full materials and the online supplemental materials for all descriptives and results). New items that captured perceptions of perceived health, self-confidence, importance of running, and commitment to running read as follows: “How healthy do you think Justin/Mark is?” (1 = *Not healthy at all*, 4 = *Moderately healthy*, 7 = *Extremely healthy*), “How self-confident do you think Justin/Mark is?” (1 = *Not self-confident at all*, 4 = *Moderately self-confident*, 7 = *Extremely self-confident*) “How committed is Justin/Mark to running?” (1 = *Not at all committed*, 4 = *Moderately committed*, 7 = *Extremely committed*), “How important is running to Justin/Mark?” (1 = *Not important at all*, 4 = *Moderately important*, 7 = *Extremely important*). After the person perception measures and the partner choice item, participants indicated on slider scales how much money they would send to each target (from \$0 to \$1) and how much money they would expect to receive back from each target (from \$0 to \$3) if they played the TG with each of them. Responses to these items were examined as part of secondary analyses relating to the partner choice task. Finally, participants completed the PWE scale and demographic questions before concluding the study.

Results

Following our preregistration, we excluded participants who completed the study in less than 3 minutes, rated the low-effort target as exerting equal or greater effort than the high-effort target, and/or failed the comprehension questions. This resulted in a final sample of 689 participants (the results are substantively similar with the full sample included in the analyses). Following our prior studies, indices of competence (Cronbach’s $\alpha_{\text{high-effort}} = .84$, $\alpha_{\text{low-effort}} = .81$), warmth ($\alpha_{\text{high-effort}} = .91$, $\alpha_{\text{low-effort}} = .91$), and morality ($\alpha_{\text{high-effort}} = .83$, $\alpha_{\text{low-effort}} = .81$) were used in the primary analyses. For the person perception measures, participants’ evaluations of both targets were entered into paired samples *t* tests.⁹ Participants in the final sample rated the high-effort target as putting significantly more effort into running ($M = 6.5$, $SD = .7$)

⁹ We preregistered one-tailed tests for the effort and morality variables because we had directional predictions. We present the results of two-tailed tests in the main text to maintain interpretative consistency across analyses and studies. The interpretation of the results does not change when using one- or two-tailed tests.

than the low-effort target ($M = 3.6$, $SD = 1.2$), $t(688) = 57.99$, $p < .001$, $d = 2.21$, 95% CI [2.07, 2.35].

As predicted, and conceptually replicating our previous studies, participants rated the high-effort runner as significantly more moral ($M = 5.0$, $SD = 1.0$) than the low-effort runner ($M = 4.6$, $SD = 1.0$), $t(688) = 13.41$, $p < .001$, $d = .51$, 95% CI [.43, .59]. In contrast, the high-effort target was rated as significantly *less* competent ($M = 4.5$, $SD = 1.0$) than the low-effort runner ($M = 4.7$, $SD = 1.0$), $t(688) = -7.19$, $p < .001$, $d = -.27$, 95% CI [-.35, -.20]. There were no significant differences between the high-effort ($M = 4.0$, $SD = 1.1$) and low-effort ($M = 4.1$, $SD = 1.1$) runners on perceived warmth, $t(688) = -.90$, $p = .37$, $d = -.03$, 95% CI [-.11, .04]. Notably, while the high-effort target was seen as less healthy ($M = 4.8$, $SD = 1.1$) than the low-effort target ($M = 5.9$, $SD = .9$), $t(688) = -23.80$, $p < .001$, $d = -.91$, 95% CI [-1.00, -.82], the high-effort target was seen as being more committed to running ($M = 5.9$, $SD = 1.0$) than the low-effort runner ($M = 5.2$, $SD = 1.2$), $t(688) = 13.60$, $p < .001$, $d = .52$, 95% CI [.44, .60].

The partner choice measure was analyzed using a binomial proportion test, with the test value set at .50. According with our partner choice account of effort moralization, 569 of 689 participants (83%) selected the high-effort target as their hypothetical TG partner, $p < .001$, 95% CI [.80, .85]. In secondary analyses using the slider scale items, we also found that participants reported they would send significantly more money to the high-effort target ($M = \$7$, $SD = \$3$) than the low-effort target ($M = \$6$, $SD = \$3$), $t(688) = 14.26$, $p < .001$, $d = .54$, 95% CI [.46, .62], and expected to receive more money from the high-effort target ($M = \$1.1$, $SD = \$6$) than from the low-effort target ($M = \$8$, $SD = \$2$), $t(687) = 15.94$, $p < .001$, $d = .57$, 95% CI [.49, .65]. In sum, participants indicated a strong preference for the high-effort target as their TG partner, despite the targets expending effort in an arena completely unrelated to the TG task.

Mediation Analysis

Similar to our previous studies, we tested whether moral character perceptions helped explain participants' partner choice decisions. Difference scores were calculated for the effort and morality variables, and a dummy-coded partner choice variable (0 = low-effort, 1 = high-effort) was entered as the outcome measure. The indirect effect through moral character judgment was significant, $b = .01$, 95% CI [.00, .02], $p < .001$, indicating that differences in moral character judgment were a driver of participants' partner choice decisions. This indirect effect through moral character judgment remained significant when controlling differences between the runners in perceived self-confidence, health, run difficulty, and suffering,¹⁰ $b = .01$, 95% CI [.00, .02], $p < .001$, illustrated in Figure 3. An unplanned sensitivity analysis using binomial logistic regression to estimate the "b" path from moral character to partner choice found convergent results (please see the online supplemental materials for full model statistics).

We also conducted an exploratory serial mediation analysis with perceived effort as the predictor, perceived goal commitment (to running) as the first mediator, moral character as the second mediator, and partner choice as the outcome. We found support for this serial indirect effect in models with and without the control variables (model statistics are presented in the online supplemental

materials). This finding conceptually replicates and extends motivation attribution accounts of effort moralization, which have found that effort amplifies moral judgments by increasing one's perceived commitment to explicitly prosocial or antisocial actions (e.g., returning or stealing a wallet; Bigman & Tamir, 2016). Here, we found that perceived commitment to a personal task—running—increased perceptions of one's moral character, which subsequently influenced partner choice judgments.

Moderation Analyses

Linear and binomial logistic regression analyses were conducted to assess whether PWE scores moderated the effects of effort on moral character judgments, effort on partner choice, and moral character on partner choice. Once again, none of the key interaction terms between the predictor variables and PWE scores were significant ($ps = .13-.69$), indicating that individual differences in PWE beliefs did not explain meaningful variation in the judgments underlying participants' partner choice decisions (see the online supplemental materials for statistics).

Discussion

Although the two targets engaged in otherwise identical behavior, a more effortful runner was perceived to be more moral and was preferred for an unrelated, trust-based task. These findings held when controlling for other relevant evaluations and did not vary by individuals' PWE endorsement. The more effortful target was preferred because their greater efforts signaled moral traits that indicate general cooperative intentions. This suggests that effort moralization may be a pervasive social heuristic that can help explain interpersonal judgments and decision-making across domains of behavior.

Study 5

Partner choice decisions often occur informally and can involve degrees of commitment to multiple partners. For instance, online fundraising platforms provide people the opportunity to assist countless others by donating to their charitable causes. How do people decide which advocates are worthy of their monetary support? Research on the "martyrdom effect" (Olivola & Shafir, 2013) finds that individuals donate more when they have exerted economically unnecessary effort for a cause, such as running a race for charity. Extending this work into the interpersonal domain and paralleling the partner choice paradigm of the previous study, we tested whether people donate more to a target who runs a longer race due to perceiving this effort as a signal of moral character.

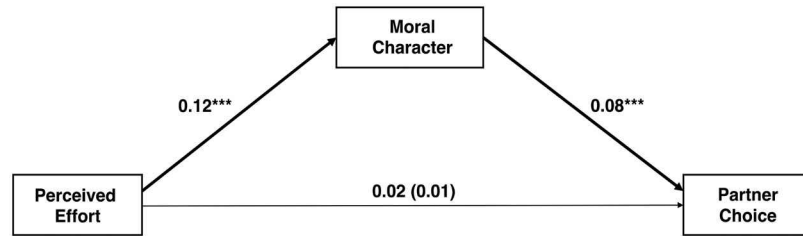
Method

Procedure

We aimed to recruit 400 participants to have at least 80% power to detect a small ($d = .13$) within-subjects effect (one-tailed, $\alpha =$

¹⁰ Analyses controlling for perceived self-confidence, health, and difficulty were preregistered; we included perceived suffering as a control variable post-hoc to maintain analytic consistency throughout our studies.

Figure 3
 Mediation Model for Study 4 Showing the Effect of Differences in Perceived Effort on Partner Choice, as Mediated by Differences in Moral Character



Note. The presented path estimates control for differences in perceived self-confidence, health, run difficulty, and suffering. Unstandardized coefficients are displayed. On the center path, the coefficient outside the parentheses is the total effect, and the coefficient inside the parentheses is the direct effect. Asterisks indicate significant paths (***) $p < .001$.

.05).¹¹ Ultimately, 405 U.S. adults (211 female, one other; age: $M = 37.1$, $SD = 11.7$) were recruited through MTurk.

After consenting, participants were asked to read about and allocate \$99 between two fundraisers. To involve real rather than hypothetical stakes for the participants' decisions, we explicitly stated that we would randomly select the allocations made by one participant and donate that money to the two fundraisers accordingly.

Participants were presented with the two fundraising pages, ostensibly organized by separate individuals on GoFundMe.com. One page described Mark's fundraiser for the Against Malaria Foundation, and the other page described Justin's fundraiser for the Deworm the World Initiative. Participants were assigned to one of two experimental conditions. In one condition, Mark ran a marathon (approximately 26-mile) race for his Against Malaria Foundation fundraiser while Justin ran a 5k (approximately 3-mile) race for his Deworm the World Initiative fundraiser; in the other condition, Mark ran a 5k while Justin ran a marathon. The order in which each fundraiser was presented to participants was counterbalanced across conditions, and participants were required to stay on each page for at least 20 seconds before proceeding.

After being presented with both stimuli, participants allocated the \$99 between the two fundraisers. Next, participants were again presented with the GoFundMe page images for Mark or Justin and asked to fill out a series of measures about each character. The images and corresponding items for each character were presented on separate pages and in counterbalanced order.

Measures

As in our prior studies, 15 trait items measured competence, warmth, and moral evaluations, along with a similar set of face-valid measures (the full materials are available on our OSF page). Notably, new items for perceived athleticism, charity importance, and charity popularity read as follows: "How athletic do you think Mark/Justin is?" (1 = *Not athletic at all*, 4 = *Moderately athletic*, 7 = *Extremely athletic*), "How important do you think the charity is that Mark/Justin is raising money for?" (1 = *Not important at all*, 4 = *Moderately important*, 7 = *Extremely important*), and "How popular do you think the charity is that Mark/Justin is raising money for?" (1 = *Not popular at all*, 4 = *Moderately popular*,

7 = *Extremely popular*). Lastly, participants completed the PWE Scale, Work Ethos scale, and demographic questions.

Results

Following our preregistration, we excluded participants who rated the target running the 5k as exerting more effort than the target running the marathon, resulting in a final sample of 360 participants (the results are substantively similar with all participants included in the analyses). As in the previous studies, indices of competence ($\alpha_{26\text{miles}} = .90$, $\alpha_{3\text{miles}} = .91$), warmth ($\alpha_{26\text{miles}} = .84$, $\alpha_{3\text{miles}} = .83$), and morality ($\alpha_{26\text{miles}} = .92$, $\alpha_{3\text{miles}} = .92$) were used in the subsequent analyses. Participants' evaluations of both targets were entered into paired samples t-tests for these analyses. Participants rated the target who ran a marathon as putting significantly more effort into fundraising ($M = 6.2$, $SD = 1.15$) than the target who ran a 5k ($M = 5.1$, $SD = 1.3$), $t(359) = 17.32$, $p < .001$, $d = .91$, 95% CI [.79, 1.04].

As predicted, participants donated more to the marathon fundraiser ($M = \$55.4$, $SD = \$17.5$) than to the 5k fundraiser ($M = \$43.6$, $SD = \$17.5$), $t(359) = 6.43$, $p < .001$, $d = .34$, 95% CI [.23, .45]. Also as predicted, the marathon runner was seen as significantly more moral ($M = 5.9$, $SD = 1.0$) than the 5k runner ($M = 5.7$, $SD = 1.1$), $t(359) = 6.89$, $p < .001$, $d = .36$, 95% CI [.26, .47]. Unlike our previous studies, here the higher-effort, marathon runner was considered more competent ($M = 5.5$, $SD = 1.0$) than the lower-effort 5k runner ($M = 5.3$, $SD = 1.1$), $t(359) = 4.20$, $p < .001$, $d = .22$, 95% CI [.12, .33]. There were no significant differences between the marathon runner ($M = 5.1$, $SD = 1.0$) and the 5k runner ($M = 5.1$, $SD = 1.0$) on perceived warmth, $t(359) = .39$, $p = .35$, $d = .02$, 95% CI [-.08, .12].

Mediation Analysis

We tested whether moral character perceptions explained the relationship between perceived effort and charitable giving using the jAMM package in jamovi. Difference scores were calculated

¹¹ As in Study 4, we present the results of two-tailed tests despite preregistering one-tailed analyses for the effort, morality, and donation variables. The interpretation of the results does not differ when using one- or two-tailed tests.

for the effort, morality, and donation variables and entered into the model. Importantly, the indirect effect through moral character judgment was significant, $b = 1.70$, 95% CI [.33, 2.84], $p = .008$. Differences in moral character evaluations partially explained the effect of differences in effort perceptions on participants' donation allocations, even when controlling for differences in perceived athleticism, meaning in life, enjoyment of running, suffering, charity importance, and charity popularity, $b = 1.17$, 95% CI [.06, 2.17], $p = .03$, illustrated in Figure 4 (additional model statistics are presented in the online supplemental materials).

Moderation Analyses

Analyses were also conducted to assess whether PWE or Work Ethos scores moderated the effects of effort on moral character judgments, effort on donations, and moral character on donations. Overall, none of the key interaction terms between the predictor variables and the work ethic individual difference variables were significant ($ps = .26-.95$), indicating no significant moderation effects (see the online supplemental materials for statistics). Individual differences in PWE and Work Ethos beliefs did not explain meaningful variation in the moral and monetary value participants ascribed to the runners' efforts.

Discussion

We found that fundraisers who engaged in more economically unnecessary effort received more charitable donations, in part because they were perceived as having greater moral character. These findings extend work on the martyrdom effect to the interpersonal realm and document a novel mediator to explain the effect (Olivola & Shafir, 2013). Specifically, participants contributed more to targets who invested more effort in part because those targets were perceived as more moral. Moreover, this finding was robust to alternative explanations, such as perceptions of suffering (Schaumburg & Mullen, 2017). In conjunction with our previous findings, these results suggest that effort moralization may underlie a broader range of interpersonal judgments and behaviors than previously theorized.¹²

Study 6

We have shown that perceived effort affects moral judgment in conditions where effort can directly influence economic outcomes (Studies 1–2c) and in situations where effort cannot directly alter economic outcomes (Studies 3–5). However, important limitations in some of our previous studies remain. First, in Studies 2a–2c, we did not control for perceptions of work value. It is possible that the moral judgment effects observed in those studies could be explained by perceived differences in the value produced by each target. Second, the composite measure of moral character we used throughout our studies combined items that can be separated into two separate types of moral traits: core goodness traits (or unconditionally moral traits, e.g., “Honesty”), and value commitment traits (or conditionally moral traits, e.g., “Dedicated”; Piazza et al., 2014). Consequently, it is not clear whether effort cues are broad signals of moral character, engendering both core goodness and value commitment judgments, or if effort cues are more precise signals of either core goodness or value commitment.

Prior research has demonstrated that effort cues signal one's commitment to moral actions, amplifying the perceived goodness of prosocial behavior and the badness of antisocial behavior (Bigman & Tamir, 2016). This is consistent with effort being a narrow signal of value commitment traits, characteristics that enable one to follow through on their moral or immoral intentions. However, we hypothesized that effort may also serve as a signal of one's cooperative intent. Thus, rather than being a narrow cue of value commitment specifically, we argue that people heuristically evaluate effort cues as broad signals of one's capacity to enact one's intentions (i.e., value commitment traits) and one's cooperative intent (i.e., core goodness traits), a process that facilitates efficient partner choice decision-making. In Study 6, we adapted the materials from Study 2a and incorporated the partner choice task from Study 4 to test the broad and narrow accounts of effort moralization.

Method

Procedure

The results of a pilot study and power analyses for indirect effects (Schoemann et al., 2017) indicated we would need at least 800 participants to have 80% power ($\alpha = .05$) to detect the hypothesized indirect effects. Since we expected to exclude several participants (exclusion criteria were preregistered and are detailed below), we aimed to recruit 1000 participants. Ultimately, 1,002 U.S. adults (573 female, 19 other; age: $M = 34.9$, $SD = 12.2$) were recruited through Prolific Academic.

After consenting and completing English comprehension items, participants were informed that they were to read about Justin and Mark, two workers who have the same job at the same widget factory. On the next page, participants read character descriptions like those presented in Studies 2a–c, one of a low-effort target (Justin) and one of a high-effort target (Mark). In this study, the descriptions were presented together in a joint-evaluation format (Hsee et al., 1999), and the vignette was edited to better control for the perceived value and quality of each worker's output. The vignette read as follows:

Justin and Mark work in the same factory and make the same widgets. Both Justin and Mark are able to produce approximately six widgets per hour, one widget around every 10 minutes.

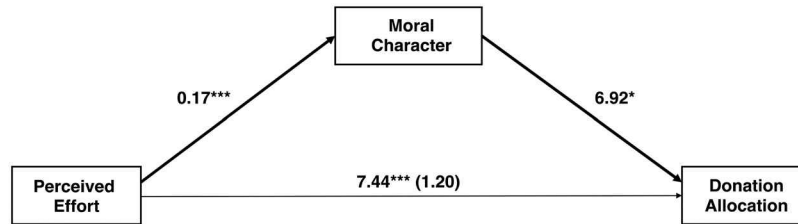
The market value for these widgets is \$4.00. Quality control inspections indicate that 96% of Justin's widgets and 96% of Mark's widgets work flawlessly, which means they can be sold. Thus, in an average hour, both Justin and Mark are able to produce \$23.04 worth of high-quality widgets.

For Justin, making widgets requires minimal effort—although he works as quickly as possible, it is easy work.

For Mark, making widgets requires a lot of effort—although he works as quickly as possible, it is hard work.

¹² Please see the online supplemental materials for a study that investigated naturalistic donations to fundraisers using data from GoFundMe.com. Although these data did not include assessments of morality, we found that fundraiser distance, a proxy for effort, was a significant predictor of the amount of donations solicited by the fundraiser.

Figure 4
 Mediation Model for Study 5 Showing the Effect of Differences in Perceived Effort on Differences in Donation Allocation, as Mediated by Differences in Moral Character



Note. The presented path estimates control for differences in perceived athleticism, meaning in life, enjoyment of running, suffering, charity importance, and charity popularity. Unstandardized coefficients are displayed. On the center path, the coefficient outside the parentheses is the total effect, and the coefficient inside the parentheses is the direct effect. Asterisks indicate significant paths (* $p < .05$. *** $p < .001$).

After reading the vignette, participants completed separate sets of dependent measures for each target in randomized order. Participants were then presented with similar trust game (TG) instructions and dependent measures as those used in Study 4 to assess cooperative partner choice. We hypothesized that significantly more than 50% of participants would select the high-effort target to be their TG partner.

Measures

Drawing from Piazza and colleagues (2014), we adjusted the trait items used in this study to create composite measures that distinguished between core goodness traits (i.e., unconditionally moral) and value commitment traits (i.e., conditionally moral) for each target. Core goodness was measured with six items: “Moral,” “Trustworthy,” “Honest,” “Respectful,” “Just,” and “Cooperative.” Value commitment was measured using seven items: “Dedicated,” “Responsible,” “Principled,” “Determined,” “Motivated,” “Disciplined,” and “Self-controlled.” Single items were used to measure competence (“Competent”) and warmth (“Warm”), respectively. All trait items were measured using 7-point scales (1 = *Does not describe Justin/Mark well*, 4 = *Describes Justin/Mark moderately well*, 7 = *Describes Justin/Mark extremely well*).

As in our previous studies, face-valid dependent measures were used to measure the perceived effort, quality of work, difficulty of work, and deserved pay of each target. In this study, we also included an item to assess the perceived value of the work generated by each target, “How valuable do you think Justin’s/Mark’s work is?” (1 = *Not at all valuable*, 4 = *Moderately valuable*, 7 = *Extremely valuable*).

Finally, the TG instructions and accompanying dependent measures were nearly identical to those used in Study 4. Participants were asked to choose the target with whom they would prefer to play the TG (Justin or Mark). Next, they used slider scales to indicate how much money they would send to each target (from \$0 to \$1) and how much money they would expect to receive back from each target (from \$0 to \$3) if they played the TG together. Participants then completed the PWE scale and demographic questions before concluding the study.

Results

Following our preregistration, we excluded participants who completed the study in less than 3 minutes, rated the low-effort target as exerting equal or greater effort than the high-effort target, and/or failed the TG comprehension questions. This resulted in a final sample of 869 participants (the results are substantively identical when including all participants). Indices of core goodness (Cronbach’s $\alpha_{\text{high-effort}} = .94$, $\alpha_{\text{low-effort}} = .95$), and value commitment ($\alpha_{\text{high-effort}} = .93$, $\alpha_{\text{low-effort}} = .94$) were used as the main dependent measures. Although the core goodness and value commitment traits were significantly correlated ($r_{\text{high-effort}} = .72$, $r_{\text{low-effort}} = .80$), confirmatory factor analyses indicated that models with core goodness and value commitment as separate factors had significantly better model fit than single-factor models of the moral character traits (see the online supplemental materials for the full set of model statistics). Further following our preregistration, we entered participants’ evaluations of both targets into paired samples t-tests for the primary analyses.

Participants rated the high-effort target as putting significantly more effort into his work ($M = 6.6$, $SD = .6$) than the low-effort target ($M = 3.0$, $SD = 1.1$), $t(868) = 89.26$, $p < .001$, $d = 3.03$, 95% CI [2.87, 3.18]. Despite indicating to participants that the two workers had the same job and produced widgets of equal quality and value, participants rated the high-effort target as having a more difficult job ($M = 5.0$, $SD = 1.2$) than the low-effort target ($M = 2.8$, $SD = 1.2$), $t(868) = 32.92$, $p < .001$, $d = 1.12$, 95% CI [1.03, 1.20], producing higher quality work ($M = 6.2$, $SD = .9$) than the low-effort target ($M = 6.0$, $SD = 1.1$), $t(868) = 5.87$, $p < .001$, $d = .20$, 95% CI [.13, .27], and producing more valuable work ($M = 5.6$, $SD = 1.2$) than the low-effort target ($M = 5.4$, $SD = 1.3$), $t(868) = 5.76$, $p < .001$, $d = .20$, 95% CI [.13, .26]. These results conceptually replicate and highlight the power of the effort heuristic (Kruger et al., 2004): Even when being told that each targets’ work was of identical economic value, participants perceived the high-effort target’s work as being more valuable.

In testing our main hypotheses, we found that, as predicted, the high-effort target was rated significantly higher on core goodness traits ($M = 5.1$, $SD = 1.1$) than the low-effort target ($M = 4.8$, $SD = 1.1$), $t(868) = 12.49$, $p < .001$, $d = .42$, 95% CI [.35, .49]. The high-effort target was also rated significantly higher on the value

commitment traits ($M = 5.8$, $SD = 1.0$) than the low-effort target ($M = 4.9$, $SD = 1.1$), $t(868) = 22.31$, $p < .001$, $d = .76$, 95% CI [.68, .83]. Thus, effort cues yielded broad moralization of both core goodness and value commitment traits rather than value commitment traits alone.

At the request of a reviewer, we conducted regression analyses to test whether differences in perceived effort predicted differences in perceived moral goodness when controlling for differences in perceived value commitment. This provided an initial assessment of the degree to which effort cues served as direct or indirect signals of cooperative intent. In a model with only effort predicting core goodness, there was a significant effect, $b = .07$, 95% CI [.03, .12], $p < .001$. This indicated that, the more effort participants perceived the high-effort target to exert compared with the low-effort target, the more they thought the high-effort target exuded core goodness traits relative to the low-effort target. However, when including value commitment in the model, the effect of effort was reduced to nonsignificance, $b = .01$, 95% CI [−.03, .04], $p = .65$, while the effect of value commitment was highly significant, $b = .45$, 95% CI [.42, .49], $p < .001$. Thus, when accounting for differences in value commitment between the two targets, differences in effort were not predictive of differences in core goodness, suggesting that effort indirectly signals cooperative intent.

The high-effort target was also seen as significantly warmer ($M = 4.7$, $SD = 1.2$) than the low-effort target ($M = 4.5$, $SD = 1.2$), $t(868) = 5.89$, $p < .001$, $d = .20$, 95% CI [.13, .27], yet the high-effort target was seen as significantly *less* competent ($M = 5.3$, $SD = 1.4$) than the low-effort target ($M = 5.9$, $SD = 1.2$), $t(868) = -12.44$, $p < .001$, $d = -.42$, 95% CI [−.49, −.35]. As in Studies 2a–2c, participants rated the high-effort target as deserving of a higher wage ($M = \$13.8$, $SD = \$2.0$) than the low-effort target ($M = \$13.7$, $SD = \$2.1$), although this effect was smaller than those in our previous studies, $t(863) = 3.55$, $p < .001$, $d = .12$, 95% CI [.05, .19].

The TG partner choice measure was analyzed using a binomial proportion test, with the test value set at .50. In line with our predictions, 630 out of 869 participants (73%) selected the high-effort target as their hypothetical TG partner, $p < .001$, 95% CI [.69, .75]. Participants also reported that they would send significantly more money to the high-effort target ($M = \$7$, $SD = \$3$) than to the low-effort target ($M = \$6$, $SD = \$3$), $t(867) = 14.10$, $p < .001$, $d = .48$, 95% CI [.41, .55], and expected to receive more money from the high-effort target ($M = \$1.1$, $SD = $.6$) than from the low-effort target ($M = $.8$, $SD = $.6$), $t(866) = 13.01$, $p < .001$, $d = .44$, 95% CI [.37, .51]. As in Study 4, participants indicated that the high-effort target had more cooperative intentions—as indicated by both the core goodness measure and the measures of expected TG behavior—which was further evidenced by their preference for the high-effort target in the cooperation task.

Mediation Analysis

Following our preregistration, we first used the jAMM package in jamovi to estimate whether perceptions of core goodness and perceptions of value commitment would mediate the relation between perceived effort and participants' decisions in the partner choice task. Difference scores were calculated for the perceived effort, core goodness, and value commitment variables, and a

dummy-coded partner choice variable (0 = low-effort, 1 = high-effort) was entered as the outcome measure. Core goodness and value commitment were entered as separate mediators in this model.

The indirect effect through core goodness was significant, $b = .004$, 95% CI [.00, .01], $p = .043$, as was the indirect effect through value commitment, $b = .011$, 95% CI [.00, .02], $p = .002$. In other words, differences in perceived effort engendered higher core goodness and value commitment judgments of the high-effort target, and these moral judgments increased the likelihood of participants selecting the high-effort target in the partner choice task. We then constructed a model that controlled for the alternative mechanisms of differences in job difficulty, work quality, and work value. In this second model, the indirect path through core goodness became marginally significant, $b = .004$, 95% CI [−.00, .01], $p = .056$, whereas the indirect effect through value commitment remained significant, $b = .01$, 95% CI [.00, .02], $p = .002$. Although the indirect effects through work quality and work value were not significant, the indirect effect through job difficulty was negative, $b = -.006$, 95% CI [−.01, −.00], $p = .04$, such that perceiving the high-effort target as having a more difficult job decreased the likelihood of selecting the high-effort target for the partner choice task. The results of these preregistered mediation analyses generally supported our partner choice account of effort moralization, and we did not find evidence in support of proposed alternative mechanisms for these effects.

We additionally ran an unplanned sensitivity analysis using binomial logistic regression and found that the linear models may have underestimated the “b” path of the core goodness indirect effect. In the linear model including covariates, core goodness was only marginally predictive of partner choice, $b = .05$, 95% CI [−.00, .10], $p = .051$; in contrast, when constructing a binomial logistic model with the same variables—likely the more appropriate analysis given the dichotomous partner choice dependent variable—core goodness was a stronger and highly significant predictor of partner choice, $OR = 1.77$, 95% CI [1.25, 2.51], $p = .001$. Indeed, the logistic model accounted for more variance in partner choice decisions ($R^2 = .13$) than the linear model ($R^2 = .05$). Thus, the true indirect effect through core goodness may be larger and more robust to covariates than what was estimated in our preregistered mediation analyses (please see the online supplemental materials for full model statistics and a discussion of the differences between the linear and logistic model estimates).

Finally, to follow up on the reviewer-requested regression analyses more directly, we also conducted a serial mediation analysis to assess an alternative causal pathway from effort to core goodness and partner choice. In this model, perceived effort was entered as the predictor, value commitment was the first mediator, core goodness was the second mediator, and partner choice was the outcome variable. The key difference in this model, compared with the parallel mediations described above, was the inclusion of a path from value commitment to core goodness. This allowed us to more strictly test whether value commitment mediated the relationship between effort and core goodness. It also enabled us to explore whether including the serial path between value commitment and core goodness increased the explanatory value of the mediation model. We constructed linear and logistic estimates of the paths to the partner choice outcome variable, given the variation in results observed due to model specifications in the

parallel mediation analyses. We ran models with and without including job difficulty, work quality, and work value as alternative mechanisms, although the inclusion of these variables added little explanatory value to these models ($\Delta R^2_{linear} = .01$, $\Delta R^2_{logistic} = .03$; see the online supplemental materials for full statistics across model specifications). These analyses were conducted in Mplus.

We focus on the model that controlled for alternative mechanisms and included logistic estimates of partner choice because this model explained the most variance in participants' partner choice decisions ($R^2 = .19$). This model is illustrated in Figure 5. Notably, there was not a significant indirect path through core goodness in predicting partner choice; instead, there was a serial indirect path from effort to value commitment, value commitment to core goodness, and then core goodness to partner choice. The high-effort target was seen as more value committed, which increased perceptions of the high-effort target's core goodness, which in turn increased the likelihood of participants choosing the high-effort target as a partner for the cooperative task. In other words, this model suggested that differences in perceived effort did not directly yield differences in perceived core goodness. There was also a significant path from value commitment to partner choice, indicating that value commitment had a direct influence on participants' partner choice decision in addition to the indirect effect it had through core goodness.

Overall, whereas the preregistered parallel mediation models suggested that effort cues had a direct effect on core goodness traits, the serial mediation analyses—which accounted for more variance in partner choice outcomes—suggested that effort had only an indirect influence on core goodness through value commitment traits. These results suggested a somewhat different causal model of the effort moralization and partner choice effects than we hypothesized. Nevertheless, the findings still provided evidence in support of our claim that effort is a broad signal of moral character that influences partner choice decisions. Targets perceived as exerting greater effort were judged as possessing greater core goodness and value commitment traits, and the full suite of analyses showed that both types of moral character judgments increased the likelihood of participants selecting the high-effort target as their hypothetical cooperation partner.

Moderation Analyses

Following our preregistration, we conducted analyses to assess whether PWE scores moderated the effects of effort on core goodness, effort on value commitment, effort on partner choice, core goodness on partner choice, and value commitment on partner choice. Unlike our previous studies, here we found clear evidence of PWE beliefs moderating effort moralization effects: the interaction between effort difference scores and the PWE composite was significant in predicting both core goodness, $b = -.11$, 95% CI $[-.21, -.02]$, $p = .024$, and value commitment, $b = -.13$, 95% CI $[-.27, -.01]$, $p = .048$. In both models, those with higher PWE scores moralized effort *less* than those with lower PWE scores.

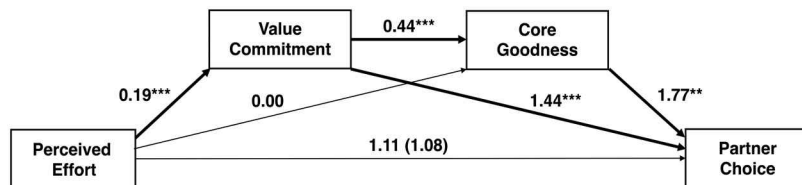
Moderation results were more mixed in the models with partner choice as the dependent variable. The interaction term between effort and PWE was not significant in predicting partner choice, $b = -.01$, 95% CI $[-.06, .04]$, $p = .71$, and neither was the interaction between core goodness and PWE in predicting partner choice, $b = .06$, 95% CI $[-.02, .12]$, $p = .14$. However, there was a significant interaction between value commitment and PWE in predicting partner choice, $b = .09$, 95% CI $[.04, .14]$, $p < .001$, such that differences in perceived value commitment between the two targets were more predictive of partner choice preferences for participants with higher PWE scores. Although all participants who judged the high-effort target as possessing greater value commitment were more likely to select that target as their cooperation partner, those value commitment judgments had a stronger influence on the partner choice decisions of participants with higher PWE scores. Sensitivity analyses using binomial logistic regression found convergent results for the three analyses with partner choice as the dependent variable.

Thus, the effects of PWE beliefs in the present results were complex: although effort cues were weaker signals of both perceived core goodness and value commitment for high-PWE participants, differences in perceived value commitment had a greater influence on the partner choice decisions of those same high-PWE individuals.

Discussion

When explicitly and statistically controlling for the value of work done by two individuals, we found that perceptions of effort

Figure 5
Mediation Model for Study 6 Showing the Effect of Differences in Perceived Effort on Cooperative Partner Choice Decisions, as Serially Mediated by Differences in Value Commitment and Core Goodness Traits



Note. The presented path estimates control for differences in perceived job difficulty, work quality, and work value. Unstandardized coefficients are displayed on paths predicting value commitment and core goodness. Odds ratios are displayed on paths predicting partner choice. On the center path, the coefficient outside parentheses is the odds ratio of the total effect, and the value inside parentheses is the odds ratio of the direct effect. Asterisks indicate significant paths (** $p < .005$. *** $p < .001$).

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

still influenced perceptions of these targets' moral character. A high-effort worker was seen as having stronger moral character, both in terms of possessing greater core goodness traits and value commitment traits, than an equally productive low-effort worker doing the same job. Counter to our hypotheses, participants' partner choice decisions were best predicted by a model in which effort cues did not directly influence perceptions of core goodness. Rather, effort cues directly influenced judgments of the targets' value commitment, which then directly influenced partner choice decisions and indirectly influenced those selections through core goodness evaluations. Although this differed from the causal pathway we proposed, these results still support our claim that effort is a broad signal of moral character. Effort cues induced value commitment and core goodness judgments that influenced participants' partner choices, resulting in nearly three-quarters of participants selected the high-effort target as their cooperation partner.

In sum, we found that effort cues serve as both direct and indirect signals of moral character, even when one's efforts do not yield any direct increases in economic value. These results lend further support to our partner choice account of intuitive effort moralization.

General Discussion

Is effort deemed socially valuable, even in situations where one's efforts have no direct economic utility? Eight studies using multiple methodologies and cross-cultural samples indicate that it is. We provided evidence of effort moralization—displays of effort increased the moral qualities ascribed to individuals (we did not, we should note, provide evidence of the specific process by which effort cues shift from having a nonmoral to moral status, a more limited definition of moralization; Rhee et al., 2019). Moreover, the moralization of effort guided participants' allocations of monetary resources and selections of cooperation partners. These data support our argument that effort moralization is a "deeply rational" social heuristic for navigating cooperation markets (Barclay, 2013; Kenrick et al., 2009). Even in circumstances where effort was economically unnecessary, people believed such efforts reflected others' inner virtues.

Our investigation advances previous research on effort evaluations in important ways. First, we extend prior research on evaluations of labor (Amos et al., 2019; Tierney et al., 2020) by showing that more effortful workers are morally valued even when they do not clearly produce concrete, economic benefits. Second, we broaden research on the martyrdom effect (Olivola & Shafir, 2013) by conceptually replicating it in paradigms focused on interpersonal judgments. We demonstrate that moral character judgments help explain why people donate more to causes that others have invested with economically unnecessary effort. Furthermore, we provide the first discriminative evidence that effort cues affect moral evaluations specifically, rather than positive character ascriptions generally. Although prior research has documented moral judgments of effort (e.g., Bigman & Tamir, 2016; Uhlmann & Sanchez-Burks, 2014), these studies could not delineate moral judgments from broader halo effects because only one domain of person perception, morality, was measured. Across our seven pre-registered experiments, manipulating effort produced consistent differences in assessments of morality but not assessments of warmth or competence. Our final experiment also demonstrated

that effort cues affect judgments of both core goodness and value commitment traits (Piazza et al., 2014), suggesting that effort moralization influences perceptions of cooperative intent as well as one's capacity to enact their intentions. In sum, consistent with research that places moral character judgments at the center of person perception (Goodwin, 2015; Uhlmann et al., 2015) and cooperative partner choice selection (Barclay, 2013; Everett et al., 2016; Gintis et al., 2001), we found that effort cues induced broad moral trait inferences that drove participants' subsequent social and economic decision-making.

Unpacking Explanations of Effort Moralization

In seeking to address methodological limitations of prior work, we revealed theoretical shortcomings of a cultural explanation for effort moralization. If PWE beliefs caused people to moralize effort (Amos et al., 2019; Uhlmann & Sanchez-Burks, 2014), then those who do not endorse those cultural beliefs should moralize effort less than those who do. Yet individual differences in work ethic beliefs rarely moderated the observed effects, and the few moderation effects that emerged (e.g., Study 6) indicated that those with lower PWE endorsement moralized effort *more* than those with stronger PWE beliefs. This evidence suggests a limited role of PWE in explaining effort moralization. Moreover, we conducted a preliminary assessment of the universality of effort moralization using the triangulation approach (Norenzayan & Heine, 2005), replicating the results from a U.S. sample (Study 2a) in South Korea and France (Studies 2b and 2c). These findings, alongside recent cross-cultural evidence from the United Kingdom, Australia, and India (Tierney et al., 2020), provide support for the notion that effort moralization may be a more cross-culturally prevalent heuristic than previously theorized.

In conjunction with specific evidence of effort moralization among members of a modern hunter-gatherer society (Smith & Apicella, 2020) and general evidence of effort valuation in humans and even nonhuman animals (e.g., Clement et al., 2000; Inzlicht et al., 2018; Lydall et al., 2010), these cross-cultural findings further suggest that effort moralization may rest on more fundamental, potentially evolutionary, origins. Humans evolved in collaborative, group-living environments where paying attention to displays of costly signaling may have been an efficient and adaptive way to assess the dedication and cooperative intent of others (Gintis et al., 2001). Just as people will engage in unnecessary prosocial behavior to differentiate themselves as a superior cooperative partner (Barclay, 2013), displays of effort, including economically unnecessary effort, may serve a similar function. While effort devoted toward explicitly prosocial ends has been shown to inform moral character judgments (e.g., returning a lost wallet; Bigman & Tamir, 2016), our findings suggest that even committing oneself to predominantly personal, self-focused endeavors (e.g., running for fitness) can engender moral judgments that, in turn, guide cooperative partner choice decision-making.

The results of Study 6 suggested that perceptions of both cooperative capacities (e.g., value commitment traits like dedication) and cooperative intent (e.g., core goodness traits like trustworthiness) work in concert to explain why perceptions of effort are moralized, yet further investigation is required to understand the causal paths between judgments of effort, cooperative capacities, cooperative intent, and partner choice. Nonetheless, the evidence we presented

—in which people derived broad moral character evaluations from the efforts that targets devoted to economically and morally neutral tasks—is more consistent with a costly signaling account of effort moralization than prior accounts of these effects.

This evolutionary perspective may provide a more parsimonious framework for integrating research on effort evaluations: the “effort heuristic” (Kruger et al., 2004) may be more functionally dynamic than previously recognized, with effort moralization constituting one of its social functions. Thus, rather than directly *causing* people to moralize effort, cultural beliefs like the PWE may be scaffolded on evolved psychological mechanisms such as shared intuitions about the value of effort. The PWE (and similar work ethics among other populations) may have emerged, then, because it benefited from a combination of being well fit to our psychology (in appealing to an underlying tendency for effort moralization) and culturally useful (in promoting cooperation and industriousness; Henrich, 2020; Henrich & Boyd, 2016). Nevertheless, replications of the current findings outside of rich and industrialized countries are needed to disentangle the universal and cultural aspects of effort moralization. Investigations into potential individual and situational boundaries of effort moralization will also be crucial in determining the mechanisms underlying these effects.

Implications for the Future of Work

We have argued that effort moralization is a “deeply rational” social heuristic (Kenrick et al., 2009); although it may yield seemingly irrational judgments in certain contexts, we believe effort moralization flows from an adaptive evolutionary logic. However, as with other folk-economic beliefs (Boyer & Petersen, 2017), deeply rational intuitions at the individual level can lead to suboptimal and harmful norms at the societal level (Li et al., 2018). Valuing those who appear hardworking and committed, even when they do not directly increase economic value compared with less effortful counterparts or automated alternatives, can create perverse incentives. Thus, effort moralization may help explain how bullshit work (Graeber, 2018) is maintained and rewarded within otherwise efficient economic systems: bullshit labor may serve as a way for millions of workers to signal moral worth through structured drudgery.

Effort moralization may also underlie opposition to policies that forward alternatives to economically unnecessary labor, such as a universal basic income. Cues of effort, and lack of effort, implicitly frame such discussions in terms of perceived deservingness, in turn activating strong social emotions and the denigration of perceived free-riders (Petersen et al., 2011, 2012). Although voters might agree that it is neither efficient nor ethical to waste human potential in economically unnecessary labor (Graeber, 2018), opposition to policies like a universal basic income (Gilberstadt, 2020) suggests that many will also have moral objections to monetary compensation divorced from work.

Yet our results also intimate that even symbolic displays of effort can be leveraged to increase monetary support for the unemployed. If economically unnecessary effort, like running for charity, is deemed worthy of reward, then framing inarguably necessary forms of unpaid labor, like caretaking, in terms of effort may also increase backing for broader social support initiatives. Direct tests of these claims are of mounting importance given the gravity of present policy concerns: labor participation continues to

be eroded by automation, pandemics, and economic shifts. Widespread job loss, whether short-term or permanent, will continue to be socially and economically destabilizing unless policy safeguards are enacted. Understanding the psychology of effort will be paramount in advancing alternatives for accruing social and monetary capital in societies where consequential employment opportunities become increasingly erratic or scarce.

Context of the Research

Our research was motivated by seismic shifts in the labor market. Prior to the COVID-19 pandemic, American labor participation rates had already been falling for twenty years. Many political and business leaders have wrestled with the material concerns posed by unemployment and underemployment. However, we became increasingly concerned with the psychological challenges involved in addressing these economic issues. What does work offer other than a paycheck? Inspired by the contributions of many researchers, we saw that work—and effort more generally—appeared to serve as a moral signal to oneself and others.

The obsession with work for work’s sake has typically been associated with American capitalism or the Protestant Work Ethic. However, influenced by research on folk economics, we suspected that this phenomenon may be more fundamental than a cultural quirk, representing a potentially universal social heuristic. Economic systems are built on these shared psychological foundations. The societal scaling up of intuitive effort moralization may be responsible for some admirable phenomena that are not otherwise materially beneficial to others, such as our appreciation of hardworking artists and athletes. However, we fear it has also created harmful incentive structures that reward workaholism and joyless devotion to mundane efforts that produce little value beyond the signal of effortful engagement. These fears are echoed in David Graeber’s heartbreaking writings on bullshit work, which deeply influenced the trajectory of this article. We will miss his incisive irreverence in political discourse.

References

- Amárah Research. (2019). *Future work*. <https://amarach.com/news-blog-articles/future-work.html>
- Amos, C., Zhang, L., & Read, D. (2019). Hardworking as a heuristic for moral character: Why we attribute moral values to those who work hard and its implications. *Journal of Business Ethics*, 158(4), 1047–1062. <https://doi.org/10.1007/s10551-017-3725-x>
- Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, 59(2), 177–181. <https://doi.org/10.1037/h0047195>
- Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior*, 34(3), 164–175. <https://doi.org/10.1016/j.evolhumbehav.2013.02.002>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bigman, Y. E., & Tamir, M. (2016). The road to heaven is paved with effort: Perceived effort amplifies moral judgment. *Journal of Experimental Psychology: General*, 145(12), 1654–1669. <https://doi.org/10.1037/xge0000230>

- Boyer, P., & Petersen, M. B. (2017). Folk-economic beliefs: An evolutionary cognitive model. *Behavioral and Brain Sciences*, *41*, e158. <https://doi.org/10.1017/S0140525X17001960>
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216. <https://doi.org/10.1177/135910457000100301>
- Celniker, J. B., Gregory, A., Koo, H. J., Piff, P. K., Ditto, P. H., & Shariff, A. F. (2020, October 9). *The moralization of effort*. osf.io/zwqbe
- Clement, T. S., Feltus, J. R., Kaiser, D. H., & Zentall, T. R. (2000). Work ethic” in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, *7*(1), 100–106. <https://doi.org/10.3758/BF03210727>
- Cosmides, L., & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology*, *64*(1), 201–229. <https://doi.org/10.1146/annurev.psych.121208.131628>
- Dahlgreen, W. (2015). “37% of British workers think their jobs are meaningless.” *YouGov*. <https://yougov.co.uk/news/2015/08/12/british-jobs-meaningless/>
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch’s t-Test instead of student’s t-Test. *Revue Internationale de Psychologie Sociale*, *30*(1), 92–101. <https://doi.org/10.5334/irsp.82>
- Everett, J. A. C., Pizarro, D. A., & 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>
- Frey, C. B., & Osborne, M. (2013). The future of employment: How susceptible are jobs to computerisation? <https://www.oxfordmartin.ox.ac.uk/downloads/academic/future-of-employment.pdf>
- Furnham, A., Bond, M., Heaven, P., Hilton, D., Lobel, T., Masters, J., Payne, M., Rajamanikam, R., Stacey, B., & Van Daalen, H. (1993). A comparison of Protestant Work Ethic beliefs in thirteen nations. *The Journal of Social Psychology*, *133*(2), 185–197. <https://doi.org/10.1080/00224545.1993.9712136>
- Gilberstadt, H. (2020). *More Americans oppose than favor the government providing a universal basic income for all adult citizens*. Pew Research Center. <https://www.pewresearch.org/fact-tank/2020/08/19/more-americans-oppose-than-favor-the-government-providing-a-universal-basic-income-for-all-adult-citizens/>
- Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology*, *213*(1), 103–119. <https://doi.org/10.1006/jtbi.2001.2406>
- Gomila, R. (2021). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*, *150*(4), 700–709. <https://doi.org/10.1037/xge0000920>
- Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, *24*(1), 38–44. <https://doi.org/10.1177/0963721414550709>
- Graeber, D. (2018). *Bullshit jobs: A theory*. Simon & Schuster.
- Harris, L. T., & Fiske, S. T. (2011). Dehumanized perception: A psychological means to facilitate atrocities, torture, and genocide? *The Journal of Psychology*, *219*(3), 175–181. <https://doi.org/10.1027/2151-2604/a000065>
- Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity: International Journal of Methodology*, *43*(1), 59–74. <https://doi.org/10.1007/s11135-007-9077-3>
- Henrich, J. (2020). *The WEIRD people in the world: How the west became psychologically peculiar and particularly prosperous*. Farrar, Straus & Giroux.
- Henrich, J., & Boyd, R. (2016). How evolved psychological mechanisms empower cultural group selection. *Behavioral and Brain Sciences*, *39*, e40. <https://doi.org/10.1017/S0140525X15000138>
- Hsee, C. K., Blount, S., Loewenstein, G. F., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin*, *125*(5), 576–590. <https://doi.org/10.1037/0033-2909.125.5.576>
- Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Sciences*, *22*(4), 337–349. <https://doi.org/10.1016/j.tics.2018.01.007>
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. <https://doi.org/10.1037/a0028347>
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. <https://doi.org/10.1037/0003-066X.58.9.697>
- Kenrick, D. T., Griskevicius, V., Sundie, J. M., Li, N. P., Li, Y. J., & Neuberg, S. L. (2009). Deep rationality: The evolutionary economics of decision making. *Social Cognition*, *27*(5), 764–785. <https://doi.org/10.1521/soco.2009.27.5.764>
- Kruger, J., Wirtz, D., Van Boven, L., & Altermatt, T. W. (2004). The effort heuristic. *Journal of Experimental Social Psychology*, *40*(1), 91–98. [https://doi.org/10.1016/S0022-1031\(03\)00065-9](https://doi.org/10.1016/S0022-1031(03)00065-9)
- Lamont, M. (2000). *The dignity of working men: Morality and the boundaries of race, class, and immigration*. Harvard University Press. <https://doi.org/10.4159/9780674039889>
- Li, N. P., van Vugt, M., & Colarelli, S. M. (2018). The evolutionary mismatch hypothesis: Implications for Psychological Science. *Current Directions in Psychological Science*, *27*(1), 38–44. <https://doi.org/10.1177/0963721417731378>
- Lydall, E. S., Gilmour, G., & Dwyer, D. M. (2010). Rats place greater value on rewards produced by high effort: An animal analogue of the “effort justification” effect. *Journal of Experimental Social Psychology*, *46*(6), 1134–1137. <https://doi.org/10.1016/j.jesp.2010.05.011>
- Markovits, D. (2019). *The meritocracy trap: How America’s foundational myth feeds inequality, dismantles the middle class, and devours the elite*. Penguin Press.
- Mirels, H. L., & Garrett, J. B. (1971). The Protestant ethic as a personality variable. *Journal of Consulting and Clinical Psychology*, *36*(1), 40–44. <https://doi.org/10.1037/h0030477>
- Nedelkoska, L., & Quintini, G. (2018). *Automation, skills use and training (OECD Social, Employment and Migration Working Papers, No. 202)*. OECD Publishing. <https://doi.org/10.1787/2e2f4eea-en>
- Norenzayan, A., & Heine, S. J. (2005). Psychological universals: What are they and how can we know? *Psychological Bulletin*, *131*(5), 763–784. <https://doi.org/10.1037/0033-2909.131.5.763>
- OECD. (2019). *Hours worked (indicator)*. <https://doi.org/10.1787/47be1c78-en> Accessed December 8, 2019.
- Olivola, C. Y., & Shafir, E. (2013). The martyrdom effect: When pain and effort increase prosocial contributions. *Journal of Behavioral Decision Making*, *26*(1), 91–105. <https://doi.org/10.1002/bdm.767>
- Petersen, M. B., Slothuis, R., Stubager, R., & Togeby, L. (2011). Deservingness versus values in public opinion on welfare: The automaticity of the deservingness heuristic. *European Journal of Political Research*, *50*(1), 24–52. <https://doi.org/10.1111/j.1475-6765.2010.01923.x>
- Petersen, M. B., Sznycer, D., Cosmides, L., & Tooby, J. (2012). Who deserves help? Evolutionary psychology, social emotions, and public opinion about welfare. *Political Psychology*, *33*(3), 395–418. <https://doi.org/10.1111/j.1467-9221.2012.00883.x>
- 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>
- Rhee, J. J., Schein, C., & Bastian, B. (2019). The what, how, and why of moralization: A review of current definitions, methods, and evidence in moralization research. *Social and Personality Psychology Compass*, *13*(12), 1–14. <https://doi.org/10.1111/spc3.12511>

- Schaumberg, R. L., & Mullen, E. (2017). From incidental harms to moral elevation: The positive effect of experiencing unintentional, uncontrollable, and unavoidable harms on perceived moral character. *Journal of Experimental Social Psychology, 73*, 86–96. <https://doi.org/10.1016/j.jesp.2017.06.016>
- Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological & Personality Science, 8*(4), 379–386. <https://doi.org/10.1177/1948550617715068>
- Schouten & Nelissen. (2017). 4 op de 10 medewerkers vinden hun werk niet zinvol [4 out of 10 employees do not think their work is meaningful]. <https://www.sn.nl/inspiratie/vitaliteit-stress-en-werkdruk/4-op-de-10-medewerkers-vinden-hun-werk-niet-zinvol>
- Shepperd, J. A., Arkin, R. M., Strathman, A., & Baker, S. M. (1994). Dysphoria as a moderator of the relationship between perceived effort and perceived ability. *Journal of Personality and Social Psychology, 66*(3), 559–569. <https://doi.org/10.1037/0022-3514.66.3.559>
- Smith, K. M., & Apicella, C. L. (2020). Hadza hunter-gatherers disagree on perceptions of moral character. *Social Psychological & Personality Science, 11*(5), 616–625. <https://doi.org/10.1177/1948550619865051>
- Stibbard-Hawkes, D. N. E., Attenborough, R. D., & Marlowe, F. W. (2018). A noisy signal: To what extent are Hadza hunting reputations predictive of actual hunting skills? *Evolution and Human Behavior, 39*(6), 639–651. <https://doi.org/10.1016/j.evolhumbehav.2018.06.005>
- Tierney, W., Hardy, J. H., III, Ebersole, C. R., Leavitt, K., Viganola, D., Clemente, E. G., Gordon, M., Dreber, A., Johannesson, M., Pfeiffer, T., & Uhlmann, E. L. (2020). Creative destruction in science. *Organizational Behavior and Human Decision Processes, 161*(July), 291–309. <https://doi.org/10.1016/j.obhdp.2020.07.002>
- Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science, 10*(1), 72–81. <https://doi.org/10.1177/1745691614556679>
- Uhlmann, E. L., & Sanchez-Burks, J. (2014). The implicit legacy of American Protestantism. *Journal of Cross-Cultural Psychology, 45*(6), 992–1006. <https://doi.org/10.1177/0022022114527344>
- van Hoorn, A., & Maseland, R. (2013). Does a Protestant work ethic exist? Evidence from the well-being effect of unemployment. *Journal of Economic Behavior & Organization, 91*, 1–12. <https://doi.org/10.1016/j.jebo.2013.03.038>
- Weber, M. (1958). *The Protestant ethic and the spirit of capitalism*. Charles Scribner's Sons. (Original work published 1904)
- Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences, 45*, E1. <https://doi.org/10.1017/S0140525X20001685>

Received April 3, 2020

Revision received May 2, 2022

Accepted May 3, 2022 ■