Registered Report: See me, judge me, pay me: Gendered effort moralization in work and care.

Roth[1,2,3], L. H. O., Tissot[1,4], T. T., Fischer[3], T., Masak[3], S., C.

[3]Department of Occupational, Economic, and Social Psychology, Faculty of Psychology, University of Vienna, Austria

[4] Department of Developmental, Personality, and Social Psychology, Faculty of Psychology and Educational Sciences, Ghent University, Ghent, Belgium

[1]the first two authors share first authorship

[2]corresponding author (leopold.roth@univie.ac.at)

Authors note

Leopold Helmut Otto Roth[1,2], 0000-0002-1120-4733

Tassilo Tom Tissot[1], 0000-0002-6215-975X

Thea Fischer, 0009-0002-0490-4894

Sophie Charlotte Masak, 0009-0005-1555-9258

Author contribution

Conceptualization: LR, TT, TF

Data curation: LR, TT

Formal analysis: LR

Funding acquisition: LR, TF

Investigation: LR, TT, TF, CM

Methodology: LR, TT

Project administration: LR, TT, TF, CM

Visualization: LR, CM

Writing - original draft: LR, TT, TF, CM

Writing - review & editing: LR, TT, TF, CM

See me, judge me, pay me

Abstract

Displaying high effort at work is rewarded with more positive moral judgments (effort moralization effect) and increased cooperation partner attractiveness. This holds, even if higher effort is unrelated to better performance. Yet, current evidence is exclusively based on males, mostly situated in the work context. This prohibits generalization to the full population and neglects critical aspects of our lives, such as the care context (e.g., unpaid care for elders). To address this gap, we conducted two Studies (Study 1: $N_{work} = 859$, Study 2: $N_{care} = 701$) testing the effect between genders and contexts—work and care. Study materials featured two actors performing the same task, requiring different levels of effort (high/low). Participants rated the actor's morality, suggested hourly salary, and reported their satisfaction to cooperate with them as assigned partners. The results confirmed the effort moralization effect in work contexts but were mixed for the care context, potentially due to the inherently moral nature of the behavior. There were no gender differences, supporting the demographic generalizability of the effect. Contrary to previous studies, effort did not influence suggested pay, and participants expressed greater satisfaction with low-effort actors as assigned cooperation partners. While further research is needed to explore the boundary conditions of effort moralization in different contexts, the findings support its role as a robust heuristic for moral judgment in the work context.

Keywords: Registered Report, effort moralization, gender, work, care, bias, judgment

Design Table

Question	Hypothesis	Sampling plan	Analysis plan	Rationale for deciding the sensitivity	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
Aim 1: Replication of core effect						
Can we replicate the effort moralization effect?	Individuals who invest **higher effort** in their **work** are judged **higher in morality.**	We will collect data through **Prolific**. The **total required** sample size is $N = \textbf{648}$, which we will **oversample** to $N = \textbf{700}$. The required sample size per t-test is $N = \textbf{272}$.	Using a **one-sided dependent Welch's t-test** and respective **Bayes Factor**, we will test for differences in **perceived moral character** (core goodness and value commitment). We will further test for differences in **perceived warmth**, **perceived competence,** and **pay deservingness**. Yet, prior research highlighted variance in these more distal measures.	Based on the **smallest effect size of interest** (Lakens, 2022) approach, we aim to power for a small effect $d = \textbf{0.20}$ (Cohen, 1988) ($a = .05$, $1\text{-}\beta = .95$, one-tailed). This was computed using G*Power 3.1.9.7 [see supplemental material, https://osf.io/s8ec5/].	If the effect is not found, the effort moralization effect is **not replicated** in the target magnitude. This can be due to the absence of the effect or due to the pooling of genders, which is tested in the following steps.	Effort moralization theory's **generalizability** could be shown undetectable under the current conditions of the study.
	Individuals who invest **higher effort** in their **care work** are judged **higher in morality.**				If the effect is not found, it is potentially **not generalizable** to care work. Yet, the following analyses test the results in a more fine-grained manner.	Effort moralization is potentially not **generalizable** to the care context.
Aim 2: moral character as a function of gender, context, and effort						
Are there	Moral character	In study 1	Using **mixed-ANOVA** with 2	Based on the **smallest effect**	The ANOVA can	If effort doesn't

differences in effort moralization in the **work context** by gender and effort?	judgment differs by gender and effort.	(work context), we will sample *N* = **350** individuals (computed *N* = **324**)	(gender: female/male) x 2 (effort: high/low) design. **Gender** serves as a **between-subject** factor, and **effort** is a **within-subject** factor. **We further test the interaction of both terms.** The respective **Bayes Factor** is computed for each term.	**size of interest** (Lakens, 2022) approach, we aim to power for a small effect η^2 = **.01** (Cohen, 1988) (*a* = .05, 1-β = .95, 2(gender)x2(effort)). This was computed, using G*Power 3.1.9.7 [see supplemental material, https://osf.io/s8ec5/].	**illustrate whether gender and/or effort** differentially influence **moral character** judgment in the **work context**.	affect moral character judgment, the effect is potentially not replicable in this context. If it shows differences by gender, the effect is potentially heterogeneous between genders (female/male).
Are there differences in effort moralization in the **care context** by gender and effort?	Moral character judgment differs by gender and effort.	In study 2 (care context), we will sample *N* = **350** individuals (computed *N* = **324**)	Using **mixed-ANOVA** with 2 (gender: female/male) x 2 (effort: high/low) design. **Gender** serves as a **between-subject** factor, and **effort** is a **within-subject** factor. **We further test the interaction of both terms.** The respective **Bayes Factor** is computed for each term.	Based on the **smallest effect size of interest** (Lakens, 2022) approach, we aim to power for a small effect η^2 = **.01** (cite Cohen, 1988) (*a* = .05, 1-β = .95, 2(gender)x2(effort)). This was computed, using G*Power 3.1.9.7 [see supplemental material, https://osf.io/s8ec5/].	The ANOVA and post-hoc tests can **illustrate whether gender and/or effort** differentially influence **moral character** judgment in the **care context**.	If effort doesn't affect moral character judgment, the effect is potentially not replicable in this context. If it shows differences by gender, the effect is potentially heterogeneous between genders (female/male).
Aim 3: cooperation partner satisfaction as a function of gender and effort						
Do gender and effort influence cooperation	**Work context:** cooperation satisfaction is	In each study, we will sample	Using **mixed-ANOVA** with 2 (gender: female/male) x 2 (effort: high/low) design.	Based on the **smallest effect size of interest** (Lakens, 2022) approach, we aim to	We will be informed as to which degree **effort** is crucial for cooperation	Effort might not be a meaningful predictor of

satisfaction?	predicted by gender and effort	sample $N = 350$ individuals (computed $N = 272$)	**Gender** serves as a **between-subject** factor, and **effort** is a **within-subject** factor. **We further test the interaction of both terms.** The respective **Bayes Factor** is computed for each term.	power for a small effect $\eta^2 = .01$ (Cohen, 1988) ($a = .05$, $1\text{-}\beta = .95$, 2(gender)x2(effort)). This was computed, using G*Power 3.1.9.7 [see supplemental material, https://osf.io/s8ec5/].	satisfaction **between women and men** in the **work context**.	cooperation satisfaction. Further, there might not be differences between females and males.
	Care context: cooperation satisfaction is predicted by gender and effort		Using **mixed-ANOVA** with 2 (gender: female/male) x 2 (effort: high/low) design. **Gender** serves as a **between-subject** factor, and **effort** is a **within-subject** factor. **We further test the interaction of both terms.** The respective **Bayes Factor** is computed for each term.		We will be informed as to which degree **effort** is crucial for cooperation satisfaction **between women and men** in the **care context**.	
Exploratory Analysis: are differences in effort moralization moderated by gender norm endorsement						
Are differences in effort moralization between genders moderated by gender norm endorsement?	This will be tested in the **work** and **care** context	This exploratory analysis will be performed on the computed sample size of **Aim 2**	Using **multilevel modeling,** we will test the effect of the **interaction of gender and gender norm endorsement** and the main effect of effort on moral judgment (core goodness & value commitment).	This exploratory analysis will be performed on the computed sample size of **Aim 2**	We will be informed whether gender norm endorsement moderates the influence of gender on effort moralization and whether the effect is generalizable in work and care contexts.	The effect is either generalizable in both contexts, context-dependent, or not observable with the present data.

			lmer-formula: morality ~ gender*gender_norm + effort + (1|subject) We will further compare the **Bayes Factor** of the model against the model without gender norm endorsement.			

See me, judge me, pay me

Introduction

The effort moralization effect

Social judgment is crucial in daily life. People frequently encounter strangers and have to make quick inferences about their character, such as deciding whether it is safe to sit next to someone on the bus. Considering how important these decisions are, it is notable that we need to rely on rough, incomplete information to make such critical assessments—it wouldn't be feasible to administer a personality test to every passenger on the bus before choosing where to sit. We navigate such social interactions as cognitive misers, using simple processing mechanisms to reduce cognitive load (Fiske & Taylor, 1991). Instead of seeking complete information, we rely on environmental cues (e.g., valence of facial expressions, Fox et al., 2002), stereotypes (Aronson et al., 2021), heuristics and resulting cognitive biases (Tversky & Kahneman, 1974), and personal learning experiences (Behrens et al., 2008).

One factor that plays a dominant role in the perception of other people is moral information (Brambilla & Leach, 2014; Goodwin et al., 2014; Wojciszke, 2005). In this context, a particular bias has gained recent attention: the effort moralization effect (Amos et al., 2019; Bigman & Tamir, 2016; Celniker et al., 2023; Fwu et al., 2014). It describes the tendency of observers to make moral character judgments based on the observed effort a person puts into a given behavior. The perceived intensity of effort amplifies moral judgments: actions perceived as "good" appear even more virtuous, while "bad" behaviors seem worse the more effort is involved (Bigman & Tamir, 2016). For example, it has been shown that donations of time are perceived as a greater (emotional) investment, and therefore better moral character, compared to donations of money (Johnson & Park, 2021; Reed et al., 2007).

Interestingly, the effort moralization effect persists, even when additional effort does not lead to increased performance (e.g., better outcomes at work, Celniker et al., 2023). This

points to the interpretation that the exertion of effort is valued by itself rather than its practical benefits. This observation was replicated well, yet it appears to vary between cultures in magnitude (Mexico: $d = .14–.28$, Germany: $d = .34–.37$, France: $d = .38$, US: $d = .60$, South Korea: $d = .71$; Celniker et al., 2023; Tissot & Roth, 2025).

Further, it was shown that the display of high effort—contrasted with low effort for the same outcome—led to an increased chance of being selected as a cooperation partner in a follow-up trust game (Celniker et al., 2023), which has meaningful implications, especially for the work and career context.

Current gaps in the effort moralization literature: context and gender

Prior literature has mostly focused on two types of contexts, in which effort moralization comes to play: work contexts (Amos et al., 2019; Celniker et al., 2023) and charity or helping behavior (Bigman & Tamir, 2016; Celniker et al., 2023). These contexts are justified targets, as these are impactful domains in our lives and commonly demand effort. Yet, it left the large domain of unpaid care work uncovered, which is estimated to make up 245 hours of annual work for the average American citizen (Mason & Robbins, 2024). Two-thirds of care work (65%) is done by women (Mason & Robbins, 2024), and often goes with little societal recognition (Antonopoulos, 2008) and high mental load (Dean et al., 2021), while it surpasses the value of 1 trillion dollars in the US per year (National Partnership for Women & Families, 2024). Further—to our knowledge—the literature on the effort moralization effect focused on either male or gender-neutral (e.g., Person A) vignettes and excluded female actors from described scenarios. Hence, investigating the role of an additional critical context, as well as between-gender effects and differential effects on moral character judgment, appears warranted for the generalizability of the effect. Understanding gender bias in the effort moralization effect is crucial for addressing inequalities (e.g., reinforcement of traditional gender roles).

See me, judge me, pay me

Celniker et al. (2023) found that individuals who exert more effort to achieve the same performance in widget-making are more likely to be chosen as partners in a trust game. However, cooperation partners are not always freely chosen but can be assigned as well (e.g., project assignments in the workplace). We, therefore, extend the literature by assessing individuals' satisfaction with assigned, instead of freely chosen, partners. This provides additional insights that reflect the cooperative dynamics frequently found in everyday life.

Gendered stereotyping in moral judgment and effort perceptions

As shown in prior research, social judgment is not immune to the influences of stereotyping, including gender biases. These extend to differing expectations of behavior and personality based on a person's gender. For instance, while men are often seen as more agentic, women are perceived as more communal (e.g., caring or helpful; Hentschel et al., 2019). These expectations may inform differences in effort moralization—for example, through backlashing—and can differ between contexts.

This describes how expectations, for instance, those formed by gender, can lead to differing social judgments (Rudman, 1998). Individuals who deviate from stereotypical behavior tend to be subjected to harsher sanctions. For example, women receive more severe disciplinary sanctions for ethical violations in the workplace (Kennedy et al., 2016; Rudman, 1998), whereas men face greater criticism for non-agentic behavior in leadership contexts (Moss-Racusin et al., 2010).

For effort moralization, these prior findings hold potential for differences in judgment between gender, effort levels, and social context. Male stereotypes of agentic behavior could cause stronger differences in moral judgment at work, as men are expected to work hard and autonomously.

The interplay between effort moralization and gender may expand to the caregiving context. Although gender roles have shifted, with more women entering the workforce

(Toossi & Morisi, 2017) and men contributing more to family labor (Sayer, 2016), women still do most of the care work (Charmes, 2019). These persistent gender role expectations shape how caregiving efforts are perceived. Research on double standards has shown that mothers face harsher criticism for low care efforts while fathers receive praise for being involved (Deutsch & Saxon, 1998).

In sum, existing literature indicates that gender biases likely play a role in effort moralization. The specific goals are outlined below.

How can the present project inform psychological theorizing?

As previously outlined, multiple overarching theories may influence the moralization of effort across gender and contextual lines. The extant literature offers substantiation for several possible directions of effects, and our study may yield different outcome patterns. These include the following: (1) A gender-based discrepancy in the moral evaluation of effort is observed in the domains of work and care work, (2) This discrepancy is only evident in the work context, (3) This discrepancy is only evident in the care context, (4) No differences in the moral evaluation of effort are found based on the gender of the actors.

The first pattern (1) would indicate that the gender of the actors in question does play a role in the moralization of effort and that stereotypical gender role expectations have an influence on the moral judgments of others across contexts. Within this pattern, different directions of effects are imaginable. For example, in the work context, gendered stereotypes could lead to a higher moral judgment of men, despite their level of effort, because their presence at work is stereotypically assumed and valued, while women are considered less capable of performing in the work context (Rudman, 1998; Sterling & Reichman, 2016). Conversely, the same stereotypes could potentially result in higher moral judgments for men in the high-effort condition but lower—compared to women—in the low-effort condition, as men are expected to demonstrate invested and agentic behavior at work. Consequently,

demonstrating low effort could lead to lower moral judgments for men (Moss-Racusin et al., 2010).

Within the context of care, it is plausible that men will receive higher morality ratings in the high-effort condition compared to women, as the social expectations of women's pro-social behavior might render their high effort less exceptional. Consequently, even low-effort behavior exhibited by men might receive higher moral ratings than high-effort behavior displayed by women (Deutsch & Saxon, 1998). Conversely, while men might receive higher moral evaluation for high effort than women, low effort might be evaluated equally negatively across genders, as care work is expected and socially valued (Samtleben & Müller, 2022).

Naturally, it is also possible that the effects will be found in only one of the contexts eliciting the result patterns (2) or (3). This would suggest that the work or care context, respectively, is or has grown to be more resilient to stereotypical gender role expectations regarding effort moralization. For instance, differences may be found only in the care context, while gender stereotypes may not play a significant role for effort moralization in the work context (anymore). Alternatively, if no differences in effort moralization are identified in both contexts, as reflected by result pattern (4), the absence of effects could indicate neglectable influences of gender stereotypes on the moral judgment of effort.

Current Studies

The current project focused on three core aims: I) to test the replicability of the effort moralization effect (Study 1) and explore its generalizability to the care context (Study 2), II) to examine how the effort moralization effect interacts with gender across different contexts, and III) to investigate whether cooperation partner satisfaction differs by gender and effort.

Method

Sample and sample size

Using the smallest effect size of interest approach (Lakens, 2022), we powered both Studies to detect a small effect (Cohen, 1988) in a 2x2 mixed ANOVA ($\eta^2 = .01$, $\alpha = .05$, $1\text{-}\beta = .95$). This resulted in a minimum sample size of $N = 324$. We conducted a second power analysis for the interaction effect with the same parameters ($d = 0.20$), resulting in a similar required sample size ($N = 325$) to countercheck between computation tools. The total target sample size across both studies was $N = 700$ to buffer against exclusions. The computation was done using G*Power 3.1.9.7 (Faul et al., 2009) and IntXPower (Sommet et al., 2023), documented in the supplemental material (https://osf.io/s8ec5/)[1]. Participants were recruited via Prolific and consisted of individuals based in the US. The majority of the sample attended university (77.24%). For more fine-grained information, please refer to the supplemental material.

Table 1

Descriptives by Study

	N	age: M (SD)	range	%female	SES
overall	1,560	42.32 (13.68)	18-86	47.56	5.20 (1.75)
Study 1 (work)	859	43.43 (13.75)	18-81	48.89	5.22 (1.77)
Study 2 (care)	701	40.97 (13.49)	18-86	45.93	5.18 (1.73)

Materials

Building on previous research, we employed and adapted materials from a seminal study in the field (Celniker et al., 2023). We designed a new vignette for the caregiving context and adapted the work vignette to feature less stereotypical tasks (office instead of factory scenario).

---

[1] We are aware that some patterns of interaction terms potentially require larger samples. Given that the pattern is not known at the time of power computation, it can happen that some interaction forms might not be sufficiently powered through our sample.

We assessed the perceived morality of actors using 13 trait items (Celniker et al., 2023) that have been demonstrated to distinguish between two types of moral virtues (Piazza et al., 2014). While core goodness traits like kindness are universally good, the moral valence of value commitment traits like dedication depends on the context—a kind murderer is "better" than an unkind one, while a dedicated murderer is "worse" than an undedicated one. All trait items were rated on a 7-point scale.

Following the procedure of Celniker et al. (2023), warmth and competence, two universal dimensions of social cognition for anticipating interdependence and status, were assessed with one item each on a 7-point scale (Fiske et al., 2007).

The perceived effort, quality, difficulty, and work value were measured with single items on a 7-point scale as manipulation checks.

The item assessing the pay deservingness of each actor differed between the work and care context study. In the work context study, participants responded on a sliding scale, anchored at a midpoint that reflected a realistic average office worker salary in the US (ERI, 2025). For care, no reference point was provided given that this work is typically unpaid. Instead, participants could freely choose a wage between $0 and $50. This allowed us to assess the perceived value of care work. Further, we assessed how satisfied participants would be on a 7-point scale to have either actor as an assigned cooperation partner in a work project (work context) or organizing a charity event (care context).

In addition, for exploratory purposes, we incorporated a short version of the gender role belief scale into our study to explore potential moderating effects of traditional gender role endorsement on effort moralization (Brown & Gladstone, 2012). All materials are available in the supplemental materials (https://osf.io/s8ec5/).

See me, judge me, pay me

Table 2

Overview of Measures

Construct (*n* items)	Example item	Low anchor	High anchor
Core goodness (6)[a] (.93/.93)[d]	Honest	Does not describe X well	Describes X extremely well
Value commitment (7)[a] (.93/.93)[d]	Dedicated	Does not describe X well	Describes X extremely well
Competence/warmth (2)[a]	Competent	Does not describe X well	Describes X extremely well
Effort (1)[b]	How much effort do you think X puts into his/her (care) work?	No effort at all	A lot of effort
Quality (1)[c]	In your opinion, how well does X perform his/her (care) work?	Very bad	Very good
Difficulty (1)[c]	Compared to other jobs/ care work, how difficult is X's (care) work?	Not at all difficult	Extremely difficult
Work value (1)[c]	How valuable do you think X's (care) work is?	Not valuable at all	Extremely valuable
Pay deservingness (work) (1)[a]	The average office worker at the company makes $24 an hour. How much do you think X should make per hour?	$12	$36
Pay deservingness (care) (1)[a]	Imagine that X was paid for his/her care work. How much should s/he be paid per hour?	$0	$50
Collaboration partner choice (1)[a]	[...] Please indicate how satisfied you would be to work with either X or Y.	Extremely dissatisfied	Extremely satisfied
Gender role beliefs (10)[e] (.88)[d]	Women with children should not work outside the home if they don't have to financially.	Strongly disagree	Strongly agree

Note. [a]These variables are the focal dependent measures; [b]This measure serves as manipulation check and exclusion criterion; [c]These measures serve as manipulation check but not as exclusion criterion; [d]Reliabilities (low effort/high effort), for values after second digit, see supplemental material; [e]pre-registered exploratory moderator.

Procedure

The data was collected in two separate Studies in January and February 2025, with participants from one study being excluded from participating in the other. In both Studies, after providing informed consent, participants were presented with a scenario from either the work (Study 1) or care (Study 2) context.

See me, judge me, pay me

The vignettes featured two individuals—either male or female—who perform the exact same tasks at the same quality level but differ in the amount of effort required. For example, the work context vignette reads as follows[2]:

Anna and Sophie work at the same company and process similar orders in the company's office. Both Anna and Sophie are able to process approximately three orders per hour, which means they complete one case every 20 minutes. The average value of a completed case for the company is $50.00. Quality control inspections indicate that 96% of Anna's and Sophie's orders are error-free and complete. On average, Anna and Sophie each process correct orders worth $144 per hour.

For Anna, processing orders requires minimal effort — although she works as quickly as possible, she finds the work easy.

For Sophie, however, processing orders requires a lot of effort — although she works as quickly as possible, she finds the work hard.

After reading the vignette, participants completed a series of dependent measures for each featured individual in randomized order and the gender role belief scale (Brown & Gladstone, 2012). Within each study, gender (male vs. female names in the vignette) served as the between-subject factor, and effort (high vs. low) was the within-subject factor.

Both Studies took approximately 7 minutes per participant, and all data were collected via Prolific. Participants received compensation according to the platform's standard rates (8.00$/hour).

Data cleaning

To ensure valid responses, participants who self-reported insufficient English proficiency (below "very good") and participants who failed one of two attention checks were excluded from the analysis. The probability of passing both attention checks by random guessing was $(\frac{1}{7} \ x \ \frac{1}{7}) = 2.04\%$. Participants completing the study 3 standard deviations faster than the average participant were excluded. There was no exclusion for slow

---

[2] The vignettes were designed to reduce stereotyped associations. Hence we adapted the vignette by Celniker et al. (2023) from a factory to an office setting and designed the care vignette in a way that non-relational tasks are in the foreground (e.g., lawn mowing instead of emotional support).

participation. In line with the procedure by Celniker et al. (2023) and Tissot and Roth (2025), we further excluded participants who rated the low-effort behavior as equally or more effortful than the high-effort behavior. Participants who did not complete the study were excluded from the final analysis.

Of the initial $N = 2,270$ participants ($N_{\text{study 1}} = 1,142$, $N_{\text{study 2}} = 1,128$), we excluded $n = 710$ participants using the pre-registered criteria. In study 1, $n = 283$ participants were excluded (incomplete participation: 27, attention checks: 69, too fast completion: 0, language skills: 12, didn't see the effortful condition as more effortful: 175). In study 2, $n = 427$ participants were excluded (incomplete participation: 17, attention checks: 59, too fast completion: 0, language skills: 14, didn't see the effortful condition as more effortful: 337).

Data analysis

Aim 1: Replication of core effect

To test whether the original effort moralization effect could be replicated in the work context and generalized to the care context, we conducted a series of dependent, one-sided Welch's t-tests, comparing moral judgments of the described actor between high- and low-effort conditions. We further computed the respective effect size (Cohen's d) and Bayes Factor (BF_{10}). Additionally, we compared perceived warmth, competence, and pay deservingness between high and low-effort actors.

Aim 2: Moral character as a function of gender, context, and effort

To examine the effects of gender and effort on moral character judgments—in both work and care contexts—we used a mixed effect ANOVA (between-subjects factor: gender, within-subjects factor: effort) with an interaction term. For all terms, the respective Bayes Factor (BF_{10}) was computed to quantify evidence of absence and presence of effects. Bayesian model comparison against the null model was performed with Laplace approximation, using JASP (JASP Team, 2020).

See me, judge me, pay me

We applied the same mixed-effects ANOVA procedure to participants' suggested hourly payment to test for evidence of gender pay gaps.

Aim 3: Cooperation partner satisfaction as a function of gender and effort

We used the same mixed-effects ANOVA procedure (as described in Aim 2) to compare satisfaction with assigned cooperation partners.

Results

Aim 1: Replication of core effect

Study 1: work context

Study 1 replicated the effort moralization effect in the work context, with higher perceived morality in the high-effort actor. Contrary to expectations, cooperation satisfaction pointed in the opposite direction, indicating higher satisfaction with the actor who required less effort for the same behavior. Surprisingly, no effect on pay deservingness was found.

Table 3

Comparison of Moral Character Judgment by Effort and Context

	Low effort: M (SD)	High effort: M (SD)	d [95% CI]	BF_{10}	$^{1}d_{robustness}$
Study 1: work context					
core goodness	5.18 (1.13)	5.38 (1.13)	**-0.30** [-0.37, -0.23]ˣ	>1,000	-0.19ˣ
value commitment	5.52 (1.08)	5.89 (0.95)	**-0.37** [-0.44, -0.30]ˣ	>1,000	**-0.22**ˣ
warmth	4.77 (1.38)	4.95 (1.36)	-0.19 [-0.26, -0.12]ˣ	>1,000	-0.12ˣ
competence	6.17 (0.98)	5.48 (1.24)	**0.53** [0.46, 0.61]ˣ	>1,000	**0.50**ˣ
pay deservingness	26.37 (3.36)	26.28 (3.38)	0.04 [-0.03, 0.11]	0.076	0.09*
cooperation	6.03 (1.11)	5.62 (1.19)	**0.30** [0.23, 0.37]ˣ	>1,000	**0.34**ˣ

Note. *p < .05, ˣp < .001, Cohen's *d* ≥ .20 (smallest effect size of interest) in **bold print**; [1]as suggested in the stage 2 review process, we repeated the analysis without excluding participants who didn't see the high effort condition as more effortful (N_{work} = 1,034 instead of N_{work} = 859).

See me, judge me, pay me

Study 2: care context

Study 2 partially replicated the effort moralization effect–only for core goodness–and showed the same reversed cooperation effect. No difference in pay deservingness was observed.

Table 4

Comparison of Moral Character Judgment by Effort and Context

	Low effort: M (SD)	High effort: M (SD)	d [95% CI]	BF_{10}	$^{1}d_{robustness}$
Study 2: care context					
core goodness	5.84 (0.97)	5.94 (0.96)	-0.14 [-0.21, -0.06]ˣ	24.90	0.06
value commitment	6.04 (0.93)	6.02 (1.00)	0.02 [-0.06, 0.09]	0.048	**0.22**ˣ
warmth	5.61 (1.24)	5.63 (1.24)	-0.02 [-0.09, 0.06]	0.046	0.09*
competence	6.26 (0.99)	5.57 (1.34)	**0.49** [0.42, 0.57]ˣ	>1,000	**0.52**ˣ
pay deservingness	23.28 (8.29)	23.40 (8.20)	-0.04 [-0.11, 0.04]	0.066	0.12ˣ
cooperation	6.34 (0.92)	5.88 (1.11)	**0.41** [0.33, 0.49]ˣ	>1,000	**0.52**ˣ

Note. *$p < .05$, ˣ$p < .001$, Cohen's $d \geq .20$ (smallest effect size of interest) in **bold print**; [1]as suggested in the stage 2 review process, we repeated the analysis without excluding participants who didn't see the high effort condition as more effortful ($N_{care} = 1,038$ instead of $N_{care} = 701$).

Aim 2: Moral character as a function of gender, context, and effort

Study 1: work context

To test whether moral character judgments differed between genders across effort levels, we used a 2x2 mixed ANOVA model (between: gender, within: effort). For core goodness, we observed no main effect of gender ($F(1,857) = 1.78$, $p = .182$, $\eta^2 = .002$) and no interaction of gender and effort level ($F(1,857) = 0.00$, $p = .981$, $\eta^2 < .001$), but the assumed main effect of effort in the expected direction (see Figure 1) with higher moral judgment for higher exerted effort ($F(1,857) = 75.51$, $p < .001$, $\eta^2 = .008$). The model including effort (low/high) as a repeated-measures factor provided the strongest evidence compared to the

null model, indicating that effort was the most likely driver of the observed variation in core

goodness ratings (BF$_{10}$ > 1,000).

Very similar results were observed regarding value commitment with no significant

main effect for gender ($F(1,857) = 0.04$, $p = .850$, $\eta^2 < .001$), no significant interaction

($F(1,857) = 0.23$, $p = .631$, $\eta^2 < .001$), but an effect for effort ($F(1,857) = 118.79$, $p < .001$, η^2

$= .032$). Again, the model with the repeated measure effect for effort was the most likely,

compared to the null model (BF$_{10}$ > 1,000).

Figure 1

Work Context: Mean Differences in Moral Judgment by Effort and Gender
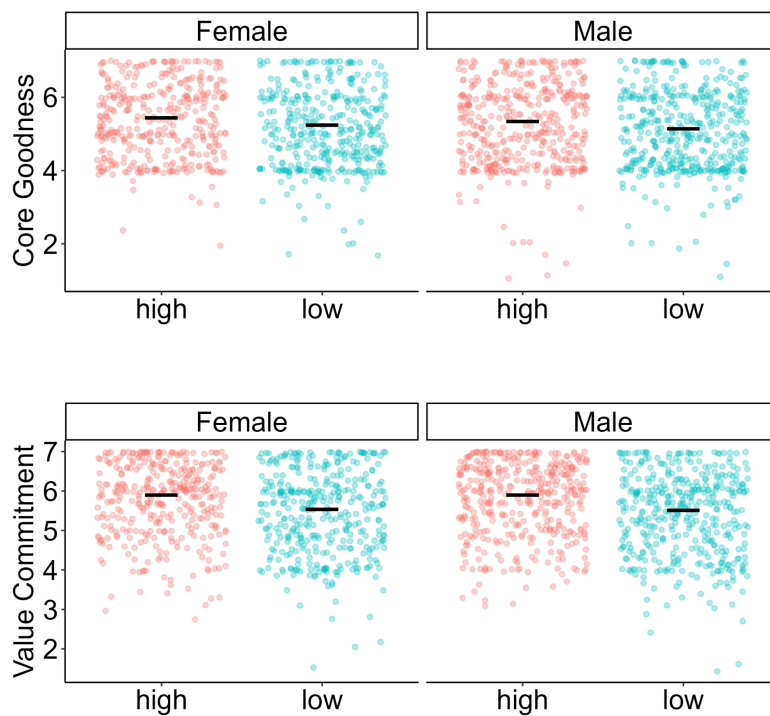
Study 2: care context

We observed similar findings as in the work context for core goodness, although the

effect was smaller. There was no significant main effect of gender ($F(1,699) = 3.37$, $p =$

$.067$, $\eta^2 = .004$) for males receiving higher ratings–and no significant interaction of gender

See me, judge me, pay me

and effort ($F(1,699) = 1.23$, $p = .269$, $\eta^2 < .001$). Again, the main effect of effort <mark>reached significance</mark> ($F(1,699) = 13.13$, $p < .001$, $\eta^2 = .002$). Compared to the null-model, the model with only effort as predictor received the strongest support ($BF_{10} = 30.96$).

For value commitment, no significant effects were found: neither the main effect of gender ($F(1,699) = 1.34$, $p = .247$, $\eta^2 = .001$), the interaction ($F(1,699) = 0.00$, $p = .946$, $\eta^2 < .001$), nor the main effect of effort ($F(1,699) = 0.25$, $p = .621$, $\eta^2 < .001$) reached significance. Strong evidence against the effort model (<mark>$BF10$ = 0.062</mark>) was observed in line with these results.

Figure 2

Care Context: Mean Differences in Moral Judgment by Effort and Gender
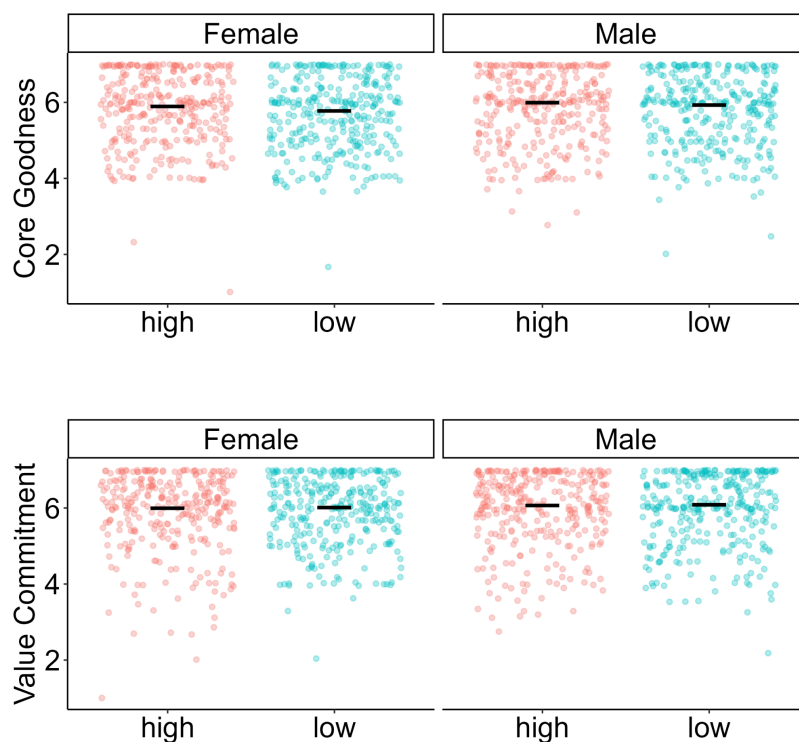
See me, judge me, pay me

Gender pay gap

Study 1: work context

The analysis did not reveal meaningful differences in pay deservingness between the effort conditions ($F(1, 857) = 1.20$, $p = .274$, $\eta^2 < .001$) or between male and female conditions ($F(1, 857) = 2.06$, $p = .151$, $\eta^2 = .002$), and there was no interaction between the factors ($F(1, 857) = 1.75$, $p = .186$, $\eta^2 < .001$).

Study 2: care context

We observed no differences in pay deservingness in the care context between the effort conditions ($F(1, 699) = 0.01$, $p = .933$, $\eta^2 < .001$) or between gender conditions ($F(1, 699) = 0.89$, $p = .346$, $\eta^2 < .001$). Also, no significant interaction was found ($F(1, 857) = 0.06$, $p = .807$, $\eta^2 < .001$).
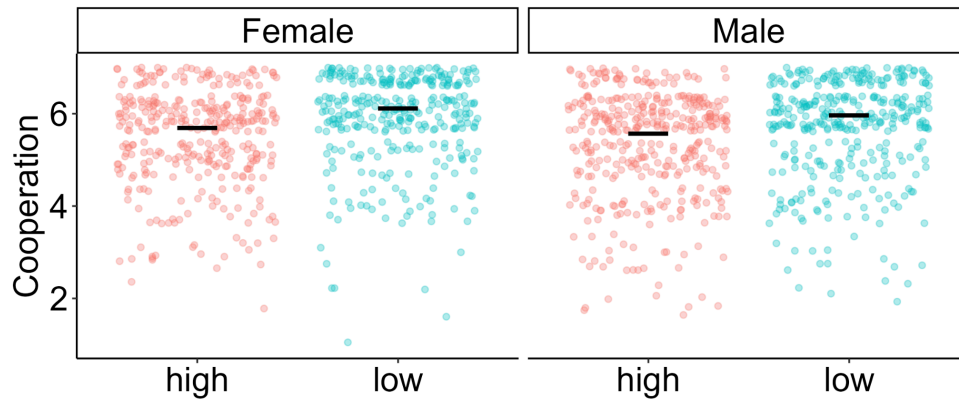
Aim 3: Cooperation partner satisfaction as a function of gender and effort

Study 1: work context

In the work context, we observed significant differences between low and high effort ($F(1,857) = 77.71$, $p < .001$, $\eta^2 = .030$), indicating higher satisfaction with the low-effort actor. Further, we observed a small gender difference ($F(1,857) = 4.57$, $p = .033$, $\eta^2 = .003$), indicating overall higher cooperation satisfaction with female cooperation partners. Yet, the interaction did not reach significance ($F(1,857) = 0.09$, $p = .760$, $\eta^2 < .001$). The model with the term for effort was the most likely mechanism behind the data ($BF_{10} > 1,000$), while the model with only gender as a predictor term did not point in an obvious direction ($BF_{10} = 0.622$).

21

See me, judge me, pay me

Figure 3

Work Context: Cooperation Partner Satisfaction

Study 2: care context

The higher average cooperation partner satisfaction with females was not observed in the care context ($F(1,699) = 2.17$, $p = .141$, $\eta^2 = .002$), but the difference between effort conditions was replicated ($F(1,699) = 116.84$, $p < .001$, $\eta^2 = .049$). Again, the interaction remained non-significant ($F(1,699) = 2.68$, $p = .102$, $\eta^2 < .001$). As for the work context, the effort model was the most likely data-generating mechanism ($BF_{10} > 1,000$), while moderate evidence against the gender model was observed ($BF_{10} = 0.221$).

Figure 4

Care Context: Cooperation Partner Satisfaction
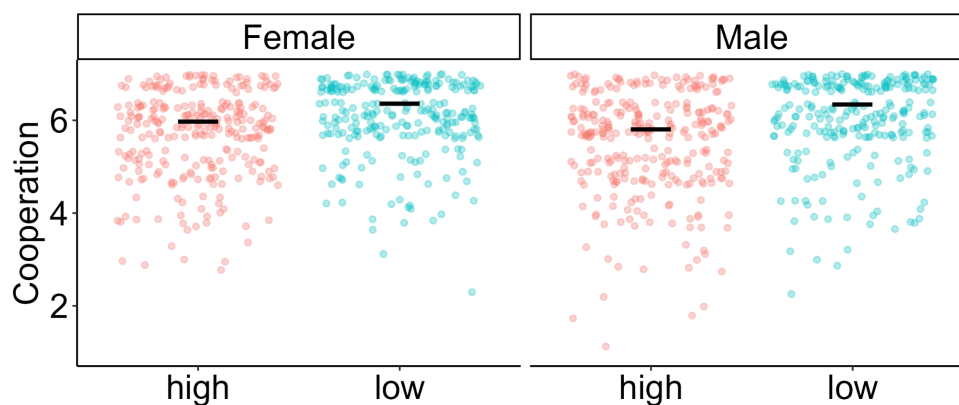
Planned Exploratory Analysis: Does gender norm endorsement moderate differences in effort moralization?

We explored whether the observed results on moral judgment could be a function of gender role beliefs (work: $M = 3.08$, $SD = 1.25$; care: $M = 2.99$, $SD = 1.26$). To do so, we tested whether differences in moral judgment were moderated by the interaction of gender and gender role beliefs. This was done by Study (work/care) and morality dimension (core goodness/value commitment).

The interaction didn't reach significance in either of the models. In both studies, we observed a small significant negative main effect of gender role beliefs on differences in value commitment (work: $\beta = -.17$, 95% CI [-.23, -.10], $p < .001$, $R^2_{adjusted} = .027$; care: $\beta = -.11$, 95% CI [-.18, -.04], $p = .004$, $R^2_{adjusted} = .008$) and for core goodness in Study 1(work: $\beta = -.08$, 95% CI [-.15, -.01], $p = .019$, $R^2_{adjusted} = .003$).

Critically, the strongest model explained below 3% of the variance in effort moralization. It is hence not likely that gender norm beliefs play a practically meaningful role, given that gender doesn't appear to meaningfully influence the effect either.

Exploratory Analysis after data collection

The present data led us to some data-driven post-hoc analyses, which we summarized below. Please note that the data is openly available for future secondary data analysis.

Interaction: participant gender and vignette gender

We computed exploratory interactions to assess the degree to which participants' gender influenced moral judgment by gender. We excluded participants who self-identified as neither female nor male (reduced samples: Study 1: $n_{work} = 849$, Study 2: $n_{care} = 693$).

In three of the models, no significant main or interaction effect was observed. Within the work context (Study 1), we observed a very small interaction, indicating a slightly stronger moralization effect in the opposite gender, which was pronounced stronger for

women (β = -.08, 95% CI [-.14, -.01], p = .025, $R^2_{adjusted}$ = .003). Hence, in the work context, women moralized the difference between high and low effort a bit stronger when evaluating men (< 1% in variance explained). Given the size of the effect, this doesn't seem to hold practical relevance.

No significant effect of participant gender was observed on differences in cooperation satisfaction, highlighting no apparent gender bias towards this domain in our data.

Discussion

Summary

The present project investigated three objectives. Firstly, it sought to test the replicability of the effort moralization effect in the work context (Study 1) and to explore its generalizability to the care context (Study 2). Secondly, it examined how the effort moralization effect interacts with gender across both contexts. Thirdly, it investigated whether cooperation partner satisfaction differs by gender and effort level. Study 1 was situated in a work context (office tasks), Study 2 in a private care situation (caring for elderly parents). While we replicated the effect in the work context, we observed mixed findings and smaller effects in the care context. Contrary to prior findings, lower effort was associated with higher cooperation partner satisfaction.

Replication in the work context

We replicated most of the effects from previous studies. The actor who exerted more effort was considered more moral. Compared to other studies, the respective effect sizes fall at the lower end. The effects were smaller than in previous studies conducted in the U.S. and were more in line with findings from Germany or France (Celniker et al., 2023, Studies 2a, 2c; Tissot & Roth, 2025). As anticipated, the high-effort actor was perceived as warmer and

less competent than the low-effort actor. In contrast to prior research, no significant results concerning pay deservingness were obtained.

An unexpected effect was observed concerning cooperation partner satisfaction: participants indicated they would be more satisfied working with the low-effort actor.

Extension to the care context

In the second study, the findings demonstrated greater variability. A significant effect was observed for core goodness, aligning with our expectations. However, the effect size was only half as large as in the work context. No significant effects were identified for value commitment and warmth.

One potential explanation is the presence of ceiling effects. Mean ratings were notably higher in the care context than in the work context (except for pay deservingness). This suggests that participants generally evaluated the actors more positively in the care setting, regardless of effort level. Notably, such effects did not emerge in previous research involving effortful behaviors with moral consequences. Celniker et al. (2023, Study 5) examined the relationship between donation levels and the distance of a respective run in fundraising events. Similarly, in a study by Bigman and Tamir (2016), participants engaged in either high- or low-effort behavior (solving math puzzles) which was rewarded with charitable donations. This situation was described to a new set of participants, who then indicated the level of reward the previous participants deserved. The findings indicated higher rewards for more effortful behavior. A fundamental distinction between these studies and the present study was that the past behaviors were fairly neutral regarding moral value—solving puzzles or running—whereas the current behavior—helping elderly parents—can be inherently described as moral. Disentangling this relationship between effort and the intrinsic morality of behavior is encouraged in future research.

In line with our expectations, the low-effort actor was perceived as more competent. No significant results were found for pay deservingness, mirroring the results of the work context study. We replicated our unexpected results regarding cooperation satisfaction: Participants reported greater satisfaction with the low-effort actor.

Interactions with gender across different contexts

Across both studies, no significant interactions emerged between effort moralization and the actor's gender. The only gender-related finding was that, in the work context, female actors received a higher mean rating as cooperation partners. Yet, male and female actors were not directly compared. This difference can hence not be interpreted as a meaningful difference.

Exploratory interactions of gender in the vignette and rater gender yielded no meaningful results. Likewise, gender role beliefs as a moderator remained non-significant. The absence of these findings is also in line with the absence of meaningful differences in effort manipulation between genders. Overall, these findings suggest that the moralization of effort applies independently of gender.

Theoretical implications and directions for future research

Our findings add to a growing body of research supporting the demographic generalizability of the effort moralization effect. Previous studies have shown that it holds across age groups (Tissot & Roth, 2025) and cultures (Bigman & Tamir, 2016; Celniker et al., 2023; Tissot & Roth, 2025), and our findings extend this pattern by demonstrating its robustness across gender.

However, our results do not fully support situational generalizability, as the findings were mixed in the care context. Potentially, the effect may be stronger in situations that provide fewer intrinsic moral cues, making heuristics more important for character judgment. Hence, inferring morality via effort as proxy might be less important in care work, which

might be seen as inherently moral behavior itself. In case the effort moralization effect is indeed a function of observability of moral behavior, future studies could systematically vary the degree of morality in given behaviors.

Finally, we made unexpected observations on cooperation satisfaction. While we expected higher satisfaction with the high-effort actor, our findings revealed the opposite effect, with participants reporting higher satisfaction with the low-effort actor. This did not align with previous studies on effort moralization, where the high-effort actor (rated as less competent) was chosen as the preferred cooperation partner (Celniker et al., 2023, Studies 4 & 6). These diverging findings might originate from differences in situative framing. While the prior scenarios relied on trust-based cooperation tasks, the current scenarios featured competence-related tasks, given their focus on common goal approaches. This resonates with recent research, indicating that cooperation partner preference depends on task affordances (e.g., trust- or competence-focused) (Matej Hrkalovic et al., 2025).

A second difference concerns the setting of the cooperation task. Whereas Celniker et al. (2023) measured partner choice in a domain-general trust game after showing the agents' effort in another domain (sports), we assessed partner satisfaction in the same domain as the prior behavior (work or care). Future research is encouraged to explicitly test the influence of differing task affordances and whether an alignment between effort and task contexts influences cooperation partner choice or satisfaction.

Limitations

To avoid stereotypical responses, we used vignettes that were not strongly associated with traditional gender roles. Hence, the work scenario was situated in an office setting, and the care scenario depicted household chores such as grocery shopping and laundry rather than emotional caregiving. This has potentially suppressed gender effects, which could have been observed in more stereotypical situations.

Further, we only studied within-gender effects and compared these between genders. Hence, no direct comparison between genders was performed by participants.

We observed unexpected findings regarding cooperation satisfaction, with higher satisfaction levels observed for the low-effort actor. These discrepancies may stem from methodological differences with earlier research. In contrast to the present study, participants in prior research (Celniker et al., 2023) were asked to select a preferred partner, as opposed to being assigned one. Since the implications of partner assignment versus partner choice on partner preference may diverge subsequent research is needed to further explore these varying dynamics.

A considerable difference in pay deservingness between the work and care contexts, with lower pay in the care context, was observed. While an immediate interpretation of this finding could be that it reflects the social reality of systematic undervaluation of care work (Antonopoulos, 2008; Charmes, 2019), it is also likely that the finding is a result of the methodological choices employed. In Study 1 (work), salaries were assessed, providing an industry benchmark (24$), while in Study 2 (care), no reference point was provided. This methodological difference most likely affected the results, as anchoring effects are known to influence numerical estimations (Tversky & Kahneman, 1974).

Conclusion

This research replicated and extended the effort moralization effect, demonstrating its robustness in the work context and generalizability across genders. However, the effect appears to be context-dependent, as results were mixed in the care context. This suggests that effort serves as a stronger moral signal in situations where moral character judgments need to be inferred from cues (i.e. effort) and cannot be inferred from behavior itself.

Notably, no differences between genders were observed in both studies, suggesting that the effect is generalizable across genders. Contrary to previous studies, participants did

not differentiate in their suggested payment based on exerted effort and demonstrated greater satisfaction with a collaboration partner who exerted less effort. While further research is needed to explore the contextual and situative boundary conditions, the present results provide further support for the effect as a mostly robust bias in moral judgment.

Competing interests

The authors report no conflict of interest.

Ethic approvement

The study was approved by the Departmental Review Board (DRB) of the Faculty of Psychology, Department of Occupational, Economic, and Social Psychology, University of Vienna (2024/M/009).

References

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

Antonopoulos, R. (2008). The Unpaid Care Work-Paid Work Connection. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1176661

Aronson, E., Wilson, T. D., Akert, R. M., & Sommers, S. R. (2021). *Social psychology* (Tenth edition. Global edition). Pearson Education Limited.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249. https://doi.org/10.1038/nature07538

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

Brambilla, M., & Leach, C. W. (2014). On the Importance of Being Moral: The Distinctive Role of Morality in Social Judgment. *Social Cognition*, *32*(4), 397–408. https://doi.org/10.1521/soco.2014.32.4.397

Brown, M. J., & Gladstone, N. (2012). Development of a Short Version of the Gender Role Beliefs Scale. *International Journal of Psychology and Behavioral Sciences*, *2*(5), 154–158. https://doi.org/10.5923/j.ijpbs.20120205.05

Celniker, J. B., Gregory, A., Koo, H. J., Piff, P. K., Ditto, P. H., & Shariff, A. F. (2023). The moralization of effort. *Journal of Experimental Psychology: General*, *152*(1), 60–79. https://doi.org/10.1037/xge0001259

Charmes, J. (2019). *The unpaid care work and the labour market: An analysis of time use*

data based on the latest world compilation of time-use surveys. ILO.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Dean, L., Churchill, B., & Ruppanner, L. (2021). The mental load: Building a deeper theoretical understanding of how cognitive and emotional labor over *load* women and mothers. *Community, Work & Family*, *25*(1), 13–29. https://doi.org/10.1080/13668803.2021.2002813

Deutsch, F. M., & Saxon, S. E. (1998). The Double Standard of Praise and Criticism for Mothers and Fathers. *Psychology of Women Quarterly*, *22*(4), 665–683. https://doi.org/10.1111/j.1471-6402.1998.tb00184.x

ERI. (2025). *Order Clerk Salary*. Economic Research Institute. https://www.erieri.com/salary/job/order-clerk

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2. ed). McGraw-Hill.

Fox, E., Russo, R., & Dutton, K. (2002). Attentional Bias for Threat: Evidence for Delayed Disengagement from Emotional Faces. *Cognition & Emotion*, *16*(3), 355. https://doi.org/10.1080/02699930143000527

Fwu, B., Wei, C.-F., Chen, S.-W., & Wang, H. (2014). Effort counts: The moral significance of effort in the patterns of credit assignment on math learning in the Confucian cultural context. *International Journal of Educational Development*, *39*, 157–162.

https://doi.org/10.1016/j.ijedudev.2014.07.010

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148–168. https://doi.org/10.1037/a0034726

Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The Multiple Dimensions of Gender Stereotypes: A Current Look at Men's and Women's Characterizations of Others and Themselves. *Frontiers in Psychology*, *10*, 11. https://doi.org/10.3389/fpsyg.2019.00011

JASP Team. (2020). *JASP* (Version 0.14) [Computer software].

Johnson, S. G. B., & Park, S. Y. (2021). Moral signaling through donations of money and time. *Organizational Behavior and Human Decision Processes*, *165*, 183–196. https://doi.org/10.1016/j.obhdp.2021.05.004

Kennedy, J., McDonnell, M.-H., & Stephens, N. (2016). Does Gender Raise the Ethical Bar? Exploring the Punishment of Ethical Violations at Work. *Academy of Management Proceedings*, *2016*(1), 11664. https://doi.org/10.5465/ambpp.2016.11664abstract

Lakens, D. (2022). Sample Size Justification. *Collabra: Psychology*, *8*(1), 33267. https://doi.org/10.1525/collabra.33267

Mason, J., & Robbins, K. G. (2024, June 27). *Americans' Unpaid Caregiving is Worth More than $1 Trillion Annually – and Women are Doing Two-Thirds of The Work*. National Partnership for Women & Families. https://nationalpartnership.org/americans-unpaid-caregiving-worth-1-trillion-annually -women-two-thirds-work/

Matej Hrkalovic, T., Li, A., Boop, M., Li, Y., & Balliet, D. (2025). Task affordances affect partner preferences. *Journal of Experimental Social Psychology*, *119*, 104751. https://doi.org/10.1016/j.jesp.2025.104751

Moss-Racusin, C., Phelan, J., & Rudman, L. (2010). When Men Break the Gender Rules: Status Incongruity and Backlash Against Modest Men. *Psychology of Men & Masculinity*, *11*, 140–151. https://doi.org/10.1037/a0018093

National Partnership for Women & Families. (2024). *New Analysis Shows Unpaid Care Work in the U.S. is Worth More Than $1 Trillion Each Year*. National Partnership for Women & Families. https://nationalpartnership.org/news_post/new-analysis-americans-unpaid-care-work-worth-more-than-1-trillion-each-year/

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

Reed, A., Aquino, K., & Levy, E. (2007). Moral Identity and Judgments of Charitable Behaviors. *Journal of Marketing*, *71*(1), 178–193. https://doi.org/10.1509/jmkg.71.1.178

Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, *74*(3), 629–645. https://doi.org/10.1037/0022-3514.74.3.629

Samtleben, C., & Müller, K.-U. (2022). Care and careers: Gender (in)equality in unpaid care, housework and employment. *Research in Social Stratification and Mobility*, *77*, 100659. https://doi.org/10.1016/j.rssm.2021.100659

Sayer, L. C. (2016). Trends in Women's and Men's Time Use, 1965–2012: Back to the Future? In S. M. McHale, V. King, J. Van Hook, & A. Booth (Eds.), *Gender and Couple Relationships* (Vol. 6, pp. 43–77). Springer International Publishing. https://doi.org/10.1007/978-3-319-21635-5_2

Sommet, N., Weissman, D. L., Cheutin, N., & Elliot, A. J. (2023). How Many Participants

Do I Need to Test an Interaction? Conducting an Appropriate Power Analysis and

Achieving Sufficient Power to Detect an Interaction. *Advances in Methods and*

Practices in Psychological Science, *6*(3), 25152459231178728.

https://doi.org/10.1177/25152459231178728

Sterling, J. S., & Reichman, N. (2016). Overlooked and Undervalued: Women in Private Law

Practice. *Annual Review of Law and Social Science*, *12*, 373–393.

https://doi.org/10.1146/annurev-lawsocsci-120814-121705

Tissot, T. T., & Roth, L. H. O. (2025). Is It Worth the Hustle? A Multi-Country Replication of

the Effort Moralization Effect and an Extension to Generational Differences in the

Appreciation of Effort. *Collabra: Psychology*, *11*(1), 132411.

https://doi.org/10.1525/collabra.132411

Toossi, M., & Morisi, T. L. (2017). *BLS Spotlight on Statistics: Women in the Workforce*

Before, During, and after the Great Recession. https://hdl.handle.net/1813/78334

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases:

Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*,

185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Wojciszke, B. (2005). Morality and competence in person- and self-perception. *European*

Review of Social Psychology, *16*(1), 155–188.

https://doi.org/10.1080/10463280500229619