

# MASTERARBEIT | MASTER'S THESIS

Titel | Title

The Moralization of Effort: Gendered Judgments in Work and  
Care Contexts

verfasst von | submitted by

Thea Marlene Fischer BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of  
Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree  
programme code as it appears on the  
student record sheet:

UA 066 840

Studienrichtung lt. Studienblatt | Degree  
programme as it appears on the student  
record sheet:

Masterstudium Psychologie

Betreut von | Supervisor:

Univ.-Prof. Dr. Veronika Job Sutnar

## Abstract

*Introduction.* Effort plays a significant role in shaping moral judgments since it is often seen as a reflection on how we evaluate individuals' character and moral integrity. Higher effort is associated with higher morality and preferred cooperation partner choice, even if higher effort does not lead to better performance. While previous research found this effect in work contexts, the role of effort moralization remains unexplored in care work and particularly gender disparities. This study investigates gender differences in the effort moralization effect across work and care contexts and how societal expectations shape moral evaluations.

*Method.* Two studies ( $N_{\text{total}} = 1,560$ ) with a mixed design were conducted (study 1 ( $N_{\text{work}} = 859$ ): work and study 2 ( $N_{\text{care}} = 701$ ): care). Participants read about either two men or two women (between-subjects factor) exerting identical tasks with the same outcome at different effort levels (high vs. low effort, within-subjects factor). Participants rated them on moral character, cooperation partner satisfaction, pay deservingness and gender role beliefs.

*Results.* The work context replicated the effort moralization effect whereas results were smaller for the care context. No gender-specific results were found, suggesting a robust effect across gender. Contrary to prior research, participants indicated greater satisfaction with the low-effort target as a cooperation partner.

*Conclusion.* Effort moralization is a robust, but context sensitive mechanism of social evaluation. Future research should further investigate how different contexts such as task requirements and emotional factors shape moral evaluations and cooperation partner choice.

This Master's thesis is part of a registered report project "See me, judge me, pay me", (see Open Science Framework: <https://osf.io/s8ec5/>).

*Keywords:* *effort moralization, gender stereotypes, work, care, moral judgment*

## **Zusammenfassung**

*Einleitung.* Anstrengung spielt eine wichtige Rolle bei der Bildung moralischer Urteile, da sie oft als Ausdruck von Charakter und moralischer Integrität gilt. Höhere Anstrengung wird mit stärkerer moralischer Bewertung sowie einer bevorzugten Wahl als Kooperationspartner:in assoziiert, selbst bei gleichem Ergebnis. Dieser Effekt wurde bereits in beruflichen Kontexten nachgewiesen, jedoch bleibt unklar, wie Anstrengung in der Care-Arbeit moralisch bewertet wird, insbesondere in Bezug auf Geschlechtsunterschiede. Die vorliegende Studie untersucht geschlechtsspezifische Unterschiede in der Moralisierung von Anstrengung in Arbeits- und Care-Kontexten sowie den Einfluss gesellschaftlicher Erwartungen.

*Methoden.* Es wurden zwei Studien ( $N_{\text{total}} = 1.560$ ) mit einem gemischten Versuchsdesign durchgeführt (Studie 1 ( $N_{\text{work}} = 859$ ): Arbeit; Studie 2 ( $N_{\text{care}} = 701$ ): Care). Die Teilnehmer:innen lasen Vignetten über entweder zwei Männer oder zwei Frauen (Zwischensubjektfaktor), die bei gleicher Aufgabe und gleichem Ergebnis unterschiedlich viel Anstrengung zeigten (Innersubjektfaktor: hoch vs. gering). Anschließend bewerteten sie die Zielperson hinsichtlich ihres moralischen Charakters, der Zufriedenheit als Kooperationspartner:in, der Entlohnung sowie der eigenen Einstellung zu Geschlechterrollen.

*Ergebnisse.* Im Arbeitskontext wurde der Moralisierungseffekt von Anstrengung repliziert, während er im Care-Kontext schwächer ausgeprägt war. Es wurden keine geschlechtsspezifischen Unterschiede gefunden, was auf einen robusten geschlechtsunabhängigen Effekt hindeutet. Im Gegensatz zu früheren Studien zeigte sich eine höhere Zufriedenheit mit derjenigen Person als Kooperationspartner:in, die weniger Aufwand zeigte.

*Schlussfolgerung.* Die Moralisierung von Aufwand ist ein robuster, aber kontextabhängiger Mechanismus sozialer Bewertung. Zukünftige Forschung sollte untersuchen, wie verschiedene Kontexte, Aufgabenanforderungen sowie emotionale Faktoren moralische Urteile beeinflussen.

Diese Masterarbeit ist Teil des Registered-Reports-Projekts “See me, judge me, pay me”, (siehe Open Science Framework: <https://osf.io/s8ec5/>).

Schlagwörter: *Moralisierung von Anstrengung, Geschlechtsstereotype, Urteile*

## Introduction

Assessing the character of unfamiliar individuals is a fundamental behavior that shapes social interactions – influencing decisions about trust, cooperation and moral evaluation in personal and professional contexts (e.g., meeting new colleagues). Since character features are not directly observable, we must rely on limited information and heuristics. For example, when meeting a new colleague, one must quickly evaluate their credibility and accessibility. For this purpose, we use simple processing mechanisms to arrive at judgments (Fiske & Taylor, 1991). Mental shortcuts including observable cues such as facial expressions (Fox et al., 2002), stereotypes (Aronson et al., 2021), heuristics and cognitive biases (Tversky & Kahneman, 1974) as well as our past personal learning experiences (Behrens et al., 2008) enable us to make rapid judgments in complex social environments.

### The effort moralization effect

Moral information plays a central role among the cues shaping the perception of other people's character (Wojciszke, 2005; Goodwin et al., 2014; Brambilla & Leach, 2014). Morality refers to a system of rules that allow individuals to distinguish right from wrong and guide human behavior in social interactions, involving judgments about fairness, harm and virtue (Ellemers et al., 2019; Gert, 2005).

The effort moralization effect is one psychological phenomenon that has gained attention in this context (Fwu et al., 2014; Bigman & Tamir, 2016; Amos et al., 2019; Celniker et al., 2023). It describes how people judge other people's moral character based on the visible effort they put into a specific task or behavior. Thus, the more effort a person invests in a specific task, the stronger the moral evaluations. Actions that are seen to be “good” become even more praiseworthy whereas actions perceived as “bad” are judged more harshly when more effort is exerted (Bigman & Tamir, 2016).

Importantly, the effect persists even if the increased effort does not lead to better performance (Celniker et al., 2023). This highlights the interpretation that effort itself is seen as valuable, regardless of its practical efficacy. The effect has been well replicated, yet studies show a variation in effect sizes between different cultures. For instance, in France ( $d = 0.38$ ), South Korea ( $d = 0.71$ ), the United States of America ( $d = 0.60$ ), Germany ( $d = .34-.37$ ) and Mexico ( $d = .14-.28$ ) (Celniker et al., 2023; Tissot & Roth, 2024).

Beyond shaping moral perceptions, the invested effort level impacts cooperation partner choice. Even if the performance of two individuals is equal, those individuals investing higher effort are more likely to be selected as a cooperation partner in a follow up trust game (Celniker et al., 2023). Although Celniker et al. (2023) have shown that those individuals performing high effort are frequently seen as less competent, they are still preferred as cooperation partners. This dynamic has important implications for professional and career-related contexts. Employees who always show high effort, might be preferred for promotions, management tasks, or cooperative customer projects, even though their performance is equal to other employees.

### **Blind spots in the literature of effort moralization**

Prior studies have predominantly focused on male or gender-neutral agents (e.g. “Person A”) in their vignettes, excluding females from their scenarios. Therefore, investigating and understanding gender differences in effort moralization is firstly important for the generalizability of the effect and secondly essential to detect potential gender biases for addressing inequalities such as reinforcing traditional gender roles.

Moreover, most studies have primarily focused on two specific contexts where effort moralization comes into effect: work contexts (Amos et al., 2019; Celniker et al., 2023) and charitable or helping behaviors (Bigman & Tamir, 2016; Celniker et al., 2023). These settings are important because of their daily presence and requested effort, but they overlook the immense contribution of unpaid care work and household duties. With responsibilities disproportionately affecting women (Mason & Robbins, 2024), this area remains unnoticed in prior research. While the average American citizen spends approximately 245 hours per year on unpaid care work, nearly two-thirds (65%) of that time is spent by women (Mason & Robbins, 2024). Whilst its substantial economic value is more than \$1 trillion annually in the US (National Partnership for Women & Families, 2024), unpaid care work tends to receive minimal societal recognition (Antonopoulos, 2008) and is associated with high cognitive and emotional strain (Dean et al., 2021).

Additionally, in the context of effort and cooperation, our study investigates how perceived effort influences individuals’ satisfaction with assigned partners to reflect real-world scenarios where cooperation is mostly determined by external structures rather than personal decisions. Celniker et al. (2023) found that people who put more effort into achieving the same performance are preferred as cooperation partners. Yet, in real world scenarios people do not

often have the opportunity of freely choosing their cooperation partner. Thus, this study addresses a gap in literature and provides insights into cooperative dynamics in professional settings, where cooperation is mostly determined by organizational hierarchies and project assignments and not by personal decisions.

### **Gender stereotypes in moral judgments across work and care contexts**

Social judgments are shaped by stereotypes and gender biases, which set expectations on how women and men should behave (Cislaghi & Heise, 2019). Traditional societal norms ascribe traits like assertiveness, independence, and competence to men, whereas women are associated with qualities such as warmth, nurturing, and cooperation (Abele et al., 2008; Hentschel et al., 2019). These expectations may influence how individuals perceive, acknowledge and penalise effort in different contexts. As a result, similar behaviors can be interpreted differently depending on gender (Rudman, 1998).

Building on the "Big Two of Personality" framework by Bakan (1966), differentiating between communion and agency, gender stereotypes align with these dimensions (Haines et al., 2016): women are generally viewed as communal (caring, compassionate), while men are seen as agentic (independence, self-confidence) (Abele et al., 2008; Hentschel et al., 2019). These stereotypical dimensions are likely to shape moral judgments in both work and care contexts.

### ***Effort moralization in work contexts***

In workplace settings, when agentic men show more effort it may be seen as an extension of their natural drive, personal advancement and self-assertion, whereas for communal women, it might be interpreted as an expression of their commitment to cooperation, community and connection. Since morality is considered as a core character trait for women, it becomes a key standard by which their effort is evaluated. As a result, women displaying high effort at work, might be judged more positively in terms of moral character.

This heightened moral expectation however, can have negative consequences. When certain behavior deviates from a traditional norm or stereotype, it may backlash unfolding through social and economic sanctions (Rudman, 1998). Research shows that women experience harder punishments for ethical violations compared to men (Kennedy et al., 2016; Rudman, 1998), while men are criticized more strongly for non-agentic behavior, particularly

in leadership roles (Moss-Racusin et al., 2010). Consequently, women may face greater criticism for low effort at work as to their male counterparts.

### ***Effort moralization in care contexts***

Beyond professional settings, the gendered moralization of effort can extend to the care context. Caregiving roles are traditionally associated with women, enhancing the societal expectation that they should naturally invest in these areas (Kennedy et al., 2016). Historically, men are viewed as primary providers engaging in paid labor whereas women are linked to unpaid care work. Despite evolving gender roles as shown by the increasing participation of women in the U.S. labor market (Toossi & Morisi, 2017; Statista Research Department, 2024) as well as a greater contribution of men to family labor (Sayer, 2016), women still bear most amount of caregiving and household duties (Charmes, 2019).

Women's contributions in this area are frequently seen as fulfilling an expected role and that their effort is obligatory rather than exceptional, which diminishes their recognition and reward (Kennedy et al., 2016). In contrast, men investing high effort in care domains are often perceived as exceeding societal expectations and going above and beyond since it deviates from their traditional role, leading to praise for being involved (Deutsch & Saxon, 1998). For men, it is therefore not judged as harshly to demonstrate low effort as it is for women since females do not meet the expected care work standards when showing low effort (Deutsch & Saxon, 1998; Samtleben & Müller, 2022).

These previous findings suggest the potential differences in judgment based on gender, levels of effort and social context.

### **Cooperation partner choice**

As shown in prior research, perceived morality influences the choice of cooperation partners (Celniker et al., 2023). However, cooperation in real-life contexts is mostly determined by external conditions which is why this study investigates individual's satisfaction with their already assigned cooperation partners. Considering the presented studies above, understanding these dynamics is important as it may lead to unequal distributions of professional opportunities. For instance, women often receive less recognition although they show the same performance and effort as men and are less likely to be involved in important projects or collaborations (McKinsey & Company, 2022). Such biases also represent a potential blind spot in the study of gender stereotyping in the workplace.

## Current Study

The current study is embedded within the registered report project “See me, judge me, pay me: Gendered effort moralization in work and care” (<https://osf.io/s8ec5/>). It focuses on three aims: 1) To replicate the effort moralization effect in the work condition and test if it is generalizable to care contexts, 2) To explore gender-specific differences in the moralization of effort across work and care contexts, and 3) To investigate whether and how cooperation partner satisfaction differs by gender and effort.

## Hypotheses

As described above and presented in Table 1, this study pursues three key aims, from which the following hypotheses can be derived:

**Table 1**

### *Summary of Hypotheses*

| Research Aim                                                                                                       | Hypotheses                                                                                                                                                                                                                                                              |
|--------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Replicate the effort moralization effect in the work context and test its generalizability to the care context. | H1a: Individuals who dedicate higher effort in their work are judged higher in morality than individuals exerting lower effort.<br>H1b: Individuals who dedicate higher effort in their care work are judged higher in morality than individuals exerting lower effort. |
| 2. Explore gender-specific differences in effort moralization across work and care contexts.                       | H2a: In the work context, moral character judgments vary based on gender and effort level.<br>H2b: In the care context, moral character judgments vary based on gender and effort level.                                                                                |
| 3. Investigate cooperation partner satisfaction by gender and effort.                                              | H3a: Cooperation satisfaction is predicted by gender and effort in the work context.<br>H3b: Cooperation satisfaction is predicted by gender and effort in the care context.                                                                                            |

## Method

### Sample

Following the smallest effect size of interest (SESOI) approach (Lakens, 2022) to determine the required sample size for detecting a small effect (Cohen, 1988) in a 2×2 mixed ANOVA ( $\eta^2 = .01$ ,  $\alpha = .05$ ,  $1-\beta = .95$ ) using G\*Power 3.1.9.7 (Faul et al., 2009) and IntXPower (Sommet et al., 2023), a minimum sample size of  $N = 324$  was calculated for both study 1 (work) and study 2 (care), which was planned to be oversampled to  $N = 350$  per study to account for potential data loss. We aimed for a gender-balanced quota sample with 50% males and 50% females in each study. Detailed specifications of the simulations can be found in the supplemental material.

The sample consisted of US participants, recruited through Prolific (<https://www.prolific.com/>) for both studies. The platform ensures fair compensation (\$8.00/hour) and guarantees high data quality. Data collection took place from 29.01.2025 to 31.01.2025. The initial sample size was  $N_{\text{initial}} = 2,270$  with  $N_{\text{work}} = 1,142$  and  $N_{\text{care}} = 1,128$  as more participants could be recruited due to a grant scholarship in the amount of 3,000 euros. After applying our exclusion criteria (see data cleaning), we eliminated  $N = 710$  participants leading to a total sample size of  $N_{\text{total}} = 1,560$ . Thus, we had a final sample size of  $N_{\text{work}} = 859$  for study 1 within the work context and  $N_{\text{care}} = 701$  for study 2 in the care context.

The total mean age was  $M_{\text{total}} = 42.32$  years ( $SD = 13.68$ ) with an average age in study 1 (work context) of  $M_{\text{work}} = 43.4$  years ( $SD = 13.76$ ) ranging from 18 to 81. In study 2 (care context) the mean age was  $M_{\text{care}} = 41.0$  years ( $SD = 13.49$ ) in the care context with ages ranging from 18 to 86 years, as presented in Table 2.

**Table 2**

*Descriptives by Study*

|               | <b>N</b> | <b>age: M (SD)</b> | <b>range</b> | <b>male</b> | <b>female</b> |
|---------------|----------|--------------------|--------------|-------------|---------------|
| overall       | 1,560    | 42.32 (13.68)      | 18-86        | 800         | 742           |
| Study 1: work | 859      | 43.43 (13.75)      | 18-81        | 429         | 420           |
| Study 2: care | 701      | 40.97 (13.49)      | 18-86        | 371         | 322           |

*Note.* Participant counts include “self-described as”, “prefer not to say” and missing values (work: 9 self-described, 1 prefer not to say, 10 missing; care: 6, 2, and 8, respectively).

In terms of participant gender, the sample is balanced with overall 47.56% of female participants.

## **Procedure**

Two independent studies (study 1: work and study 2: care) were conducted, with participants who participated in one study excluded from participating in the other study. The decision for two separate studies is based on technical and organisational reasons concerning participant management and data collection logistics on Prolific as well as to ensure balanced gender distributions within each context. In both studies, participants first read the instructions and then provided informed consent for the study procedure. Following this, they were presented with a vignette – either about a work or care context. The vignettes involve two individuals, randomly presented as male or female, performing identical tasks at the same quality level. However, the effort differs in their required amount. The work vignette describes the two individuals who are processing company orders with identical productivity and quality, but who differ in the level of effort they invest. The care vignette depicts the individuals as twins assisting their elderly parents with household chores. These tasks such as grocery shopping and cooking were deliberately chosen to avoid emotionally laden elements such as child care as this might be biased by gender. For instance, the vignette outlining the work context states:

*Michael and Lukas work at the same company and process similar orders in the company's office. Both Michael and Lukas 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 Michael's orders and 96% of Lukas's orders are error-free and complete. On average, Michael and Lukas each process correct orders worth \$144 per hour.*

*For Michael, processing orders requires minimal effort — although he works as quickly as possible, he finds the work easy.*

*For Lukas, however, processing orders requires a lot of effort — although he also works as quickly as possible, he finds the work hard.*

Following the scenario presentation, participants completed a series of measures assessing each individual in a randomized order. The measures are presented in Table 3.

Afterwards, they responded to a series of demographic questions such as age, gender and nationality. Both studies lasted approximately seven minutes.

## **Measures**

Established measurement instruments and vignettes were adapted and refined for the current study (Celniker et al., 2023). Celniker et al.'s (2023) work vignette was revised to include less stereotypical tasks (the factory scenario was changed to an office scenario) to mitigate possible gender biases for traditional male or female roles and to establish a more contemporary scene. In addition, we designed a new vignette for the care context.

The study assesses perceived morality using 13 trait items from prior research (Celniker et al., 2023) which differentiates between two types of moral virtues: core goodness traits (or unconditionally moral traits, e.g., “Honest” or “Trustworthy”), and value commitment traits (or conditionally moral traits, e.g., “Motivated” or “Dedicated”; Piazza et al., 2014). Each trait item is rated on a 7-point scale. To capture dimensions of social perception, warmth and competence are measured on a 7-point scale by Fiske et al. (2007). Manipulation checks include assessment of effort, quality, difficulty and work value, rated on a 7-point scale.

Pay deservingness was evaluated differently across the two study contexts as shown in Table 3. Salary benchmarks were informed by public data shared on the website of the ERI Economic Research Institute (<https://www.erieri.com>).

Cooperation partner choice differentiates between a future work project (work context) and organizing a charity event (care context), using a 7-point scale.

To explore potential moderating effects on effort moralization, a brief version of the Gender Role Belief Scale will be implemented to measure participants’ level of gender role endorsement (Brown & Gladstone, 2012). An overview of the measures can be found in Table 3 and detailed information in the supplemental material.

**Table 3***Measures Overview*

| <b>Construct (<i>n</i> items)</b>             | <b>Example item</b>                                                                                        | <b>Low anchor</b>        | <b>High anchor</b>         |
|-----------------------------------------------|------------------------------------------------------------------------------------------------------------|--------------------------|----------------------------|
| core goodness (6) <sup>a</sup>                | honest                                                                                                     | does not describe X well | describes X extremely well |
| value commitment (7) <sup>a</sup>             | dedicated                                                                                                  | does not describe X well | describes X extremely well |
| competence/warmth (2) <sup>a</sup>            | competent                                                                                                  | does not describe X well | describes X extremely well |
| effort (1) <sup>b</sup>                       | How much effort do you think X puts into his/her (care) work?                                              | no effort at all         | a lot of effort            |
| quality (1) <sup>c</sup>                      | In your opinion, how well does X perform his/her (care) work?                                              | very bad                 | very good                  |
| difficulty (1) <sup>c</sup>                   | Compared to other jobs/ care work, how difficult is X's (care) work.                                       | not at all difficult     | extremely difficult        |
| work value (1) <sup>c</sup>                   | How valuable do you think X's (care) work is?                                                              | not valuable at all      | extremely valuable         |
| pay deservingness (work) (1) <sup>a</sup>     | 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) <sup>a</sup>     | 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) <sup>a</sup> | [...] Please indicate how satisfied you would be to work with either X or Y?                               | extremely dissatisfied   | extremely satisfied        |
| gender role beliefs <sup>a</sup>              | Women with children should not work outside the home if they don't have to financially                     | strongly agree           | strongly disagree          |

*Note.* <sup>a</sup>These variables present the primary dependent measures, <sup>b</sup>This variable serves as manipulation check and exclusion criterion, <sup>c</sup>These variables serve as manipulation check but not as exclusion criteria.

## Data Cleaning

Participants with incomplete data sets as well as those missing one of two attention checks were excluded from the analysis. The probability for random guessing and therefore passing both attention checks was  $(\frac{1}{7} \times \frac{1}{7}) = 2.04\%$ . Following Celniker et al. (2023) and Tissot & Roth (2024), elimination also occurred for participants judging the low-effort condition to be just as or more effortful than the high-effort condition. Participants who finished the study three standard deviations (SD) faster than the average were excluded, because it may reflect a lack of sufficient attention to the task. No exclusion occurred for slow participation. As a result, after our initial data collection with 2,270 participants,  $N = 1,560$  remained for the final analysis (Roth et al., 2025).

## Data Analysis

To test the replication of the original effort moralization effect and its extension to the care context, we conducted several one-sided, dependent Welch's  $t$ -tests. Respective effect sizes (Cohen's  $d$ ) and Bayes Factors ( $BF_{10}$ ) were also computed. We compared the moral judgments as well as perceived warmth, competence, and pay deservingness of the targeted actors between high and low effort conditions.

To analyze moral character as a function of gender, context, and effort, we conducted a mixed effect ANOVA with gender serving as a between-subjects factor and effort as the within-subjects factor with an interaction term included. To assess the absence or presence of an effect, all Bayes Factors ( $BF_{10}$ ) were calculated.

Finally, cooperation partner satisfaction was analyzed using the same mixed-effects ANOVA as mentioned above.

## Results

### Reliabilities

The reliability analyses using Cronbach's alpha showed consistently high internal consistencies of the scales used. For the Core Goodness dimension, the scale showed an excellent reliability of  $\alpha = .93$  in the low-effort condition and a Cronbach's alpha of  $\alpha = .94$  in the high-effort condition. The Value Commitment dimension also showed excellent values in both conditions:  $\alpha = .94$  (low effort) and  $\alpha = .93$  (high effort). The scale for measuring gender role beliefs achieved a good reliability of  $\alpha = .88$  (Arof et al., 2018). Therefore, all scales fulfil the current recommendations for psychological survey instruments, which classify values of .80 and above as good and values above .90 as excellent (Arof et al., 2018).

### Research Aim 1: Replicate the effort moralization effect

Dependent one-sided Welch's sample t-tests were conducted to determine the effort moralization effect in the work context and its transmission to the care context. Participants' judgments of the low effort and the high effort condition were compared across the dimensions core goodness etc.

#### *Study 1: Work context*

In the work context, the effort moralization effect was replicated. Targets in the high-effort condition were judged higher in their core goodness than those in the low-effort condition, as seen in Table 4. Similar results were observed for value commitment and warmth, but with a smaller effect size. Concerning competence, targets in the low-effort condition were judged as to be more competent than targets in the high-effort condition. Cooperation was also rated higher for targets in the low-effort condition compared to high-effort targets. There was no significant difference between the conditions for pay deservingness.

**Table 4***Moral Character Judgment in the Work Context by Effort Level*

|                   | <b>Low<br/>Effort</b> | <b>High<br/>Effort</b> | <b>d</b> | <b>CI</b>      | <b>BF<sub>10</sub></b> | <b>p</b> | <b>Interpretation</b> |
|-------------------|-----------------------|------------------------|----------|----------------|------------------------|----------|-----------------------|
| Core Goodness     | 5.18<br>(1.13)        | 5.38<br>(1.13)         | -0.30    | [-0.37, -0.23] | >1,000                 | < .001   | Higher high effort    |
| Value Commitment  | 5.52<br>(1.08)        | 5.89<br>(0.95)         | -0.37    | [-0.44, -0.30] | >1,000                 | < .001   | Higher high effort    |
| Warmth            | 4.77<br>(1.38)        | 4.95<br>(1.36)         | -0.19    | [-0.26, -0.12] | >1,000                 | < .001   | Higher high effort    |
| Competence        | 6.17<br>(0.98)        | 5.48<br>(1.24)         | 0.53     | [0.46, 0.61]   | >1,000                 | < .001   | Higher low effort     |
| Pay Deservingness | 26.37<br>(3.36)       | 26.28<br>(3.38)        | 0.04     | [-0.03, 0.11]  | 0.076                  | 0.243    | No difference         |
| Cooperation       | 6.03<br>(1.11)        | 5.62<br>(1.19)         | 0.30     | [0.23, 0.37]   | >1,000                 | < .001   | Higher low effort     |

***Study 2: Care context***

In the care context, targets in the high-effort condition were also judged higher in their core goodness than those in the low-effort condition but half the effect size. The same pattern as in the work context was found for cooperation and competence. No significant differences were found for value commitment, warmth and pay deservingness. These findings are summarized in Table 5.

**Table 5***Moral Character Judgment in the Care Context by Effort Level*

|                      | <b>Low<br/>Effort</b> | <b>High<br/>Effort</b> | <b>d</b> | <b>CI</b>          | <b>BF<sub>10</sub></b> | <b>p</b> | <b>Interpretation</b> |
|----------------------|-----------------------|------------------------|----------|--------------------|------------------------|----------|-----------------------|
| Core Goodness        | 5.84<br>(0.97)        | 5.94<br>(0.96)         | -0.14    | [-0.21, -<br>0.06] | 24.90                  | < .001   | Higher high<br>effort |
| Value<br>Commitment  | 6.04<br>(0.93)        | 6.02<br>(1.00)         | 0.02     | [-0.06, 0.09]      | 0.048                  | 0.622    | No difference         |
| Warmth               | 5.61<br>(1.24)        | 5.63<br>(1.24)         | -0.02    | [-0.09, 0.06]      | 0.046                  | 0.671    | No difference         |
| Competence           | 6.26<br>(0.99)        | 5.57<br>(1.34)         | 0.49     | [0.42, 0.57]       | >1,000                 | < .001   | Higher low<br>effort  |
| Pay<br>Deservingness | 23.28<br>(8.29)       | 23.40<br>(8.20)        | -0.04    | [-0.11, 0.04]      | 0.066                  | 0.349    | No difference         |
| Cooperation          | 6.34<br>(0.92)        | 5.88<br>(1.11)         | 0.41     | [0.33, 0.49]       | >1,000                 | < .001   | Higher low<br>effort  |

**Gender role beliefs scale**

In order to gain an overview about participants' attitudes towards gender role beliefs, descriptive statistics were computed for the work and the care context. In each context, participants indicated a medium level of endorsement on a scale from 1 to 7. In the work context, participants indicated a mean value of  $M = 3.08$  ( $SD = 1.25$ ), ranging from 1 to 7. In the care context, the mean value was  $M = 2.99$  ( $SD = 1.26$ ), ranging from 1 to 6.9, presented in Table 6.

**Table 6***Descriptives by Study for Gender Role Beliefs*

|               | <b>N</b> | <b>M</b> | <b>SD</b> | <b>range</b> |
|---------------|----------|----------|-----------|--------------|
| Study 1: work | 859      | 3.08     | 1.25      | 1-7          |
| Study 2: care | 701      | 2.99     | 1.26      | 1-6.9        |

## Research Aim 2: Gender Differences in Effort Moralization

2×2 mixed ANOVAs (between-subjects factor: gender, within-subjects factor: effort) were conducted to test moral judgments between the target gender and the effort levels.

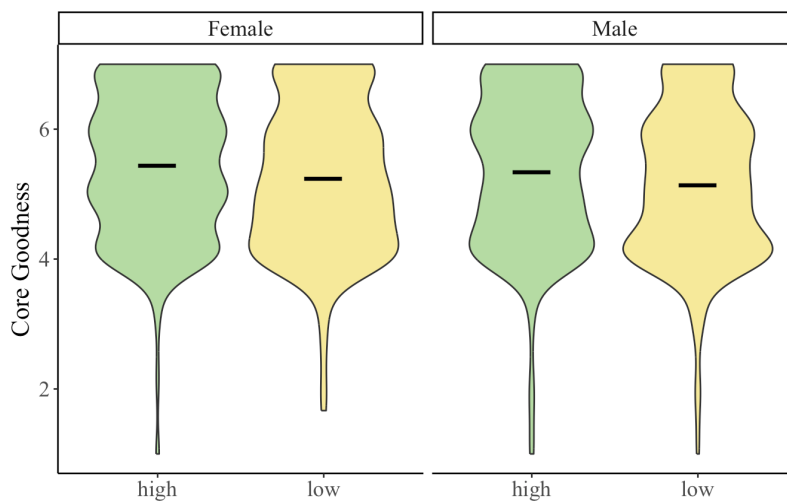
### *Study 1: Work context*

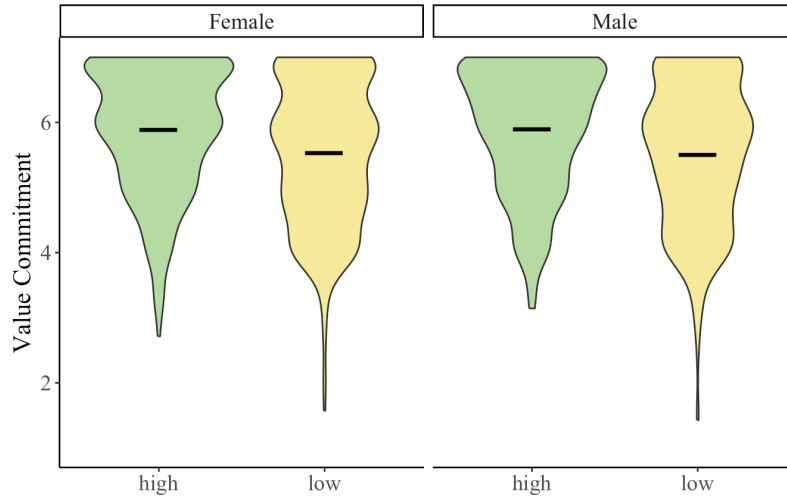
For core goodness, we found a significant main effect of effort ( $F(1,857) = 75.51, p < .001, \eta^2 = .008$ ) with higher moral ratings for the high-effort target as anticipated and seen in Figure 1. There was no significant main effect for gender ( $F(1,857) = 1.78, p = .182, \eta^2 = .002$ ) as well as no interaction of gender and effort. Post-hoc tests showed a significant effect of effort for both male ( $p < .0001$ ) and female targets ( $p < .0001$ ) with comparable results ( $\Delta M = 0.20$ ).

For value commitment, almost identical results we observed with a significant main effect of effort ( $F(1,857) = 118.79, p < .001, \eta^2 = .032$ ), but no main effect for gender ( $F(1,857) = 0.04, p = .850, \eta^2 < .001$ ) and no significant interaction ( $F(1,857) = 0.23, p = .631, \eta^2 < .001$ ).

**Figure 1**

*Differences in Moral Judgments Across Gender and Effort in the work context*





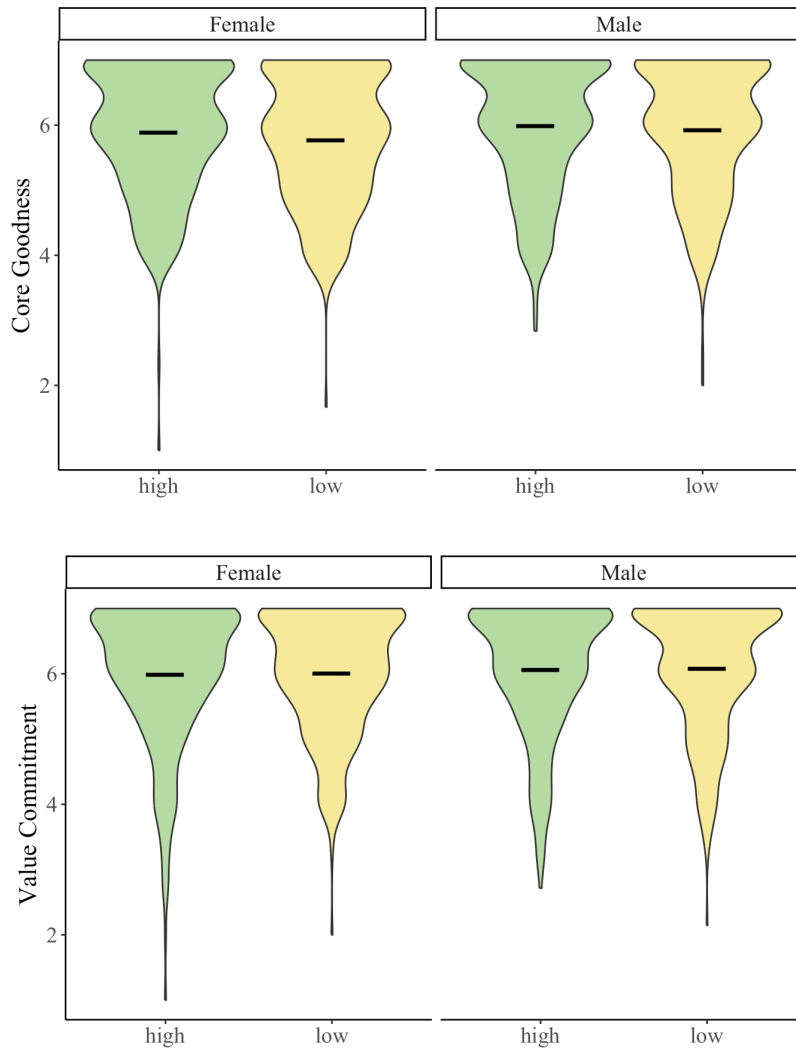
### ***Study 2: Care context***

In the care context, we found similar yet smaller effects on core goodness with a significant main effect for effort ( $F(1,699) = 13.13, p < .001, \eta^2 = .002$ ). We found a marginal tendency for the main effect of gender, yet no significant effect ( $F(1,699) = 3.37, p = .067, \eta^2 = .004$ ) with slightly higher ratings for male targets ( $M = 5.96$ ) than for female targets ( $M = 5.83$ ) as presented in Figure 2. Again, we found no significant interaction between gender and effort ( $F(1,699) = 1.23, p = .269, \eta^2 < .001$ ).

Surprisingly, there was no main effect for effort concerning value commitment ( $F(1,699) = 0.25, p = .621, \eta^2 < .001$ ) as well as no main effect of gender ( $F(1,699) = 1.34, p = .247, \eta^2 = .001$ ) and none for the interaction term ( $F(1,699) = 0.00, p = .946, \eta^2 < .001$ ).

**Figure 2**

*Differences in Moral Judgments Across Gender and Effort in the care context*



### **Pay deservingness**

In addition to the analyses mentioned above, we also explored whether the suggested pay differed in virtue of the target gender, effort level and context.

We found no significant main effect of gender ( $F(1,1556) = 0.19, p = .665, \eta^2 < .001$ ), which contradicts our expectations. Further, we found no significant interaction with gender, effort or context (work or care). However, we found that targets in the work context received higher suggested pay than in the care context, across both gender and effort conditions ( $F(1,1566) = 97.30, p < .001, \eta^2 = .056$ ).

### Research Aim 3: Cooperation partner satisfaction by gender and effort

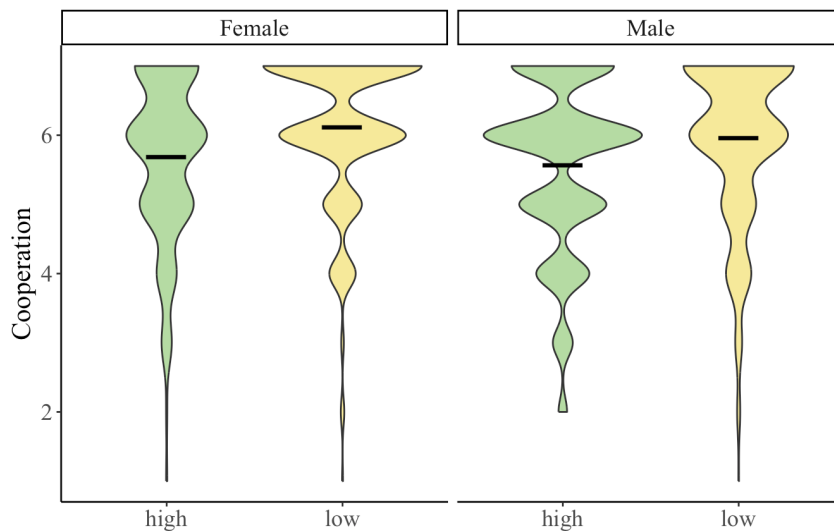
2×2 mixed ANOVAs with gender as between-subjects factor and effort as within-subjects factor were used to investigate if the satisfaction with an assigned cooperation partner was influenced by effort and target gender.

#### *Study 1: Work context*

In the work context, results revealed a significant main effect of effort ( $F(1,857) = 77.71, p < .001, \eta^2 = .030$ ) meaning low-effort targets were rated more positively (see Figure 3). This suggests that participants showed higher satisfaction with the low-effort target as an assigned cooperation partner which differs from prior research by Celniker et al. (2023). We found a significant main effect of gender ( $F(1,857) = 4.57, p = .033, \eta^2 = .003$ ), revealing a higher satisfaction to work with female targets as cooperation partners. The interaction was not significant ( $F(1,857) = 0.09, p = .760, \eta^2 < .001$ ).

**Figure 3**

*Cooperation Partner Satisfaction Across Gender and Effort in the work context*

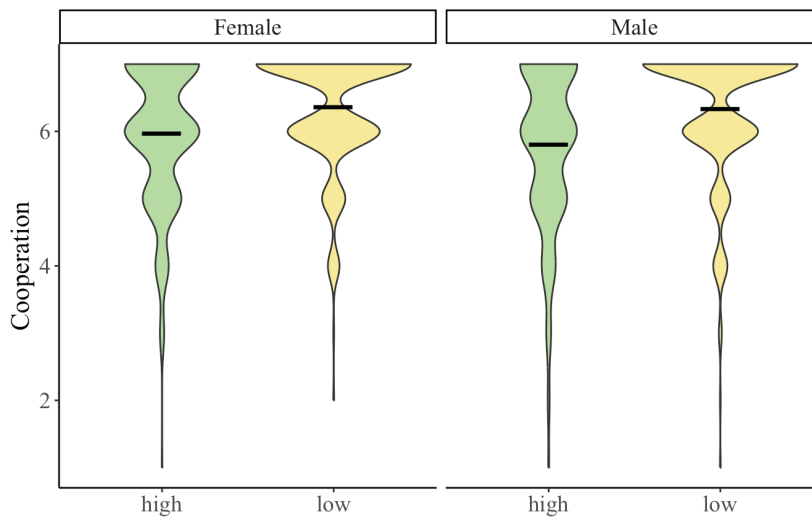


#### *Study 2: Care context*

We found similar results in the care context with a significant main effect for effort ( $F(1,699) = 116.84, p < .001, \eta^2 = .049$ ) and no significant effort × gender interaction ( $F(1,699) = 2.68, p = .102, \eta^2 < .001$ ). However, the main effect of gender reached no significance ( $F(1,699) = 2.17, p = .141, \eta^2 = .002$ ), indicating that the higher satisfaction for female cooperation partners was not replicated in the care context. Figure 4 visualizes this result.

**Figure 4**

*Cooperation Partner Satisfaction Across Gender and Effort in the care context*



### Exploratory Analyses

Based on the collected data, subsequent exploratory analyses are presented below.

#### *Two-Way Interaction (Vignette Gender × Gender Role Beliefs)*

To test whether differences in gender role beliefs influence the effort moralization effect (dimensions of core goodness and value commitment), linear regression analyses were conducted with gender role beliefs (GRB), the target gender and their interaction. We observed no significant effect of gender nor the interaction of gender and gender role beliefs. Yet, we found a minimal negative main effect of gender role beliefs on differences in value commitment in the work context ( $\beta = -.17$ , 95%-CI  $[-.23, -.10]$ ,  $p < .001$ ,  $R^2_{\text{adjusted}} = .027$ ) as well as the care context ( $\beta = -.11$ , 95%-CI  $[-.18, -.04]$ ,  $p = .004$ ,  $R^2_{\text{adjusted}} = .008$ ). Thus, participants with higher scores in gender role beliefs scale (more traditional) indicated a smaller difference in moral judgments between the high and the low effort target.

### ***Three-Way Interaction (Participant Gender × Gender Role Beliefs × Vignette Gender)***

The three-way interaction of Participant Gender × Vignette Gender × GRB as a predictor of differences in moral judgment was also conducted for both contexts. In none of the four models significance of the three-way interaction was observed. Similarly, there were no significant two-way interactions. Within both contexts, a significant negative main effect of gender role beliefs on value commitment differences was conducted as before (work context:  $\beta = -.16$ , 95%-CI [-.23, -.09],  $p < .001$ ,  $R^2_{\text{adjusted}} = .021$ ; care context:  $\beta = -.10$ , 95%-CI [-.18, -.03],  $p = .008$ ,  $R^2_{\text{adjusted}} = .006$ ).

## Discussion

### Summary

Our results replicated the effort moralization effect in the work context but showed smaller effect sizes in the care context. Contrary to our expectation, based on previous research, low-effort targets were rated as more satisfactory cooperation partners.

### Aim 1

The replicated effort moralization effect in the work context and its limited generalizability to the care context reveal potential context specific mechanisms.

### *Study 1: Work context*

Our findings mostly replicate the core effort moralization effect in the work context, in line with previous research that suggests people attribute higher morality to those who exert more effort (Bigman & Tamir, 2016; Celniker et al., 2023; Tissot & Roth, 2024). Participants attributed higher levels of morality for both core goodness and value commitment to the high-effort target. Although our effect sizes are lower than in prior studies undertaken in the U.S, they are more similar to those found in France (Celniker et al., 2023, Studies 2a, 2c). Interestingly, while effect sizes were moderate for core goodness and value commitment, warmth displayed smaller effect sizes, suggesting that effort plays a smaller role in judgments about people's warmth or it may be influenced by further characteristics such as friendliness or emotional accessibility (Fiske et al., 2007). As expected and shown in previous research, the high-effort target was perceived as less competent than the low-effort target. This may be due to an 'effort discounting' heuristic where individuals perceive those achieving outcomes with apparent ease as inherently more competent compared to those with greater effort (Heider, 1958). Contrary to earlier research, we found no significant results for pay deservingness.

An important observation is that high effort was not consistently rated as more positive. Surprisingly, participants would be more satisfied to work with the low-effort target. This is contrary to prior findings, such as Celniker et al. (2023, Studies 4 & 6), where the high-effort target was preferred as a cooperation partner to choose for a follow-up trust game about money. Choosing the high-effort target seems rational, as perceived morality increases trustworthiness (Delgado et al., 2005).

A potential explanation is the involuntary nature of our cooperation task. While Celniker et al. (2023) examined the free choice of a cooperation partner in a general trust game, our study measured the satisfaction with an assigned cooperation partner for an important task in the same area as shown in the vignette (work or care). Consequently, participants might have focused on practical aspects of their cooperation, like competence, rather than their moral judgments. In fact, the high-effort condition could evoke negative concerns about efficiency and the ability to deal with challenges, essential in work contexts. Hence a person putting in a lot of effort for the same performance might seem less proficient and effective for teamwork. This is supported by Amos et al. (2019) pointing out that effort can create ambivalent effects, either considered positively in terms of engagement or negatively in terms of inefficiency, depending on context and interpretations of competence or overcompensation.

### ***Study 2: Care context***

Our attempt to generalize the effort moralization effect to the care context was not fully successful with slightly different results and greater variability in the observed effect sizes. Even though we found a significant effect for core goodness, supporting the effort moralization effect, the effect size was only half the size ( $d = -0.14$ ) of the work context ( $d = -0.30$ ). The analyses revealed no statistically significant effects regarding value commitment. One explanation might be that effort in the care context might not be perceived as an extra moral value but rather as an expected matter of course (Kennedy et al., 2016). Additionally, social expectations might differ between work and care contexts. Performance and effort are a central evaluation criteria in the business environment and success-oriented areas such as sports (Celniker et al., 2023). While this is in line with the logic of meritocracy, care work might be more associated with empathy, caring and intrinsic motivation in judgments rather than mere visible effort. In fact, visible effort in care scenarios might be perceived negatively, suggesting that people struggle with tasks that should feel natural and genuine. If moral evaluation in the care context is influenced by emotional connection, effort might lose importance in moral judgments.

Just as in the work context, the low-effort target was considered as more competent and participants preferred them as a cooperation partner. Another possible explanation for the lower satisfaction with the high-effort target might be a subtle form of reactance, where people react to perceived social pressure with resistance and rejection (Miron & Brehm, 2006). Participants

may feel that the high-effort target overexerts, leading to lower satisfaction for cooperation, despite the same outcome.

## **Aim 2**

The lack of gender-specific differences in effort moralization across work and care contexts suggests that effort can be interpreted independently of gender, which is contrary to predictions derived from stereotypical role norms.

### ***Study 1: Work context***

We found no significant interactions between the moralization of effort and the target gender. This suggests that effort in the work context is judged equally across genders. This might be due to a growing equality and balance at work (Schaerer et al., 2023). Indeed, a meta-analysis of field audit studies (1976-2020) from Schaefer et al. (2023) found that there is no longer any discrimination between men and women in workplaces with a mixed gender ratio, and in some cases, there is even a slight advantage for women. However, a study by McKinsey & Company from 2024 shows that although women have made progress in leadership positions, they are still underrepresented, especially women of color (Krivkovich et al., 2024).

### ***Study 2: Care context***

In the care context, no significant interactions were found either. Men and women were judged equally regardless of effort levels which contradicts our theoretical considerations assuming less appreciation for high-effort women in the care context as they would merely fulfil societal expectations (Kennedy et al., 2016). Notably, there was a slight tendency for men receiving higher moral values, reinforcing our assumptions of men being perceived as going beyond and deviating positively from the norm in care contexts whereas female effort is being taken for granted (Deutsch & Saxon, 1998). However, this tendency was not statistically significant and should be interpreted with caution.

Yet, it is possible that the attempt to remove emotionally charged elements from the care vignette could be the reason for the absence of gender differences. The tasks might have been designed in a too functional manner for a care context, potentially suppressing stereotypical attributions (Fiske & Taylor, 1991).

### **Aim 3**

Our findings from Aim 3 seem to be contrary to the central assumption of the effort moralization theory, where higher effort typically leads to higher moral evaluations and therefore to higher willingness to cooperate and satisfaction. In contrast to prior research by Celniker et al. (2023), participants in our study consistently rated higher satisfaction for the low-effort target as a cooperation partner. This was true for both contexts: work and care and regardless of gender. This inverted effort effect was moderate ( $\eta^2 = .049$  in the care context) (Cohen, 1988), suggesting that high effort is not rated positively in cooperative situations. On the contrary: people who show high effort might be perceived as dominant and controlling, insecure or incompetent, while people exerting low effort are perceived as more competent, efficient and confident – which reflects itself in a higher satisfaction for collaboration. This means competence might override moral worth in certain situations. Particularly in the care context where emotional intelligence, calmness and resilience might play an important role, high effort might be interpreted as a sign of emotional imbalance, nervousness and agitation.

Interestingly, in the work context, participants indicated higher satisfaction to cooperate with the female target in both effort levels. This might be due to the fact that women are often perceived as inherently more communal and cooperative, qualities valued in cooperative contexts (Abele et al., 2008; Hentschel et al., 2019). However, male and female targets were not compared directly with each other and thus cannot be interpreted as sound.

The consideration of gender role beliefs as a moderator had also no significant results for both contexts. This supports one central observation of this research: the effort moralization effect appears to be independent of gender.

### **Limitations**

First, this study examines how the moralization of effort differs by the target gender in our vignettes (male vs. female) and the context (work vs. care). However, it does not make a direct comparison between men and women on how they evaluate effort differently. Same-gender biases such as judging effort more favorably when reporting for individuals' own gender or cross-gender valuation (e.g. men expecting women's effort in care work, women being more skeptical of men's effort in caregiving roles) could play a role in their moral judgments (Trepte & Loy, 2017).

A further limitation concerns the alignment of the vignettes between the work and the care scenario. The work vignette contains quantitative performance indicators such as orders processed per hour and error rates, while the care vignette uses more qualitative descriptions like grocery shopping or doing laundry. This might have distorted our results and contributed to the higher mean values in the care context. Closer alignment, for instance by adding information about task completion or parent satisfaction in the care vignette would increase internal validity and ensure that perceived effort and outcomes are presented in equal detail across contexts. Moreover, the elimination of tasks that were occupied with traditional stereotypical descriptions in our vignettes leading household chores rather than emotional caregiving in the care vignette might have caused the lack of observed gender differences.

Additionally, an important limitation refers to the choice of cooperation partners. While Celniker et al. (2023) examined the active and voluntary choice of a cooperation partner as part of a trust game, we investigated the satisfaction with an assigned partner for a future task as this seems more realistic for the real world. Unlike Celniker et al. (2023) where the trust game context was independent from the original effort context, our cooperation satisfaction was conducted within the same thematic context (work or care) as described in the previous vignette. This methodological difference might have led to different evaluation criteria. Perceived morality increases trustworthiness and trustworthiness lays the foundation for partner choice in a trust game (Delgado et al., 2005; Eckel & Wilson, 2004). This might be in favor of the high morality and thus high effort target in the setting of Celniker et al. (2023). However, in our fixed and task-related contexts, functional criteria such as perceived competence could hold more importance and thus explain the preference for the low effort target despite their lower morality.

Another limitation of this study might lie in our measure of pay deservingness. Although we found lower pay suggestions in the care context, this might be due to methodological differences. While we gave an industry benchmark of \$24 in study 1 as an anchor, there was no point of reference in study 2, yet participants had to suggest the pay on the same sliding scale ranging from \$0 to \$50. Since anchoring biases alter numerical estimations (Tversky & Kahneman, 1974), this change in methodology likely influenced our results.

Finally, participants consist of US citizens recruited through the online platform prolific. They do not fully represent the general population and may not be generalizable to

non-western or collectivistic cultures. This might be important though since effort and gender might be perceived differently in those cultures, placing great value on communal contributions rather than individual achievement (Cheng et al., 2020). In addition, prior research has already shown cultural distinctions in the moralization of effort (Celniker et al., 2023; Tissot & Roth, 2024).

## **Future Research**

This study extends existing research and contributes to the developing research about the effort moralization effect to previously underrepresented contexts. However, several promising aspects remain open for future research.

Firstly, future studies should make a direct comparison between men and women on how they evaluate effort, instead of focusing only on the target gender. Investigating same-gender and cross-gender biases could reveal preferential treatment for certain groups and stereotypes in work and care contexts.

Secondly, our findings raise interesting questions concerning the interaction of moral worth, competence and cooperation partner preference. In contrast to Celniker et al.'s (2023) trust games focusing on central moral traits like trustworthiness, our future task scenario shows a possible trade-off between moral esteem and practical use. Future research should investigate when and under which circumstances individuals favor one over the other for example in high-pressure settings, tight deadlines or obligations to achieve results. Thus, when choosing a cooperation partner, this could help clarify whether morality or competence is weighted more heavily depending on context and task requirements, or whether both factors are considered together.

Lastly, future studies should explicitly investigate how different contexts such as (active) partner choice in trust games and (assigned) partners for a future task influence perceived morality, competence and the evaluation of cooperation partner. Beyond work and care, further social and organisational settings should be explored to determine the plausible situational boundaries of the effort moralization effect and its resulting consequences.

## **Conclusion**

This study addresses a gap in literature by investigating how the moralization of effort varies between work and care contexts and how gender differences shape moral evaluations.

Our findings replicated the effort moralization effect in the work context. In the care context, our effects were less strong, indicating different moral priorities and possibly different social expectations in different contexts. Gender-specific differences could not be demonstrated, nor moderation effects by gender role beliefs suggesting a gender generalizable effect. Contrary to our expectations and prior studies, participants preferred the low-effort target as an assigned cooperation partner for a future task which suggests different evaluation criteria for cooperation choice depending on the context.

Our results suggest that effort moralization is a robust, yet context sensitive mechanism of social evaluation. Future research should further examine how different contexts and cooperation conditions influence the moralisation of effort and resulting social preferences as it holds important implications for real-world scenarios such as (career) opportunities, social professions, reward and fairness in everyday life.

## References

- Abele, A. E., Uchrowski, M., Suitner, C., & Wojciszke, B. (2008). Towards an operationalization of the fundamental dimensions of agency and communion: Trait content ratings in five countries considering valence and frequency of word occurrence. *European Journal of Social Psychology*, 38(7), 1202–1217. <https://doi.org/10.1002/ejsp.575>
- 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>
- Arof, K. Z. M., Ismail, S., & Saleh, A. L. (2018). Contractor's performance appraisal system in the Malaysian construction industry: Current practice, perception and understanding. *International Journal of Engineering & Technology*, 7(3.9), 46. <https://doi.org/10.14419/ijet.v7i3.9.15272>
- Aronson, E., Wilson, T. D., Akert, R. M., & Sommers, S. R. (2021). *Social psychology* (Tenth edition. Global edition). Pearson Education Limited.
- Bakan, D. (1966). *Duality of human existence: An essay on psychology and religion*. Beacon Press.
- 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 rRole

- 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.
- Cheng, A. W., Rizkallah, S., & Narizhnaya, M. (2020). Individualism vs. collectivism. *The Wiley Encyclopedia of Personality and Individual Differences*, 287–297. <https://doi.org/10.1002/9781119547181.ch313>
- Cislaghi, B., & Heise, L. (2019). Gender norms and social norms: differences, similarities and why they matter in prevention science. *Sociology of Health & Illness*, 42(2), 407–422. <https://doi.org/10.1111/1467-9566.13008>
- 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>
- Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, 8(11), 1611–1618. <https://doi.org/10.1038/nn1575>
- 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>
- Eckel, C., & Wilson, R. (2004). Whom to trust? Choice of partner in a trust game. *Department of Economics, Virginia Tech*.
- Ellemers, N., Van der Toorn, J., Paunov, Y., & Van Leeuwen, T. (2019). The psychology of morality: A review and analysis of empirical studies published from 1940 through

2017. *Personality and Social Psychology Review*, 23(4), 332–366.  
<https://doi.org/10.1177/1088868318811759>
- 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), 355–379.  
<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>
- Gert, B. (2005). Morality. In *Oxford University Press eBooks*.  
<https://doi.org/10.1093/0195176898.001.0001>
- 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>
- Haines, E. L., Deaux, K., & Lofaro, N. (2016). The times they are a-changing ... or are they not? A comparison of gender stereotypes, 1983–2014. *Psychology of Women Quarterly*, 40(3), 353–363. <https://doi.org/10.1177/0361684316634081>
- Heider, F. (1958). The psychology of interpersonal relations. In *John Wiley & Sons, Inc. eBooks*. <https://doi.org/10.1037/10628-000>
- 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. <https://doi.org/10.3389/fpsyg.2019.00011>

- 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>
- Krivkovich, A., Field, E., Yee, L., McConnell, M., & Smith, H. (2024, September 17). *Women in the workplace 2024: The 10th-anniversary report*. McKinsey & Company.  
<https://www.mckinsey.com/featured-insights/diversity-and-inclusion/women-in-the-workplace>
- 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/>
- McKinsey & Company. (2022). *Women in the workplace 2022*. McKinsey & Company.
- Miron, A. M., & Brehm, J. W. (2006). Reactance theory - 40 years later. *Zeitschrift für Sozialpsychologie*, 37(1), 9–18. <https://doi.org/10.1024/0044-3514.37.1.9>
- 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/](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>

- Roth, L. H. O., Tissot, T., Fischer, T., & Masak, S. C. (2025, February 28). See me, judge me, pay me: Gendered effort moralization in work and care. [https://doi.org/10.31219/osf.io/wkbs6\\_v1](https://doi.org/10.31219/osf.io/wkbs6_v1)
- 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](https://doi.org/10.1007/978-3-319-21635-5_2)
- Schaerer, M., Du Plessis, C., Nguyen, M. H. B., Van Aert, R. C., Tiokhin, L., Lakens, D., Clemente, E. G., Pfeiffer, T., Dreber, A., Johannesson, M., Clark, C. J., & Uhlmann, E. L. (2023). On the trajectory of discrimination: A meta-analysis and forecasting survey capturing 44 years of field experiments on gender and hiring decisions. *Organizational Behavior and Human Decision Processes*, 179, 104280. <https://doi.org/10.1016/j.obhdp.2023.104280>
- 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>
- Statista Research Department. (2024, July 5). *U.S. female labor force participation rate 1990-2023*. Statista. <https://www.statista.com/statistics/191737/us-female-labor-force-participation-rate-since-1990/>
- Tissot, T. T., & Roth, L. H. O. (2024). 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. <https://osf.io/5tkfr>

- 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>
- Trepte, S., & Loy, L. S. (2017). Social identity theory and self-categorization theory. *The International Encyclopedia of Media Effects*, 1–13.  
<https://doi.org/10.1002/9781118783764.wbieme0088>
- 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>

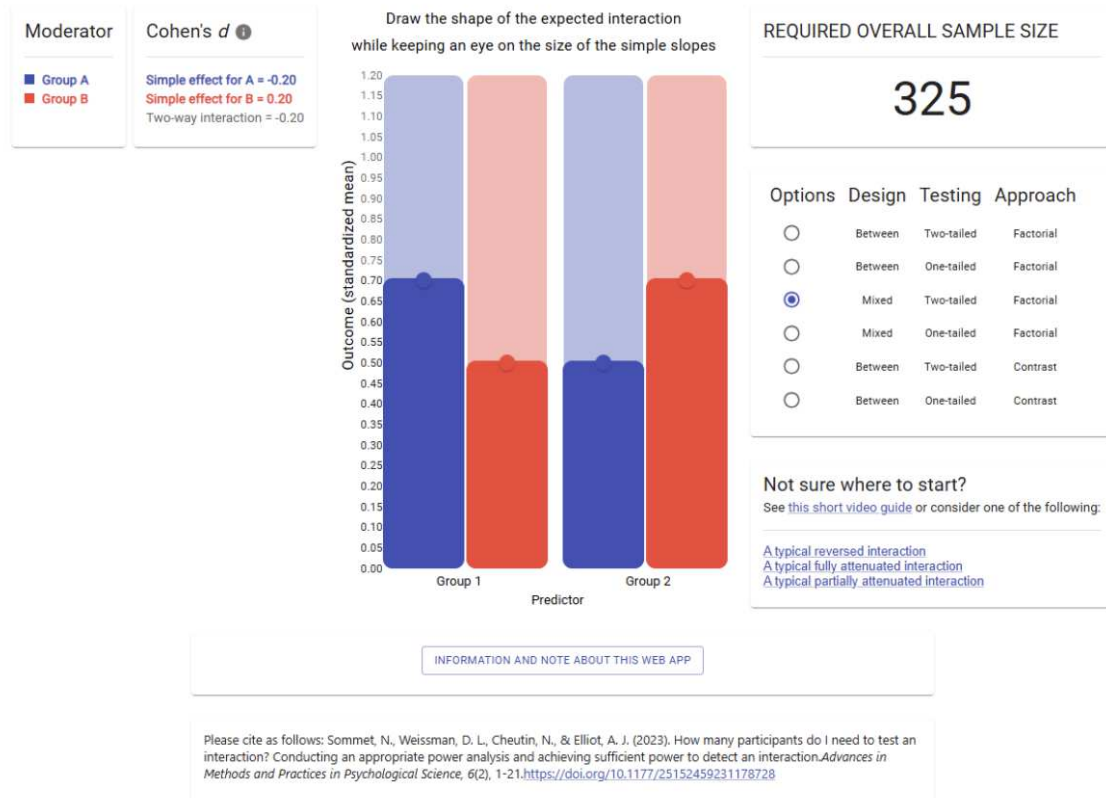
Note:

For some language refinement and minor stylistic adjustments, AI tools such as ChatGPT were used.

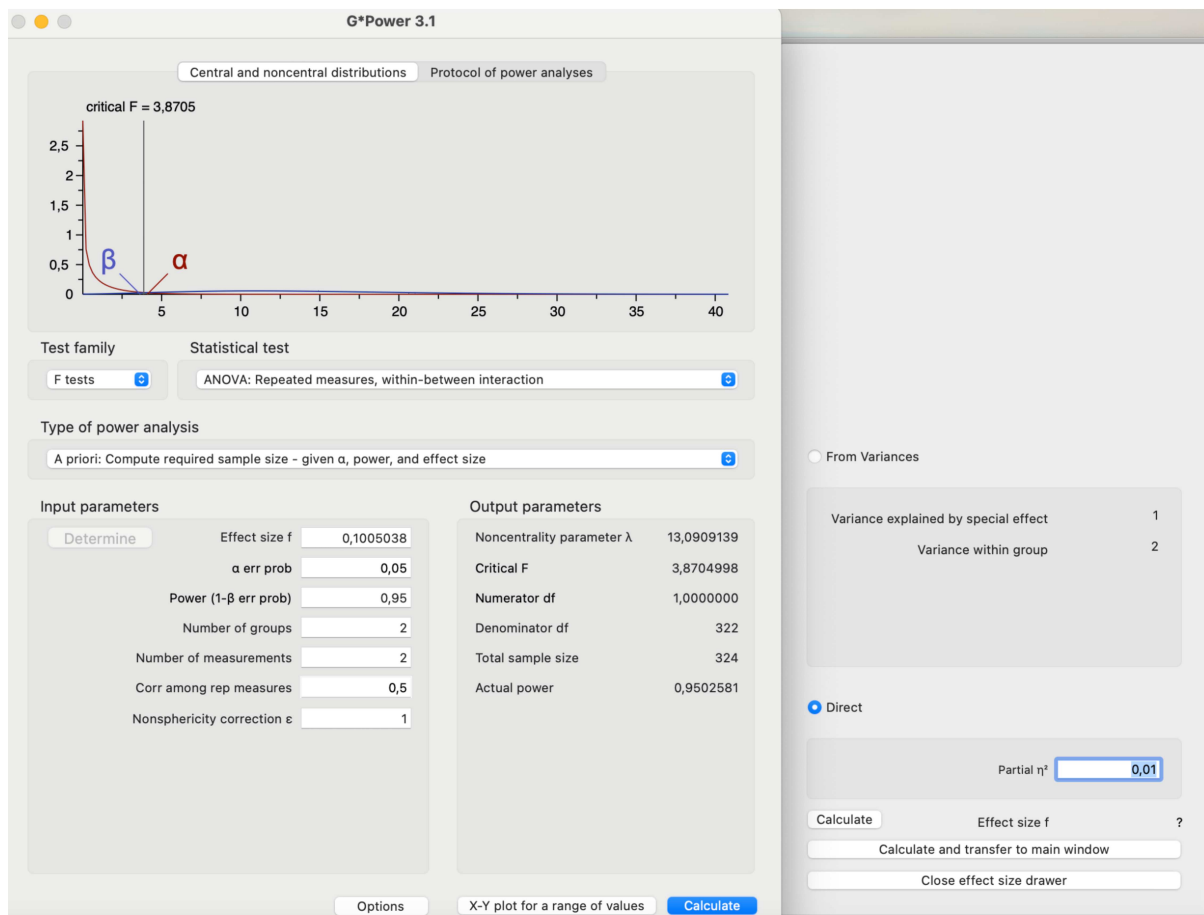
## Supplemental Material

### Power Analysis Interaction (Sommet et al., 2023)

INTxPower: Finding the target sample size to detect a two-way interaction with power .95 (α = .05)



## Power Analysis ANOVA (Faul et al., 2009)



## Vignettes Work context

### Male

On the following pages, you will be asked to read about Michael and Lukas, two employees at the same company.

Michael and Lukas have the same job processing similar orders in the company's office.

To ensure that you have read and understood the following scenario, you will need to remain on the next page for at least 20 seconds before proceeding with the study.

Please read the scenario below with care and attention:

*Michael and Lukas work at the same company and process similar orders in the company's office. Both Michael and Lukas 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 Michael's orders and 96% of Lukas's orders are error-free and complete. On average, Michael and Lukas each process correct orders worth \$144 per hour.*

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

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

## Female

On the following pages, you will be asked to read about Anna and Sophie, two employees at the same company.

Anna and Sophie have the same job processing similar orders in the company's office.

To ensure that you have read and understood the following scenario, you will need to remain on the next page for at least 20 seconds before proceeding with the study.

Please read the scenario below with care and attention:

*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 orders and 96% of 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 also works as quickly as possible, she finds the work hard.

## Vignettes Care context

### Male

On the following pages, you will be asked to read about Michael and Lukas, two twin brothers who take turns helping their elderly parents with household chores.

Michael and Lukas perform the same tasks such as grocery shopping, doing laundry, cooking meals, etc.

To ensure that you have read and understood the following scenario, you will need to remain on the next page for at least 20 seconds before proceeding with the study.

Please read the scenario below with care and attention:

*Michael and Lukas are twin brothers who take turns helping their elderly parents on weekends with household chores. Each weekend, one of them spends approximately three hours assisting with various tasks such as grocery shopping, doing laundry, and cooking meals. Both Michael and Lukas manage to complete all the necessary tasks during their time, ensuring that the house is clean, the laundry is fresh, and meals are prepared. By the end of the day, the home is in perfect order, with no visible differences in the cleanliness or organization achieved by either brother — both leave the house well-maintained and with the parents' needs fully met.*

For Michael, performing these household chores requires minimal effort — although he works as quickly as possible, he finds the tasks easy.

For Lukas, however, performing these chores requires a lot of effort — although he also works as quickly as possible, he finds the tasks hard.

## Female

On the following pages, you will be asked to read about Anna and Sophie, two twin sisters who take turns helping their elderly parents with household chores.

Anna and Sophie perform the same tasks such as grocery shopping, doing laundry, cooking meals, etc.

To ensure that you have read and understood the following scenario, you will need to remain on the next page for at least 20 seconds before proceeding with the study.

Please read the scenario below with care and attention:

*Anna and Sophie are twin sisters who take turns helping their elderly parents on weekends with household chores. Each weekend, one of them spends approximately three hours assisting with various tasks such as grocery shopping, doing laundry, and cooking meals. Both Anna and Sophie manage to complete all the necessary tasks during their time, ensuring that the house is clean, the laundry is fresh, and meals are prepared. By the end of the day, the home is in perfect order, with no visible differences in the cleanliness or organization achieved by either sister — both leave the house well-maintained and with the parents' needs fully met.*

For Anna, performing these household chores requires minimal effort — although she works as quickly as possible, she finds the tasks easy.

For Sophie, however, performing these chores requires a lot of effort — although she also works as quickly as possible, she finds the tasks hard.

## Core goodness, value commitment and competence/warmth scale

(Celniker et al., 2023)

|                                               | Does not describe Anna very well |                       |                       | Describes Anna moderately well |                       |                       | Describes Anna extremely well |
|-----------------------------------------------|----------------------------------|-----------------------|-----------------------|--------------------------------|-----------------------|-----------------------|-------------------------------|
| Moral                                         | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Motivated                                     | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Respectful                                    | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Responsible                                   | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Warm                                          | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Disciplined                                   | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
|                                               | Does not describe Anna very well |                       |                       | Describes Anna moderately well |                       |                       | Describes Anna extremely well |
| Self-controlled                               | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Please select "Describes Anna extremely well" | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Competent                                     | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Determined                                    | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Cooperative                                   | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Trustworthy                                   | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
|                                               | Does not describe Anna very well |                       |                       | Describes Anna moderately well |                       |                       | Describes Anna extremely well |
| Dedicated                                     | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Just                                          | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Honest                                        | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |
| Principled                                    | <input type="radio"/>            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>          | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>         |

### Perceived effort, quality, difficulty and work value

How much effort do you think **Anna** puts into her care work?

- ☐ No effort at all
- ☐
- ☐
- ☐ An average amount of effort
- ☐
- ☐
- ☐ A lot of effort

In your opinion, how well does **Anna** perform her care work?

- ☐ Very bad
- ☐
- ☐
- ☐ Average
- ☐
- ☐
- ☐ Very good

Compared to other care work, how difficult is **Anna's** care work?

- ☐ Not at all difficult
- ☐
- ☐
- ☐ Moderately difficult
- ☐
- ☐
- ☐ Extremely difficult

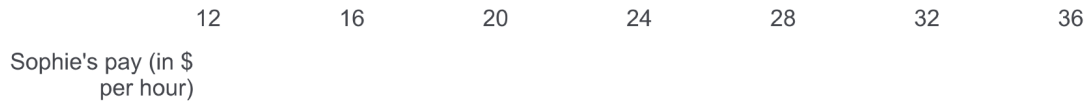
How valuable do you think **Anna's** care work is?

- ☐ Not valuable at all
- ☐
- ☐
- ☐ Moderately valuable
- ☐
- ☐
- ☐ Extremely valuable

## Pay deservingness (<https://www.erieri.com>)

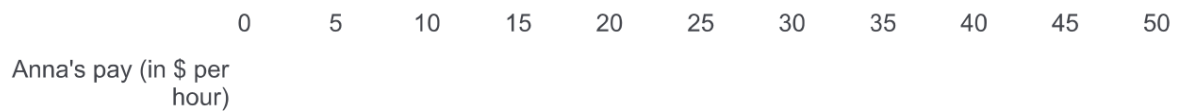
### Work

The average office worker at the company makes \$24 an hour. How much do you think **Sophie** should make per hour?



### Care

Imagine that **Anna** was paid for her care work. How much should she be paid per hour?



### Cooperation partner satisfaction

|                                                         | Extremely<br>dissatisfied | Moderately<br>dissatisfied | Slightly<br>dissatisfied | Neither<br>satisfied<br>nor<br>dissatisfied | Slightly<br>satisfied | Moderately<br>satisfied | Extremely<br>satisfied |
|---------------------------------------------------------|---------------------------|----------------------------|--------------------------|---------------------------------------------|-----------------------|-------------------------|------------------------|
| How satisfied would you be to work with <b>Anna</b> ?   | <input type="radio"/>     | <input type="radio"/>      | <input type="radio"/>    | <input type="radio"/>                       | <input type="radio"/> | <input type="radio"/>   | <input type="radio"/>  |
| How satisfied would you be to work with <b>Sophie</b> ? | <input type="radio"/>     | <input type="radio"/>      | <input type="radio"/>    | <input type="radio"/>                       | <input type="radio"/> | <input type="radio"/>   | <input type="radio"/>  |

**Gender role beliefs scale** (Brown & Gladstone, 2012)

10 Items, with seven answer possibilities ranging from 1 = “strongly agree”, 4 = “undecided” to 7 = “strongly disagree” for each item

1. It is disrespectful to swear in the presence of a lady.
2. The initiative in courtship should usually come from the man.
3. Women should have as much sexual freedom as men.
4. Women with children should not work outside the home if they don't have to financially.
5. The husband should be regarded as the legal representative of the family group in all matters of law.
6. Except perhaps in very special circumstances, a gentleman should never allow a lady to pay the taxi, buy the tickets, or pay the check.
7. Men should continue to show courtesies to women such as holding open the door or helping them on with their coats.
8. It is ridiculous for a woman to run a train and for a man to sew clothes.
9. Women should be concerned with their duties of childrearing and house tending, rather than with desires for professional and business careers.
10. Swearing and obscenity is more repulsive in the speech of a woman than a man.