Stage 1

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

Tassilo T. Tissot^{1*} & Leopold H. O. Roth^{2*}

¹Ghent University, Faculty of Psychology and Educational Sciences, Department of Developmental, Personality, and Social Psychology, Ghent, Belgium,

²University of Vienna, Faculty of Psychology, Department of Occupational, Economic, and Social Psychology, Vienna, Austria

*authors share first authorship and contributed equally to the manuscript

Contact:

Tassilo T. Tissot, tassilo.tissot@ugent.be, 0000-0002-6215-975X

Leopold H. O. Roth, leopold.roth@univie.ac.at, 0000-0002-1120-4733 (corresponding)

The approved STAGE 1 version of the manuscript is available on OSF at this link:

 $\underline{https://osf.io/643gt/}. \ The \ project \ is \ available \ at this \ link: \ \underline{https://osf.io/k3f4y/}.$

Abstract

Inferring moral character of individuals is an adaptive need for social decision-making. The effort moralization effect describes the finding that people who expore effort in a task are seen as more moral, even if higher effort does not enhance the outcome (e.g., higher performance or better quality). We replicated this effect, based on Celniker et al. (2023, Study 6), in countries not yet investigated (Germany and Mexico). Further, drawing on discussions around workforce participation (see 'great resignation' or 'quiet quitting') criticizing the lower work ethic of younger individuals (e.g., the so-called Gen Z), we tested whether (ineffective) effort is moralized less in younger individuals. Our findings support the generalizability of effort moralization to Germany and Mexico, yet point to heterogeneity in effect strength. However, our findings do not support the hypothesis that effort moralization increases with age. This indicates that young people do not see effort as less valuable than older people. We discuss further implications and limitations of these findings and suggest avenues for future research on effort moralization.

Keywords: Registered Report, Effort moralization, Generation effect, Replication, Multicountry, Work ethic

PCIRR-Study Design Table

Table 1

PCI-RR Design Table

Question	Hypothesis	Sampling plan	Analysis plan	Rationale for sensitivity	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
Can the effort moralization effect be replicated in the overall sample?	Participants will rate the person, showing high-effort behavior, as more moral, even though the added effort doesn't increase the productivity or quality	We will use the services of online panel providers (e.g., Prolific) as well as social media sampling to reach a total sample of 680 complete cases (340 per country) to have at least <i>N</i> = 327 valid cases (passing	We will use two- sided dependent t- tests on the pooled data as well as by country to test the differences in moral evaluation (core goodness and value commitment) by effort condition and differences in	We aim for a sensitivity of $d = 0.20$ as the smallest effect of interest. Targeting a power of .95, a respective minimum sample size of $N = 327$ was computed. We plan to oversample every country to 340 to	We will use the criteria by LeBel et al. (2019) to evaluate the replication, utilizing the reported $d = 0.42$ from Celniker et al. (2023, p.73, study 6) for core goodness and $d = 0.76$ for value commitment.	Effort moralization theory's generalizability could be shown undetectable under the current conditions of the study.

	These effects will be found in every by-country analysis.	exclusion criteria) by country. We will aim to achieve a roughly equal distribution of participants in the following age groups: 18 - 30, 31 - 45, 46 - 60,	deserved pay by effort condition. We will further test differences in perceived warmth and perceived competence. Yet, prior research indicated a variance of effects between countries here.	compensate for potential exclusions. Power analysis was conducted, using G*Power 3.1.9.7 [see supplemental material].	This will be done, using the criteria signal, consistency, and direction.	
Does the strength of the effort moralization effect depend on the age of the evaluator in situations where further effort does not improve the quality of the outcome?	The effort moralization effect will be predicted positively by the age of the participants (higher age, stronger effect). Pay deservingness differences by condition will be predicted positively by the age of the participants (higher age, higher deservingness).	> 60 years We will fill the sample in other groups if the sample in a respective age group is not completed within 3 weeks after the beginning of data collection.	We will extend the analyses regarding the previous hypothesis by conducting by-country regression analyses with age as predictor and the discrepancy in moral evaluation, and deserved pay between ratings as dependent variables.	To detect the smallest effect of interest of $R^2 = .15$ with one predictor, power = .95 and α = .05, a minimum sample size of $N = .76$ complete and valid cases by countries is required. As described above, we aim to sample 340 complete cases per country.	If the effect is significant, pointing in the expected direction and of expected magnitude in all countries, we interpret the effect as generalizable to all respective countries. If the effect is only observed in one of the countries, it is not generalizable to	The idea that effort moralization is less expressed in younger participants could be shown to be undetectable under the current circumstances or not generalizable across all investigated countries.

These effects will be found in every by-		all sampled countries.	
country analysis.		If the effect is never observed, the effect cannot be assumed	
		under the given conditions.	

Introduction

'It seems like nobody wants to work these days.' (Kardashian, K. in Variety, 2022).

The ideological debate about the lack of qualified workforce specifically amongst younger potential employees has become a common theme of the news (Medlar et al., 2022). While there is a series of systemic reasons that reduce the supply of workforce to certain fields, such as demographic changes, stagnating wages, working conditions, and delayed effects of the COVID-19 pandemic (Műrage et al., 2022; Pillai, 2023; Silverstein, 2008; Smith, 2022), debates often focus on constructs like 'work ethic' or 'laziness' and commonly target the youngest generation in the workforce (Formica & Sfodera, 2022). While the idea that younger generations are lazy and morally inferior to preceding generations is a recurring theme throughout history, perspectives on work are indeed changing, leaving companies with potential employees who are less willing to provide unpaid services or excessive overtime (Chillakuri, 2020; Xueyun et al., 2023). One avenue to approach this topic is the so-called effort moralization effect (Bigman & Tamir, 2016; Celniker et al., 2023), which describes the translation of observed behavioral effort into a moral judgment of the agent. This effect persists even if the effort is not productive (Celniker et al., 2023). The effort moralization effect can inform the debate around the lack of qualified and willing workforce. We hypothesize that younger individuals indicate less effort moralization of ineffective labor. YoungerRather, younger individuals do not judge ineffective effort as a sign of higher morality based solely on the effort. We hence aim to replicate the findings by Celniker et al. (2023, Study 6) in two countries (Mexico and Germany) and test whether participants' age explains differences in the effort moralization effect. Replicating the effort moralization effect offers a new perspective on debates around the supposedly lower 'work ethic' of younger generations.

Impressions of character as a function of behavior

To infer character values of new encounters is an adaptive and inherent behavior, concerning philosophy and psychology for the longest (Doris & The Moral Psychology Research Group, 2010). Several philosophical traditions suggest that morality can be inferred only from the actions of individuals (Fengyan, 2004; Johnson & Cureton, 2004; Telfer, 1989). Moral judgment is crucial for social decision-making—an often automatic process by which we form impressions about the morality of others' behavior (Uhlmann et al., 2015)—such as in cooperation settings (Celniker et al., 2023; Everett et al., 2016; Van Lange & Kuhlman, 1994).

To attribute morality to others, most individuals depend on approximations of character virtue through observation in daily life. While individuals rely on a variety of cues for this purpose, including facial and body expressions (Horberg et al., 2013), stereotypical appearance (Grizzard et al., 2018), or religious beliefs (Gervais, 2011), one of the main signals for inferring the morality of others remains behavioral observation (Mickelberg et al., 2022; Pizarro & Tannenbaum, 2012). Observations These observations create inferences about a person's morality based on trivial actions in daily life.

The Moralization of Effort

The effort motion attion effect describes one process of moral judgment. It describes delineates how people appear to use effort invested in given tasks as information on the morality of agents. The core idea of the moralization of effort is the heuristic that higher effort results in higher performance. The Yet, the focal interest in the effect concentrates on a special case of effort moralization: when effort makes no difference in the outcome.

Tamir (2016) delivered foundational insights into this effect across seven studies. These showed that perceived effort intensified judgments of both immoral and moral er effort on moral behavior led to higher moral judgment of the described agents (e.g., person). This was still observed when the action did not lead to the desired oute behavior (e.g., successfully returning a found wallet to the owner). Further research suggested that effort moralization follows certain norms as boundary conditions (Berry & Lucas, 2022). In four studies, it was shown that the effort moralization effect does not linearly increase character judgment but plateaus when agents recruit 'excessive' effort that reaches beyond societal standards of effort investment (e.g., revisiting the spot of the found wallet three days in a row). Celniker et al. (2023) tested the effort moralization effect across eight studies ruling out potential biases such as differences in quality of work or effort withholding. Further, they reported that participants were more likely to choose individuals who invested higher effort in a task as cooperation partners even when the behavior did not lead to better or more outcomes. This finding is similar to Barclay's (2013) remarks on the nature of altruism, which is thought to be expressed in order to be seen as a more attractive option in the market of available cooperation partners. Research on the 'martyrdom effect' finds that people report greater willingness to donate to a charitable cause when the contribution process is expected to be effortful rather than easy (Olivola & Shafir, 2013). Thus, the expression of effort despite being an inconsistent indicator of ability or productivity (Markovits, 2019; Shepperd et al., 1994) as a heuristic signal for judgment of character as well as cooperation seems to be utili intentions.

Changes in effort valuation at work

Older generations accusing younger generations of being 'lazy' is not a new phenomenon (Lang, 2023; Royle, 2024). While such perceptions of perceived freeloaders

activate strong social emotions (Petersen et al., 2011, 2012) there is little evidence of generational differences in actual productivity. -Yet, recent developments in labor markets as well as employee values appear to indicate some real changes in the perception of effort at work (see 'great resignation' or 'quiet quitting'). These trends imply that certain sections of today's workforce are tired of meaningless work and are striving for change (Medlar et al., 2022). We suggest that generational differences in the moralization of effort may help explain shifts in work values among younger generations that lead to conflicts with established norms.

Celniker and colleagues (2023) discuss effort moralization as a 'deeply rational' heuristic process (Kenrick et al., 2009), that enables individuals to easily incline cooperative intent and facilitates social decision-making and judgments of moral character (Celniker et al., 2023). Such heuristics reduce decision-making effort and required time. However, even though such mechanisms might prove fruitful on the individual level, they might lead to harmful norms on a societal level (Li et al., 2018). For instance, Celniker et al. (2023) theorize that effort moralization might explain the maintenance of virtue signaling by engaging in unproductive work. Virtue signaling aims to enhance one's moral reputation by publicly displaying actions that are socially perceived as moral, while the motivating source for this is status-seeking and not the moral expression itself (Westra, 2021). Signaling morality through (ineffective) effort might also foster resistance to less effortful processes or automated alternatives, and to policies that promote alternatives to economically redundant labor such as universal basic income (Celniker et al., 2023). It is plausible to assume that recent movements such as the 'great resignation', 'quiet quitting', etc. represent responses to such resistance fueled by a generational shift in work values and changing perceptions of (necessary) effort among younger generations.

We hypothesize that variations in the moralization of effort across age groups may explain these changes. Such differences provide a better understanding of social movements

than the popular (and unsubstantiated) notion that young people are inherently lazier than previous generations. Younger generations may view unproductive work differently than older generations, which may explain their tendency to reject work they consider unproductive. Consequently, our research seeks to contribute to the ongoing discussion about workforce participation.

Replication and Extension

The current study aims to replicate and extend the original findings by Celniker et al. (2023), specifically Study 6. The procedure included one vignette describing two workers, controlling for economic output, quality, and working on maximum capacity, and all earlier discussed possible biases in effort moralization (e.g., the output is identical, but the quality is higher, when the effort is higher; the low-effort individual is withholding effort by working slowly; etc.). Only the required effort for the work differs between the described workers. Note, that we focus on the focal effort moralization effect and don't test the second part of the experiment about preferred cooperation partners. We further apply the same measures for perceived moral virtue, separated by core goodness and commitment (see Piazza et al., 2014). While we replicate the procedure of Celniker et al. (2023), we will test the effect in countries that to our knowledge have not been included in earlier effort moralization research (Germany and Mexico¹). Testing the effect in different countries will inform its the generalizability-of the phenomenon. Prior research by Celniker et al. (2023) reported in Studies 2a-c that the magnitude of the effect may differ between populations (France: d = 0.38, South Korea: d = 0.71, United States of America: d = 0.60; see also Tierney et al., 2020, for cross-national

_

¹ Both countries are important economic entities in their respective region and have a combined population of > 200 mio. Further, the authors had the language abilities necessary for the translation process.

evidence of similar effects). We will further assess perceived warmth, competence and pay deservingness between effort conditions as tangenti easures.

We will further extend the available evidence by testing potential differences in effort moralization by participants' age. If younger individuals are less prone to moralize unproductive effort, the effect should be observable as a function of age. We will test this using age as a continuous (non-)linear predictor of the magnitude of effort moralization.

Deviations

We will deviate from the original study (Celniker et al., 2023) in three aspects. We will not apply the second part of the experiment (choice of cooperation partner), and neither assess ethnicity nor income.

Method

Important links

The table below Table 2 includes all necessary links to access the materials of the study.

Table 2
Summary of links to materials, code, data, and supplemental material

content	link
code and data (GitHub)	https://github.com/rothl16/mev
project (OSF)	https://osf.io/k3f4y/
code and data (OSF)	https://osf.io/zcq7m/
supplemental material (OSF)	https://osf.io/jxecn/
Qualtrics (OSF)	https://osf.io/98p7r/

Open Science



All materials, code, and data will be made openly accessible https://osf.io/k3f4y/ except data, which can identify individuals, such as mail addresses.

Power computation

The power computation for mean differences was based on the smallest effect size of interest (d=0.20) (Lakens, 2022). The smallest effect, reported by Celniker et al. (2023), critical for our study was d=0.42, quantifying differences in moral judgment (core goodness). We used G*Power 3.1.9.7 (Faul et al., 2009) to compute the required minimum sample size to detect the effect, using a dependent two-sided t-test ($1-\beta=.95$, $\alpha=.05$) resulting in a minimum sample size of N=327 by country. The sample size for the regressions, used for the effect of

age on effort moralization was computed using the *pwrss*-package (Bulus, 2023)(Bulus, 2023) (N = 76 per country). Both computations are documented in the supplemental material. We decided to oversample the number of complete cases to 340 by country to compensate for possible exclusions (see Data cleaning).

Data collection

As the study aimsed to test for age effects, we tried to reach approximately equal cell sizes within each country by the following branches: $18 - 30\frac{1}{12}$, $31 - 45\frac{1}{12}$, $46 - 60\frac{1}{12}$, and > 60. If one cell was not filled after three weeks of data collection, the next highest cell was oversampled to the by-country sampling goal. We recruited participants via the platform BeSample in Germany and Mexico and complete the cell sizes through Prolific. Individuals participating via Besample were compensated depending on the respective country's income levels (\$1.13 for Germany, and \$\$\$0.25 for Mexico). Participants who took part via Prolific received a compensation of £0.5 for completing our study.

Data cleaning

We applied a series of measures to ensure high data quality. Participants, indicating a respective language proficiency level below 'very good' (Germany: German, Mexico: Spanish) were excluded from participation in the study as well as participants who indicated to not currently live in the respective target country. Participants, failing one of the two attention checks distributed across the experiment were excluded from the analysis (labeled with AC in the materials, e.g., please choose 'describes him very well'). The chance of correctly solving both attention checks at random filling behavior was $\frac{1}{7} \times \frac{1}{7} = 2.04\%$. We excluded participants who completed the study three standard deviations (SD) faster than the average by country or who did not complete the study. There was no exclusion for slow participation. Following the

procedure by Celniker et al. (2023), we further excluded all participants who rated the low-effort condition as equally or less effortful compared to the high-effort condition as a manipulation check. The number of exclusions by reason and sample is documented in the supplemental material (https://osf.io/k3f4y/).

Samples

We collected data from two countries, where, to our knowledge, the effort moralization effect hasn't been studied (Germany and Mexico). We aimed to collect 340 complete cases per country. Table 3 gives an overview of the collected data.

 Table 3

 Overview of samples and demographic properties

	sampling period	sampled/valid	age $M(SD)$	min	max	med	f/m/o/n ¹
Germany	07.24 08.24.	574/392	43.36 (13.69)	20	75	41	216/166/6/4
Mexico	07.24 08.24.	399/290	39.13 (11.94)	18	70	39	172/116/0/2
overall	07.24 08.24.	973/682	41.56 (13.13)	18	75	41	388/282/6/6

Note. ^{1}f = female, m = male, o = other, n = no information indicated.

Procedure

After completing an informed consent form, participants were informed that they would be presented with a scenario on the following page, followed by several queen one about the actors depicted in these scenarios. The vignettes used in this study were adopted from Celniker et al. (2023, Study 6). They featured two employees, Marc and Justin, who workworking in a widget factory and haveing identical jobs. On the next page, participants read character descriptions, one of the low-effort target (Justin) and one of the high-effort target (Mark). The vignette reads as follows:

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

14

The market value for these widgets is \$4.00. Quality control inspections indicate that 96% of Justin's widgets and 96% of Mark's widgets work flawlessly, which means they can be sold. Thus, in an average hour, both Justin and Mark are able to produce \$23.04 worth of high-quality widgets. For Justin, making widgets requires minimal effort—although he works as quickly as possible, it is easy work. For Mark, making widgets requires a lot of effort—although he works as quickly as possible, it is hard work.

Participants completed separate sets of dependent measures for each target in randomized order after reading the vignette. We translated the vignettes to the respective languages using a team translation approach (Behr & Braun, 2023). We worked, working closely with native speakers of the respective languages (German and Spanish). The questionnaire was translated into each language by two independent translators, including one author of this paper. The initial translations were then thoroughly discussed in joint review sessions between the two authors and the native speakers until a consensus on the final translation was reached. To ensure not only a correct translation but also an appropriate adaptation to the target countries, we considered the choice of wording, names, currency, and product values of the respective countries.

Measures

To replicate Celniker et al. (2023) we employed identical instruments (study 6). Table 4 summarizes the employed concepts with example items and measurement anchors Study 6). Table 4 summarizes the employed concepts with example items and measurement anchors. All items were measured on 7 point scales, except for one item that asked about the deserved pay for each actor in the used scenario. For this item, participants responded on a sliding scale, anchored at a midpoint that was based on a realistic average salary in the respective countries where we conducted our study. For estimating realistic salaries in the target countries we relied

on data shared on the webpage of the ERI Economic Research Institute (https://www.erieri.com).

Overview of measures

Table 4



Construct	N items	Example item	Low anchor	High anchor
core goodness ^a	6	honest	does not describe X well	describes X extremely well
value commitment ^a	7	dedicated	does not describe X well	describes X extremely well
competence ^a / warmth ^a	2	competent	does not describe X well	describes X extremely well
deserved salary ^a	1	How much do you think X should make per hour?	Germany: 6 €; Mexico: \$30	Germany: 18 €; Mexico: \$90
perceived effort ^b	1	How much effort do you think X puts into his work?	no effort at all	a lot of effort
quality of work ^c	1	What quality of widgets do you think X produces?	very low quality	very high quality
job difficulty ^c	1	Compared to other jobs, how difficult is X's job?	not at all difficult	extremely difficult
work value ^c	1	How valuable do you think X's work is?	not valuable at all	extremely valuable

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.

. All items were measured on 7-point scales, except for one item that asked about the deserved pay for each actor in the used scenario. For this item, participants responded on a sliding scale, anchored at a midpoint based on a realistic average salary in the respective

countries where we conducted our study. For estimating realistic salaries in the target countries we relied on data shared on the webpage of the ERI Economic Research Institute (https://www.erieri.com).

Data analysis

Replication of effort moralization effect

To test the signal consistency of the effort moralization effect between the original study and the replication, we conducted mean comparisons across the entire sample (N = [add in Stage]) and by country. While having directional assumptions for effects in perceived morality (higher effort: higher morality and higher deservingness), prior research has shown between-country variance (e.g. Celniker et al., 2023, Study 2a-c). We therefore aimed for a considerably lower effect size, to reach adequate sensitivity (d = 0.20). To quantify the results, we computed Cohen's d with its respective 95% confidence interval as well as the log-transformed Bayes Factor.

Evaluation of replication

We used the criteria by LeBel et al. (2019) with the original effect size of d = 0.42 for core goodness and d = 0.76 for value commitment (Celniker et al., 2023, p. 73, right column) as a reference. The criteria include are constituted by the dimensions signal (was a significant result detected?), consistency (is the original effect size within the confidence interval of the current estimate?), and direction (is the effect smaller, larger, or opposite?) $\frac{2}{3}$?).

Extension to age as a predictor of effort moralization

 $^{^2}$ s = signal, ns = no signal; c = consistency, nc = no consistency; sm = smaller, la = larger, op = opposite.

To test the hypothesis that effort moralization is an age-dependent effect with possible variations between countries, we ran a series of regression models (overall and by country) with the difference of moral judgment between the vignettes by participant as dependent variable (Δ high effort, low effort), predicted to ge (linear and quadratic term) as a continuous measure. To quantify the evidence, we report the adjusted R^2 as well as the log-transformed Bayes Factor, compared to the null model (for the linear model) and against the linear model (for the quadratic model). Additionally, we ran an exploratory random-effects multi-level model, including fixed effects interactions of country and age⁽²⁾ vell as random intercepts for country along random slopes for age.

Summary of hypotheses

Table 5 summarizes the key hypotheses of the current Study. Note that it does not include assumptions for perceived warmth and perceived competence, as prior Studies showed incoherent results. Further, we haved no specific hypothesis on the superiority of a non-linear quadratic model above the linear model. Hence, the table only includes the hypothesis, that both models outperform the null model.

Specific hypotheses tested

Table 5

ID	hypotheses
1	esting more effort will be judged as more moral by participants
2	investing more effort leads to judgment of higher pay deservingness
3	age predicts the effort moralization effect positively

We haved no a priori assumptions on between-country differences and hence expected the same effect in each country as well as in the overall sample.

Results

Manipulation checks and exclusion criteria

Out of the initial 973 participants, N = 682 participants remained in the final sample. Country-specific sample sizes can be retrieved from Ω 3. Across samples, there was no significant difference in perception of work quality (p = .220, d = 0.047, 95% CI [-0.028, 0.122], BF₁₀ = 0.091). Higher effort was associated with a higher ranking of job difficulty (p < .001, d = 0.649, 95% CI [0.566, 0.731], BF₁₀ > 100,000), and a small effect on perceived value of the work (p = .001, d = 0.124, 95% CI [0.048, 0.199], BF₁₀ = 7.417).

Replication of effort moralization effect

For the moral dimension of core goodness, we found a significant effect in the expected direction in the Mexican and German samples and the pooled data for both countries. Only the observed effect in the German sample was consistent with the effect, reported by Celniker et al. (2023), as it fell in the confidence interval of the current effect sizes. The observed effect was smaller in the Mexican sample and the pooled data than in the original study. Table 6 illustrates these results.

 Within-subject effort moralization effect by low/high effort case (core goodness)

	M(=					
	high	low	p	d	CI_{low}	CI_{high}	BF_{10}	replication
Global sample	5.23 (1.42)	4.95 (1.40)	***	0.311	0.234	0.387	> 1,000	s-nc-sm
Germany	5.11 (1.29)	4.85 (1.18)	***	0.365	0.262	0.467	> 1,000	s-c
Mexico	5.39 (1.57)	5.08 (1.63)	***	0.278	0.161	0.395	> 1,000	s-nc-sm

Note. *** = <.001; s = signal, ns = no signal; c = consistency, nc = no consistency; sm = smaller, la = larger, op = opposite.; reference effect: d = 0.42.

Table 7



For value commitment, significant effects were observed in both samples, (see Table

7). Yet, the effect sizes were considerably smaller in each test, compared to the reported effect

by Celniker et al. (2023). All data regarding moral judgment—with means as horizontal

bars—was plotted in Figure 1 by country and high or low effort

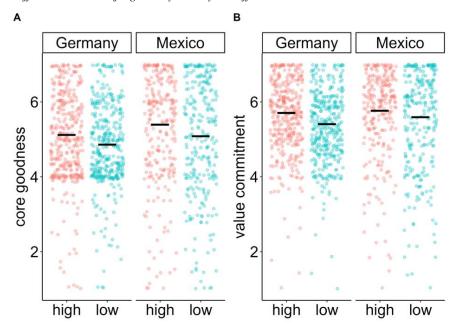
Within-subject effort moralization effect by low/high effort case (value commitment)

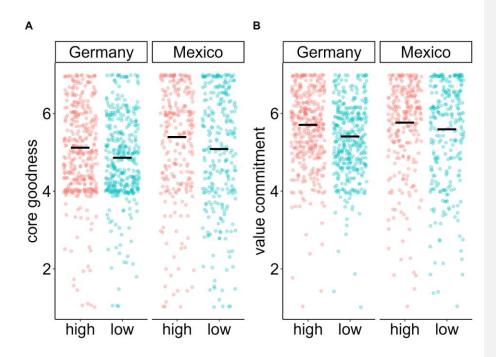
	М (SD)						
	high	low	р	d	CI_{low}	CI_{high}	BF_{10}	replication
Global sample	5.72 (1.12)	5.48 (1.16)	***	0.231	0.155	0.307	> 1,000	s-nc-sm
Germany	5.70 (1.02)	5.40 (1.00)	***	0.335	0.233	0.437	> 1,000	s-nc-sm
Mexico	5.76 (1.25)	5.58 (1.35)	.019	0.138	0.022	0.254	0.993	s-nc-sm

Note. *** = <.001; s = signal, ns = no signal; c = consistency, nc = no consistency; sm = smaller, la = larger, op = opposite.; reference effect: d = 0.76.

Figure 1

Differences in moral judgment by country and effort





Consequences of effort display on pay deservingness, warmth, and competence

Contrary to earlier reports, the deserved hourly pay did not differ between high and low effort in the German sample. Rather, the absence of a difference is supported by the BF_{10} = 0.248. Opposite to Celniker et al. (2023), Mexican participants suggested -higher pay deservingness for the low-effort individual_{τ} (see Table 8).

Table 8

Within-subject difference in pay deservingness by low/high effort case

	М (SD)					
	high	low	p	d	CI_{low}	CI_{high}	BF_{10}
Germany	14.02 (1.89)	13.95 (1.86)	.085	0.087	-0.12	0.186	0.248
Mexico	65.81 (8.96)	66.84 (8.94)	.008	-0.157	-0.272	-0.041	2.146

Note. Given the different scales of currency, no test across countries is reported, yet it is available in the supplemental material.

The effect on perceived warmth followed earlier reports, indicating an increased perception of warmth in the high-effort individual. This effect was almost twice as large in the German sample, compared to the Mexican one. (see Table 9).

Within-subject difference in perceived warmth by low/high effort case

	М ((SD)					
	high	low	p	d	CI_{low}	CI_{high}	BF_{10}
Global sample	4.69 (1.67)	4.43 (1.64)	***	0.222	0.146	0.297	> 1,000
Germany	4.73 (1.45)	4.42 (1.34)	***	0.328	0.226	0.429	> 1,000
Mexico	4.63 (1.94)	4.45 (1.97)	.027	0.130	0.014	0.246	0.733

Note. *** = <.001.

Table 9

Last, strong evidence was found for differences in perceived competence for the loweffort person. Both samples perceived the person who needed less effort for the same task as more competent with a medium-sized effect, (see Table 10).

Table 10

Within-subject difference in perceived competence by low/high effort case

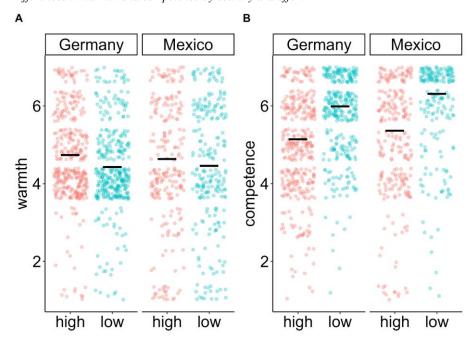
	М (SD)					
	high	low	p	d	CI _{low}	CI _{high}	BF_{10}
Global sample	5.23 (1.50)	6.12 (1.17)	***	-0.573	-0.653	-0.492	> 1,000
Germany	5.13 (1.40)	5.98 (1.13)	***	-0.602	-0.710	-0.494	> 1,000
Mexico	5.36 (1.61)	6.31 (1.20)	***	-0.546	-0.423	-0.423	> 1,000

Note. *** = <.001.

All data regarding competence and warmth—with its mean means as a-horizontal bars—was plotted in Figure 42 by country and high or low effort

Figure 2

Differences in warmth and competence by country and effort



Extension to age as a predictor of effort moralization

The extension aimed to test whether the effect of effort moralization would vary as a function of age. Our initial hypothesis was that younger participants would show lower effort moralization, as the invested effort did not lead to improved performance on the task. For core goodness, no such effect was observed. Neither as a linear nor as a quadratic function did age 23

Formatiert: Zeilenabstand: Doppelt

predict lifference in moral judgment between high and low effort. These results are substantiated in Table 11.

 Table 11

 Explanatory value of age on effort moralization effect (core goodness)

-	β	95% CI	p	R ² adj.	BF ₁₀
Global sample					
age	04	1203	.272	.000	0.070
age^2	.04	0411	.344	.000	0.060
Germany					
age	02	1208	.729	002	0.053
age^2	03	1308	.638	004	0.056
Mexico					
age	06	1805	.294	.000	0.102
age ²	.07	03 – .18	.179	.003	0.146

For value commitment, we observed no significant effect in the Mexican sample and a small linear trend in the German sample, which pointed in the opposite direction of our hypothesis. Contrary to our prediction, we observed a small decrease in effort moralization as a function of age. Table 12 illustrates these results.

Table 12

Explanatory value of age on effort moralization effect (value commitment)

	β	95% CI	p	R ² adj.	Log(BF ₁₀)
Global sample					
age	05	1203	.197	.001	0.088
age age ²	03	1004	.393	.001	0.055
Germany					
age	13	2303	.010	.014	1.451
age age ²	07	1804	.207	.016	0.112
Mexico					
age	.01	1013	.800	003	0.060
age age ²	.03	0813	.625	006	0.066

The difference in effort moralization was plotted across age by moral dimension and country.

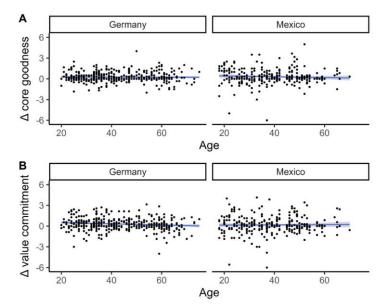
Even though the effect reached significance for value commitment in the German sample (Figur 2Figure 3, Panel B), it is evident that the difference by age was small, which is substantiated by the small $BF_{10} = 1.451$.

Formatiert: Einzug: Erste Zeile: 0 cm

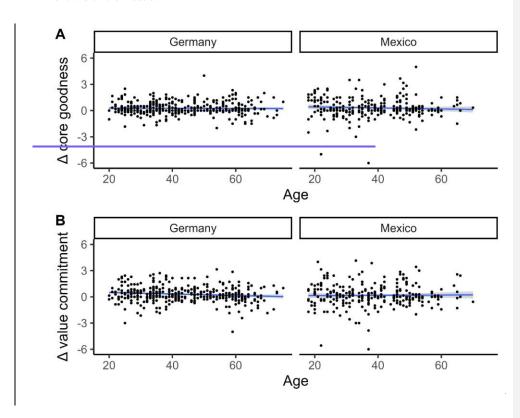
Figure 2

Figure 3

Effort moralization by dimension and country across age



Formatiert: Zeilenabstand: Doppelt



Discussion

Summary

We aimed to replicate and extend prior work on the effort moralization effect. This effect describes the attribution of moralization of moralization to current public debates, often termed as 'the great resignation' or 'quiet quitting'. These describe phenomena of individuals leaving the effort moralization effect. This effect describe phenomena of individuals leaving the workforce or reducing their amount of work.

The effort moralization effect offers a novel perspective on these debates that often propose a lower "work ethic" for younger generations. One potential explanation for these phenomena may be that younger individuals exhibit less effort moralization of ineffective labor. Consequently, they might not perceive ineffective effort as a marker of higher morality, and thus do not engage in work efforts they do not find purposeful. Hence, the strength of the effect should be a function of age.

We were able to-replicate the effects on effort moralization by Celniker et al. (2023) in both countries, yet with mostly smaller effect sizes. Ye observed some differences in secondary measures, which we discuss below.

Although we found a small effect of age on value commitment in the German sample, which was in the opposite direction of our expectations, there was otherwise no substantial evidence for an effect of age on the effort moralization effect.

Replication evaluation

Our findings largely align with those reported in the original study by Celniker et al, (2023). Yet, most effects were smaller than the target effects. In the following section, we will present our results following the replication criteria proposed by LeBel et al. (2019).

Signal

Reling moral judgment (core goodness and value commitment), we found significant results in the expected direction in the Mexican and German samples and the pooled data for both countries. Participants rated the high-effort agent as more moral than the low-effort agent across both countries.

Consistency

While significant and pointing in the expected direction, most observed effects were smaller than those reported by Celniker et al.(2023). Only the core goodness ratings in the German sample demonstrated consistency with the previous study, falling within the confidence interval of the reported effect sizes.

Replication summary

All effects on moral judgment were successfully replicated by exhibiting significant effects in the hypothesized direction. Yet, most effects (75%) were smaller than the original effects (Celinker et al., 2023). While these results indicate generalizability of the effects to Mexico and Germany, they also illustrate considerable between-country heterogeneity in effect magnitude.

Comparing effects on warmth, competence, and pay deservingness

The effect of higher perceived warmth in the person investing high effort was consistent in Mexico and Germany and following prior reports from the US and South Korea. Interestingly, the sample from France- reported a flipped effect, perceiving the low-effort person as warmer (Celniker et al., 2023, Study 2a-c)

In line with previous studies, Mexican and German participants rated the high-effort person as less competent in their task.

Notably, the effect for pay deservingness did not align with prior studies, as it indicated no significant distinction between high and low effort in the German sample and a small, albeit significant, opposite effect in the Mexican sample, suggesting that individuals exhibiting low effort were here perceived as deserving of higher compensation.

Notably, the effect for pay deservingness did not align with prior studies, as it indicated no significant distinction between high and low effort in the German sample and a small, albeit significant, opposite effect in the Mexican sample, suggesting that individuals exhibiting low effort were here perceived as deserving of higher compensation. In the current state of the literature, we can not provide an evidence-based reason for this divergence from prior data. While differences in moral values across societies have been demonstrated (Atari et al., 2023), post-hoc explanations of the findings remain speculative at this point. Future studies could assess the replicability of this finding and test with more specific assessments potential mechanisms of these cross-cultural differences. Further, it is evident that the literature still lacks sufficient primary data from WEIRD countries to address unexpected cross-cultural differences adequately.

Extension evaluation

Extension evaluation regarding target countries

-While most effect sizes observed in this study were smaller than those reported in the original investigation conducted by Celniker et al. (2023), our findings provide further evidence supporting the generalizability of the effort moralization effect. Compared to other countries where the effect has been investigated, the effect sizes found in this study can be placed at the lower end of the spectrum. Germany exhibits effect sizes regarding core goodness and value commitment that are similar to the effects observed in France (see Celniker et al., 2023, Study 2c) but slightly smaller. In Mexico, we identified smaller effects for core goodness and substantially smaller effects for value commitment.

Extension evaluation: age as predictor of Younger people value effort moralization as much as older people

Building on current societal discourses concerning the supposedly lower work engagement of younger generations, we hypothesized a lower valuation of inefficient effort in tasks as a function of age. Following these applications, we expected a weaker effort moralization effect in younger individuals. Yet, this wasn't supported by the data. The Mexican sample showed no effect of age (linear and quadratic) on effort moralization in both dimensions, with substantial evidence for the H₀. The same applied to the pooled data and the German sample, apart from a very small negative effect in value commitment. Yet, this effect pointed in the opposite direction (smaller effect in older participants).

In conclusion, the data did not provide substantial evidence of generational differences in effort moralization. Instead, the effect appeared to be age—invariant. This finding suggests that, in contrast to narratives positing discrete generational variations in work attitudes, the

inclination to moralize effort may represent a consistent bias that persists beyond age-related shifts in work values or social trends.

Theoretical Implications

The present findings carry several theoretical implications. While it was hypothesized that age predicts effort moralization, contrary evidence was found. Conversely, Bayes factor model comparison yielded moderate to strong evidence of a null age effect. As outlined, it was hypothesized that if shifts in the labor market and work values (see "great resignation" or "quiet quitting") were attributable to increased aversion towards ineffective effort in younger individuals, lower effort moralization should be observed compared to older icipants. Yet, changes depend on changes in labor supply, Covid-19, and pears more likely that the other macro phenomena, or not yet identified psychological mechanisms, instead of reduced effort moralization. The present findings suggest the generalizability of effort moralization across generations, constituting a stable bias across ages. Yet, the narrative of the effortaversive youth did not find support in our data. In addition, the present findings expand the generalizability of the effort moralization effect to two additional countries and across age demographics. This supports Celniker and colleagues' (2023) statement that effort moralization is a profoundly anchored heuristic. It appears to serve as a normative signal for evaluating character as well as intentions regarding cooperation, both inter-culturally and ageinvariant.

Limitations

The use of a single, vignette-based scenario involving factory work provided a controlled context for the assessment of the effort moralization effect. However, effort moralization may vary across different contexts or tasks, particularly those requiring different

skill sets or carrying diverse social values (e.g., caregiving roles versus technical professions). This contextual constraint might limit the generalizability of our findings to real-world situations where multiple, complex factors often inform judgments of effort. To address this limitation, future studies could incorporate multiple scenarios across varied work domains, enhancing external validity.

Further, the design of our study presented minimal information to participants, which facilitated comparisons between -conditions. Nevertheless, we acknowledge that participants may have encountered difficulties forming well-founded judgments with the limited contextual information. This minimalism—___while a central aspect of our study's design—__potentially introduces limitations in external validity. Participants in naturalistic settings may utilize additional context cues when evaluating the moral worth of effort. Consequently, our results should be interpreted within the controlled scope of the current design.

Future directions

The present project highlighted that the effort moralization effect is replicable and generalizable but shows considerable heterogeneity in magnitude between countries. Future research should aim to identify variables driving these differences. Further, it appears justified to test whether larger effects are associated with country-level differences in policy and economic development, as suggested by the recent work of Chen et al. (2024).

Furthermore, while our findings indicate that effort moralization is relatively stable across generations, a longitudinal approach could help verify this across different historical or economic periods. Longitudinal research could investigate whether macro-level factors, such as economic recessions or shifts toward automation, alter the perceived moral value of effort across age groups over time.

Meanwhile, the effort moralization effect warrants further investigation to better understand its boundary conditions and mechanics. For instance, most evidence is centered around male or gender-neutral agents, neglecting the potential impact of gender stereotyping on effort perceptions at the workplace. Further, its generalizability to more (work) contexts remains open for future research.

Last, future projects should aim to understand the role of individual differences in the strength of the bias within individuals. Personal moral beliefs and effort-directed attitudes could be candidate variables here.

Conclusion

The effort moralization effect appears to be a bias in social judgment, which is here to stay. The current project replicated the effect in countries—where the effect hadn't been—studied before and found consistent (yet mostly smaller) results. Building on this strengthened basis of evidence, the effect can be used to study different pop—ons, variations, and contexts in future projects. Further, the present study didn't deliver evidence that the effect works as a positive function of age. Rather, most analyses indicate the absence of variation—is age absence of variation across age. Given the current data, we find evidence that contradicts the assumption that younger individuals place less value on effort compared to older populations. Consequently, these presumed differences cannot be regarded as valid reasons for observed shifts within the labor market or corporate behavior.

Acknowledgments

We thank Gema Souvenir and Hilde Depauw for their help with translating the materials as well as Celine Bencker for her feedback on an earlier version of the draft.

Ethics

The study was approved by the Departmental Review Board (DRB) of the Faculty of Psychology, Department of Occupational, Economic, and Social Psychology (2024/M/001).

Conflict of interest

The authors report no conflict of interest

Monetary Support

The study was supported by the University of Ghent, Belgium (received by TT), and BeSample (received by LR).

References

- Atari, M., Haidt, J., Graham, J., Koleva, S., Stevens, S. T., & Dehghani, M. (2023). Morality beyond the WEIRD: How the nomological network of morality varies across cultures.

 Journal of Personality and Social Psychology, 125(5), 1157–1188.

 https://doi.org/10.1037/pspp0000470
- Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans.

 *Evolution and Human Behavior, 34(3), 164–175.

 https://doi.org/10.1016/j.evolhumbehav.2013.02.002
- Behr, D., & Braun, M. (2023). How Does Back Translation Fare Against Team Translation?
 An Experimental Case Study in the Language Combination English–German. *Journal of Survey Statistics and Methodology*, 11(2), 285–315.
 https://doi.org/10.1093/jssam/smac005
- Berry, Z., & Lucas, B. J. (2022). How Much Is Enough? The Relationship Between Prosocial Effort and Moral Character Judgments. *Personality and Social Psychology Bulletin*, 01461672221135954. https://doi.org/10.1177/01461672221135954
- 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
- Bulus, M. (2023). pwrss: Power and sample size calculation tools. R package version 0.3. 1.
- 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

- Chen, A. X., Sun, S., & Yu, H. (2024). Moral attitudes towards effort and efficiency: A comparison between American and Chinese history. *Humanities and Social Sciences Communications*, 11(1), 1085. https://doi.org/10.1057/s41599-024-03603-3
- Chillakuri, B. (2020). Understanding Generation Z expectations for effective onboarding.

 Journal of Organizational Change Management, 33(7), 1277–1296.

 https://doi.org/10.1108/JOCM-02-2020-0058
- Doris, J. M., & The Moral Psychology Research Group, T. M. P. R. G. (2010). *The Moral Psychology Handbook*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199582143.001.0001
- Everett, J. A. C., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, *145*(6), 772–787. https://doi.org/10.1037/xge0000165.supp
- 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
- Fengyan, W. (2004). Confucian thinking in traditional moral education: Key ideas and fundamental features. *Journal of Moral Education*, *33*(4), 429–447. https://doi.org/10.1080/0305724042000327984
- Formica, S., & Sfodera, F. (2022). The Great Resignation and Quiet Quitting paradigm shifts:

 An overview of current situation and future research directions. *Journal of Hospitality Marketing & Management*, 31(8), 899–907.

 https://doi.org/10.1080/19368623.2022.2136601
- Gervais, W. M. (2011). Finding the Faithless: Perceived Atheist Prevalence Reduces Anti-Atheist Prejudice. Personality and Social Psychology Bulletin, 37(4), 543–556. https://doi.org/10.1177/0146167211399583

- Grizzard, M., Huang, J., Fitzgerald, K., Ahn, C., & Chu, H. (2018). Sensing Heroes and Villains: Character-Schema and the Disposition Formation Process. *Communication Research*, 45(4), 479–501. https://doi.org/10.1177/0093650217699934
- Horberg, E. J., Kraus, M. W., & Keltner, D. (2013). Pride displays communicate self-interest and support for meritocracy. *Journal of Personality and Social Psychology*, 105(1), 24–37. https://doi.org/10.1037/a0032849
- Johnson, R., & Cureton, A. (2004). Kant's moral philosophy. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/kant-moral/?ref=quillette
- Kenrick, D. T., Griskevicius, V., Sundie, J. M., Li, N. P., Li, Y. J., & Neuberg, S. L. (2009).
 Deep Rationality: The Evolutionary Economics of Decision Making. *Social Cognition*,
 27(5), 764–785. https://doi.org/10.1521/soco.2009.27.5.764
- Lakens, D. (2022). Sample Size Justification. Collabra: Psychology, 8(1), 33267. https://doi.org/10.1525/collabra.33267
- Lang, H. E. (2023, November 7). It's not just a stereotype: Gen Z really does "have a work ethic problem." MarketWatch. https://www.marketwatch.com/story/its-not-just-a-stereotype-gen-z-really-does-have-a-work-ethic-problem-162faa4a
- LeBel, E., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A Brief Guide to Evaluate Replications. *Meta-Psychology*, 3. https://doi.org/10.15626/MP.2018.843
- Li, N. P., van Vugt, M., & Colarelli, S. M. (2018). The Evolutionary Mismatch Hypothesis: Implications for Psychological Science. *Current Directions in Psychological Science*, 27(1), 38–44. https://doi.org/10.1177/0963721417731378
- Markovits, D. (2019). The meritocracy trap: How America's foundational myth feeds inequality, dismantles the middle class, and devours the elite. Penguin Press.

- Medlar, A., Liu, Y., & Glowacka, D. (2022). Nobody Wants to Work Anymore: An Analysis of r/antiwork and the Interplay between Social and Mainstream Media during the Great Resignation (No. arXiv:2210.07796). arXiv. http://arxiv.org/abs/2210.07796
- Mickelberg, A., Walker, B., Ecker, U. K. H., Howe, P., Perfors, A., & Fay, N. (2022).
 Impression formation stimuli: A corpus of behavior statements rated on morality, competence, informativeness, and believability. *PLOS ONE*, 17(6), e0269393.
 https://doi.org/10.1371/journal.pone.0269393
- Mürage, A., Oyekunle, A., Ralph-Opara, U., Agada, P., Smith, J., Hawkins, K., & Morgan, R. (2022). Gendered and differential effects of the COVID-19 pandemic on paid and unpaid work in Nigeria. *Cogent Social Sciences*, 8(1), 2117927. https://doi.org/10.1080/23311886.2022.2117927
- Olivola, C. Y., & Shafir, E. (2013). The Martyrdom Effect: When Pain and Effort Increase Prosocial Contributions. *Journal of Behavioral Decision Making*, 26(1), 91–105. https://doi.org/10.1002/bdm.767
- Petersen, M. B., Slothuus, R., Stubager, R., & Togeby, L. (2011). Deservingness versus values in public opinion on welfare: The automaticity of the deservingness heuristic. *European Journal of Political Research*, 50(1), 24–52. https://doi.org/10.1111/j.1475-6765.2010.01923.x
- Petersen, M. B., Sznycer, D., Cosmides, L., & Tooby, J. (2012). Who Deserves Help?

 Evolutionary Psychology, Social Emotions, and Public Opinion about Welfare.

 Political Psychology, 33(3), 395–418. https://doi.org/10.1111/j.1467-9221.2012.00883.x
- Pillai, M. (2023). "Who am I? What am I doing?" The impact of the COVID-19 pandemic on work identities. *Employee Relations: The International Journal*, 45(6), 1527–1547. https://doi.org/10.1108/ER-01-2023-0048

- Pizarro, D. A., & Tannenbaum, D. (2012). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In *The social psychology of morality: Exploring the causes of good and evil* (pp. 91–108). American Psychological Association. https://doi.org/10.1037/13091-005
- Royle, O. R. (2024, January 6). Gen Z are over having their work ethic questioned: 'Most boomers don't know what it's like to work 40+ hours a week and still not be able to afford a house.' Fortune.Com. https://fortune.com/2024/01/26/gen-z-over-having-work-ethic-questioned-most-boomers-dont-know-work-40-hours-week-not-afford-house/
- Shepperd, J. A., Arkin, R. M., Strathman, A., & Baker, S. M. (1994). Dysphoria as a moderator of the relationship between perceived effort and perceived ability. *Journal of Personality and Social Psychology*, 66(3), 559–569. https://doi.org/10.1037/0022-3514.66.3.559
- Silverstein, M. (2008). Meeting the challenges of an aging workforce. *American Journal of Industrial Medicine*, 51(4), 269–280. https://doi.org/10.1002/ajim.20569
- Smith, J. (2022). From "nobody's clapping for us" to "bad moms": COVID-19 and the circle of childcare in Canada. *Gender, Work & Organization*, 29(1), 353–367. https://doi.org/10.1111/gwao.12758
- Telfer, E. (1989). The Unity of the Moral Virtues in Aristotle's "Nicomachean Ethics." Proceedings of the Aristotelian Society, 90, 35–48.
- Tierney, W., Hardy, J. H., Ebersole, C. R., Leavitt, K., Viganola, D., Clemente, E. G., Gordon, M., Dreber, A., Johannesson, M., Pfeiffer, T., & Uhlmann, E. L. (2020).
 Creative destruction in science. *Organizational Behavior and Human Decision Processes*, 161, 291–309. https://doi.org/10.1016/j.obhdp.2020.07.002

- Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A Person-Centered Approach to Moral Judgment. *Perspectives on Psychological Science*, 10(1), 72–81. https://doi.org/10.1177/1745691614556679
- Van Lange, P. A. M., & Kuhlman, D. M. (1994). Social value orientations and impressions of partner's honesty and intelligence: A test of the might versus morality effect. *Journal* of Personality and Social Psychology, 67(1), 126–141. https://doi.org/10.1037/0022-3514.67.1.126
- Variety (Director). (2022, October 14). Kim Kardashian's Advice for Women in Business [Video recording]. https://www.youtube.com/watch?v=Daq55GVdGJc
- Westra, E. (2021). Virtue Signaling and Moral Progress. *Philosophy & Public Affairs*, 49(2), 156–178. https://doi.org/10.1111/papa.12187
- Xueyun, Z., Al Mamun, A., Masukujjaman, M., Rahman, M. K., Gao, J., & Yang, Q. (2023).
 Modelling the significance of organizational conditions on quiet quitting intention
 among Gen Z workforce in an emerging economy. *Scientific Reports*, 13(1), Article 1.
 https://doi.org/10.1038/s41598-023-42591-3