

May 2025

The Homo Economicus Model of Work Describes Men More than Women, But Only in WEIRD Cultures

Thomas Talhelm

University of Chicago Booth School of Business

Danila Medvedev

Northwestern University Kellogg School of Management

Yin Li

Yale University School of Management

Follow this and additional works at: <https://lawandinequality.org/>

Recommended Citation

Thomas Talhelm, Danila Medvedev & Yin Li, *The Homo Economicus Model of Work Describes Men More than Women, But Only in WEIRD Cultures*, 43 L. & INEQUALITY 29 (2025).

Available at: <https://scholarship.law.umn.edu/lawineq/vol43/iss3/3>

The Homo Economicus Model of Work Describes Men More than Women, But Only in WEIRD Cultures

Thomas Talhelm, Danila Medvedev & Yin Li[†]

Introduction

Money is an obvious reason people work, but people work for many reasons that are not directly about money. People work for social approval, because of a calling, because being productive is a value in itself, and other reasons besides money. We tested several ways to encourage people to work, and we put those into two broad buckets—money and psychology. We argue that the money model of work applies more to men than women—at least in Western cultures. In contrast, gender differences are weaker (and even flip directions) in non-Western cultures.

The money bucket is the classic “homo economicus” model of work.¹ People work for their own self-interest. People work to earn money for themselves, and (in the most simplistic model) they should ignore non-monetary motivations, like social approval, judgment, or calling in life. The psychology bucket includes the social motivations that the simplest economic theory says that people should ignore—social approval, feelings of satisfaction, competition, charity, and so on. These motivations also include methods to motivate people to change their behavior in the nudge movement.

One example of a classic psychological incentive is when a University of California research team from the Department of

[†]. Thomas Talhelm is an associate professor of behavioral science at the University of Chicago Booth School of Business. Danila Medvedev is a postdoctoral research fellow at the Dispute Resolution Research Center at the Northwestern University Kellogg School of Management. Yin Li is a PhD student at the Yale University School of Management. This research was supported by a William Ladany Fellowship awarded to Thomas Talhelm. Correspondence about this article should be addressed to Thomas Talhelm, University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637. Email: talhelm@uchicago.edu. Open Data: Talhelm. Data and analysis scripts are available in the Open Science Framework. They are also available upon request from Thomas Talhelm talhelm@uchicago.edu.

1. Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis & Richard McElreath, *In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies*, 91 AM. ECON. REV. 73, 73 (2001).

Psychology sent letters to participants comparing how much power they use to their neighbors and to their neighbors who were the most energy efficient.² After receiving the letters, participants used less energy. This was despite the fact that they received no money or prize for reducing their electricity (beyond having smaller electricity bills, which everyone across conditions could always benefit from). The social norm was enough to change people's behavior.

A. A Fair Shake: Following Psychological Motivations May Make People Vulnerable

Of course, psychological motivations like following social norms can be a strategy for making money in the long run. Many rewards in life depend on social approval, such as job promotions and admissions to college. However, even if people follow psychological motivations just to earn money, they do leave people open to exploitation more than pure money motivations. Being a generous or well-liked person and expecting to be paid back eventually is a risky strategy. It requires other people to pay attention to our good behavior, remember it later, and decide to pay it back. For anyone concerned about being taken advantage of, following psychological motivations leaves people more vulnerable than the homo economicus model of money. To be sure, wage theft is real. Employers can promise workers money and then refuse to pay. Yet the shorter pay timeline makes it easier to detect dishonest bosses than with psychological motivations.

B. Is Homo Economicus Male?

We wanted to know whether there are gender differences in money versus psychology. Our goal here is not to criticize the homo economicus model. Economists since the "father of economics," Adam Smith, have recognized that humans care about things other than money.³ Instead, our goal is to ask *how well* this simple model of human behavior describes men and women across cultures.

We suspected that the homo economicus model might apply more to men than women. For one, the economics field has long had more men than women, even compared to other social science

2. See Wesley Schultz, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein & Vladas Griskevicius, *The Constructive, Destructive, and Reconstructive Power of Social Norms*, 18 PSYCH. SCI. 429, 431 (2007).

3. See ADAM SMITH, *THE WEALTH OF NATIONS* (1776) (recognizing the importance of morality and social responsibility in achieving a flourishing society).

fields.⁴ This could happen if men are more drawn to economics in the first place, or if learning about economics makes people behave more like homo economicus. Both explanations would lead to the prediction that the homo economicus model applies more to men than women.

Studies in psychology have found that people perceive women as warmer than men.⁵ Women score higher than men on tests of interpreting other people's emotions through their facial expressions,⁶ which suggests they are more sensitive. Women also have lower rates of diagnosed autism, which involves lower sensitivity to the emotions of others and social etiquette in general.⁷ These findings make it logical to think that women would respond more to psychological incentives than men.

C. Culture

There is some evidence for cultural differences in how people respond to pay-for-performance incentives. A study in the United States and China randomly assigned students to take a math test with performance incentives or without.⁸ In the incentive condition,

4. Bettina J. Casad, Christina E. Garasky, Taylor R. Jancetic, Anne K. Brown, Jillian E. Franks & Christopher R. Bach, *U.S. Women Faculty in the Social Sciences Also Face Gender Inequalities*, 13 FRONTIERS PSYCH., May 2022, at 3–4.

5. See Sabine Sczesny, Christa Nater & Alice H. Eagly, *Agency and Communion: Their Implications for Gender Stereotypes and Gender Identities*, in AGENCY AND COMMUNION IN SOCIAL PSYCHOLOGY 103 (2019) (discussing results of studies indicating that women are perceived as “more communal and less agentic” than their male counterparts); see *id.* at 111 (citing Andrea E. Abele, Nicole Haukel, Kim Peters, Eva Louvet, Aleksandra Szymkow & Yanping Duan, *Facets of the Fundamental Content Dimensions: Agency with Competence and Assertiveness – Communion with Warmth and Morality*, 7 FRONTIERS PSYCH., Nov. 2016)) (describing warmth or sociability and morality as two components of communion).

6. Simon Baron-Cohen, Sally Wheelwright, Jacqueline Hill, Yogini Raste & Ian Plumb, *The “Reading the Mind in the Eyes” Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-Functioning Autism*, 42 J. CHILD PSYCH. & PSYCHIATRY & ALLIED DISCIPLINES 241 (2001); Marcello Vellante, Simon Baron-Cohen, Mariangela Melis, Matteo Marrone, Donatella Rita Petretto, Carmelo Masala & Antonio Preti, *The “Reading the Mind in the Eyes” Test: Systematic Review of Psychometric Properties and a Validation Study in Italy*, 18 COGNITIVE NEUROPSYCHIATRY 326, 346 (2013).

7. See Meng-Chuan Lai, Michael V. Lombardo, Bonnie Auyeung, Bhismadev Chakrabarti & Simon Baron-Cohen, *Sex/Gender Differences and Autism: Setting the Scene for Future Research*, 54 J. AM. ACAD. CHILD & ADOLESCENT PSYCHIATRY 11 (2015) (addressing prior research indicating higher prevalence of autism in men and proposing a four-level conceptual framework to clarify the sex/gender differences embedded in autism research).

8. See Uri Gneezy, John A. List, Jeffrey A. Livingston, Xiangdong Qin, Sally Sadoff & Yang Xu, *Measuring Success in Education: The Role of Effort on the Test*

students received money for each question they answered correctly. In the United States, incentives worked in line with classic economic intuition. Students answered more questions correctly when they could earn money. But in China, the incentives did not boost performance. At half the schools in China, money seemed to lower performance, although the difference was not significant.

We also found evidence for cultural differences in incentives in our earlier research.⁹ We gave thousands of people across cultures a simple work task online. We paid everyone a base pay for completing the minimum requirement for the task. Then we randomly assigned some people to receive extra pay for completing more of the task. We assigned other people to receive psychological incentives, such as a social norm that suggested most people complete many trials of the task. People in individualistic cultures showed a larger “money advantage.” In other words, they worked much harder in response to the pay-for-performance than for the non-pay psychological incentives. For example, in one study, pay-for-performance boosted effort by 109% in the United Kingdom and 20% in China. In short, the homo economicus model applied more to the individualistic Western cultures than the non-Western cultures.

We also tracked whether workers quit at the first opportunity they had. We told workers they could collect their full base pay after completing ten images. In the psychological conditions, workers received no extra pay for working past ten images. In the United States, 50% of workers quit right after ten images. In Mexico, just 8% of workers did. In China, 14% quit after ten.

As one way to understand the differences, we split cultures into “WEIRD” and “non-WEIRD” cultures. This is based on the argument of Henrich and colleagues, who reviewed evidence that people in individualistic Western cultures are often outliers on psychological measures. For example, people in Western cultures tend to score on the extreme end of measures of analytic thought and offers to strangers in economic games. To describe that tendency, the researchers created the acronym “WEIRD,” which stands for Western, educated, industrialized, rich, and democratic. We use this contrast between WEIRD and non-WEIRD cultures in this study to describe cultural differences. However, we recognize that WEIRD cultures share many dimensions, such as

Itself, 1 AM. ECON. REV.: INSIGHTS 291 (2019).

9. See Danila Medvedev, Diag Davenport, Thomas Talhelm & Yin Li, *The Motivating Effect of Monetary over Psychological Incentives Is Stronger in WEIRD Cultures*, 8 NATURE HUM. BEHAV. 456 (2024).

individualism, democratic institutions, and more frequent interactions with strangers. Our goal in this study is to document cultural differences. Our data does not allow us to pull apart competing explanations for these differences beyond these broad cultural distinctions.

The difference between WEIRD individualistic cultures and non-WEIRD collectivistic cultures in motivation might explain differences in how companies structure pay around the world. Researchers have found that pay-for-performance is more common in individualistic cultures.¹⁰ One explanation is that practices tend to be adapted to the styles that fit local cultures around the world. Another possibility is that differences in the structure of pay are the cause behind why people respond differently to incentives across cultures.

The previous findings also raise questions about how people conceptualize work contracts. One way to understand contracts is that they are strict, literal, and limited to the elements that are explicitly in the contract. This view of contracts fits with the high quit rates we found in our study in the United States and United Kingdom.¹¹ The instructions said people could quit after ten images, and many of our participants in the United States and United Kingdom did just that. They followed the contract literally and maximized their pay and effort in line with the contract.

Another view of contracts is that they are ongoing relationships, flexible, and interpreted over time based on changing needs. People in some cultures may be more likely to see contracts as rough outlines. If contracts are rough outlines, people may expect each other to be forgiving if a need arises that isn't covered in the contract. And people may feel it is expected to read between the lines. That explanation fits with the behavior in our non-WEIRD cultures, like China and South Africa. People there were more likely than Americans to keep working beyond the contractual minimum. They may have thought, "They said just ten images, but they probably want me to keep working."

10. Luis R. Gomez-Mejia & Theresa Welbourne, *Compensation Strategies in a Global Context*, 14 HUM. RES. PLAN. 29, 33–35 (1991); Randall S. Schuler & Nikolai Rogovsky, *Understanding Compensation Practice Variations Across Firms: The Impact of National Culture*, 29 J. INT'L BUS. STUD. 159, 165 (1998).

11. Medvedev et al., *supra* note 9.

I. Study Overview

To test these questions, we analyzed two previous large-scale studies for gender differences.¹² In the studies, workers on crowdsourcing websites completed simple work tasks for money. Everyone received the same base pay, but some people received monetary incentives for doing extra work, and other people received psychological incentives to encourage extra work. By “monetary incentives,” we mean explicit pay-per-performance incentives. All participants received a base pay, so this is constant across conditions. Our monetary incentives are piece-rate pay beyond the base pay.

We compared how effective these monetary incentives and psychological incentives were at motivating people to work. It is important to note that there is no pure control condition in this setup. Although there is a condition that gives only base pay and no additional incentive, we think it would be inaccurate to say that this is a pure non-psychological condition. Even paying people money and asking them to do something in return involves some amount of psychological motivation, such as doing the right thing, doing a good job, or not letting down the employer. Thus, following our prior research, we calculate the psychological conditions including the base-pay-only condition.¹³

A. Study 1: Mashing Buttons for Ten Minutes

i. Methods

In Study 1, we analyzed data that two behavioral economists collected in a large-scale online study.¹⁴ They gave monetary and psychological incentives to workers on Amazon’s MTurk platform, a platform where companies or researchers can post small tasks. Workers receive small amounts of money for these micro tasks. In their study, the researchers asked participants to push the “a” and “b” buttons on their keyboard as fast as possible for ten minutes. Each time they pressed “a” and then “b,” they earned a point. We categorized their conditions into two types:

Monetary: The four monetary conditions gave participants an immediate, linear pay-per-performance incentive. For example, one condition paid participants 4 cents for every 100 points they scored.

12. *Id.*; Stefano DellaVigna & Devin Pope, *What Motivates Effort? Evidence and Expert Forecasts*, 85 REV. ECON. STUD. 1029 (2018).

13. Medvedev et al., *supra* note 9, at 458.

14. See DellaVigna & Pope, *supra* note 12.

Psychological: The seven psychological conditions gave participants no extra money beyond the base pay. To make it clear that they would not earn a piece rate, the researchers told all participants, “Your score will not affect your payment.” In the social norm condition, participants read, “Many participants scored more than 2,000.” In the “please try” condition, the researchers simply asked participants, “Please try as hard as you can.”

Table S1 and S2 describe all the conditions in more detail. This table also describes the seven conditions we did not analyze because they did not clearly fit our criteria.

Culture: One benefit of their choice of platform is that it allowed us to test for cultural differences. The original researchers did not do this intentionally, but Amazon’s MTurk platform includes a large percentage of workers in India. In the sample, 81% of workers are in the United States, but 13% are in India. In the eleven conditions we analyzed, we had a sample 6,294 participants (United States = 5,526, India = 768).

Demographics: The United States’s sample was more female (57.8%) than the sample in India (32.2%). The India sample was also more educated than the United States sample. About half (50.6%) of the United States sample had completed an undergraduate degree or were currently pursuing one, compared to 84.0% in India. In our analysis, we control for age and education.

ii. Analysis

In our main comparison of gender, we ran hierarchical linear models with participants nested in countries. We ran models using the LMER function in R:

$$\text{Points} \sim \text{Female} * \text{Monetary Incentive} + \text{Age} + \text{Education} + (1 | \text{Country})$$

When we tested the interaction between gender and culture, we ran simple regressions in this format:

$$\text{Points} \sim \text{Female} * \text{Monetary Incentive} * \text{Country} + \text{Age} + \text{Education}$$

Note that inputting the interactions automatically enters the main effects and sub-interactions for each variable. We did not nest participants in these latter models because we only had two groups, but results were similar with or without nesting.

iii. Open Data

The data is available on the Open Science Framework, along with code for processing the raw data.¹⁵ Analysis syntax for these gender analyses is available on a separate Open Science Framework page.¹⁶

iv. Results

Monetary incentives increased performance across the board ($B = 310.11$, $SE = 19.09$, $t = 16.24$, $p < .001$, from a regression without an interaction with gender). On average, the pay-for-performance conditions led to 310 more button presses. However, money was more effective for men than women (Table 1). Monetary incentives increased effort by 21% for men and 16% for women (Figure 1). Thus, money worked for both men and women, but it increased effort more for men than women.

Table 1

The Motivating Power of Money Is Stronger for Men than Women

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Female	-100.65	23.50	-4.28	< .001***
Monetary Condition	372.85	28.41	13.12	< .001***
Age	-70.96	7.36	-9.64	< .001***
College Degree	-34.69	7.81	-4.44	< .001***
Female x Monetary Condition	-114.30	38.34	-2.98	.003**
<i>N</i> = 6,294, DV = number of points in the button-pushing task				

Note: The analysis is a hierarchical linear model with participants nested in cultures. Female is coded as 0 = male, 1 = female. Monetary is coded as 0 = non-monetary (psychological), 1 = monetary. College degree is coded as 0 = no undergraduate degree, 1 = completed or currently studying for an undergraduate degree. Participants reported age in categories from 1 (18-24) to 6 (65 or

15. Danila Medvedev, Diag Davenport, Thomas Talhelm & Yin Li, *The Motivating Effect of Monetary Over Psychological Incentives Is Stronger in Western Cultures*, OSFHOME (last updated Oct. 17, 2023), <https://osf.io/8yu95/> [https://perma.cc/DBN2-7EUQ].

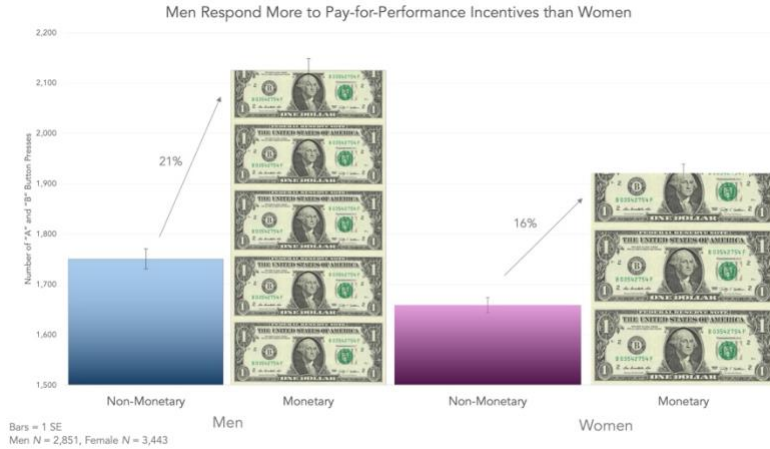
16. Thomas Talhelm, *Money Advantage - Gender Differences*, OSFHOME (last updated Dec. 17, 2024), https://osf.io/2wnvy/?view_only=171e6a4e92a946fbaac74259d6391c95 [https://perma.cc/7Z34-UJR4].

higher). A point in the button-pushing task is hitting the “A” key and then “B” key.

* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 1

Men Respond More to Pay-for-Performance Incentives than Women



Note: Men worked harder than women in response to pay-for-performance incentives in a simple button-pushing task on Amazon’s MTurk platform.

Next, we tested whether gender differences varied across cultures. The interaction between gender and monetary incentives did not differ significantly across cultures ($p = .146$, Table 2). This could be because the India sample was only 13% of the entire sample, and women only made up 32% of the India sample. That left us with fewer than 100 female participants in India in the monetary condition, which decreased our statistical power. However, looking at the averages, the U.S. pattern replicated overall patterns. Money increased effort by 24% for men and 16% for women. In India, the averages were in the opposite direction. Money increased effort by 10% for men and 14% for women.

Table 2
Interaction Between Culture, Gender, and Pay-for-Performance Incentives

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Female	-74.40	25.02	-2.97	.003**
Monetary Condition	420.73	31.34	13.42	< .001***
India	-136.01	44.32	-3.07	.002**
Age	-71.18	7.35	-9.68	< .001***
College Degree	-34.60	7.81	-4.43	< .001***
Female x Monetary	-156.06	41.18	-3.79	< .001***
Female x India	-164.54	74.91	-2.20	.028*
Monetary x India	-264.10	73.82	-3.58	< .001***
Female x Monetary x India	180.06	123.89	1.45	.146

N = 6,294, DV = number of points in the button-clicking task

Note: Female is coded as 0 = male, 1 = female. Monetary is coded as 0 = non-monetary (psychological), 1 = monetary. College degree is coded as 0 = no undergraduate degree, 1 = completed or currently studying for an undergraduate degree. Participants reported age in categories from 1 (18-24) to 6 (65 or higher).

* $p < .05$, ** $p < .01$, *** $p < .001$

v. Discussion

Our re-analysis of a large-scale study found that men responded more to pay-for-performance incentives than women. There was suggestive evidence that this gender difference only applied to the United States, not India. However, those differences were not significant, perhaps because splitting the sample across conditions, gender, and culture lowered statistical power.

B. Study 2

To test whether this was an issue of sample size, we analyzed gender differences across cultures in our larger follow-up study. This larger study gave us more statistical power to test for gender differences across cultures. In addition, our Study 2 sample included more cultures. In addition to the United States, we tested participants in the United Kingdom, China, Mexico, and South

Africa. This allowed us to test whether the findings in Study 1 were limited to India and the United States or whether the findings reflect a larger difference between WEIRD and non-WEIRD cultures. Finally, Study 2 tested whether the difference was specific to the button-pushing task. One strange feature of the button-pushing task is that it is transparently meaningless. It is possible people respond to incentives differently when work tasks are meaningless. The pointlessness of the task makes it different from most work tasks, which usually serve some purpose. Thus, we created a picture identification task and told workers that their input would help us design a machine-learning algorithm.

i. Methods

We re-analyzed data from our study of monetary incentives across cultures.¹⁷ In that study, we tested for culture differences in the effectiveness of monetary versus psychological incentives across cultures. We found that people in WEIRD cultures responded more to monetary incentives than people in non-WEIRD cultures. Here, we focus instead on gender differences. Unlike Study 1, Study 2 had an option for non-binary gender (1.8% of participants). This allowed us to contrast people who identify as male (44.3%) versus female (53.9%) directly. We could not do this in Study 1 because the previous researchers did not include a non-binary option.

ii. Sample

We analyzed the data for Study 2 from our earlier paper, which included 7,269 people from the United States, U.K., China, Mexico, and South Africa. We followed our prior study and categorized the United States and the United Kingdom as WEIRD cultures and China, Mexico, and South Africa as non-WEIRD cultures. However, results are similar if we replace the categorical WEIRD variable with a continuous variable, the Global Collectivism Index.¹⁸ Participants came from the Prolific platform for each culture except China, which does not have Prolific. Participants in China were

17. See Medvedev et al., *supra* note 9.

18. See Brett Pelham, Curtis Hardin, Damian Murray, Mitsuru Shimizu & Joseph Vandello, *A Truly Global, Non-WEIRD Examination of Collectivism: The Global Collectivism Index (GCI)*, 3 CURRENT RSCH. ECOLOGICAL & SOC. PSYCH. 1, 2022 (defining the Global Collectivism Index (GCI) GCI as an updated and globally comprehensive measure that eliminates biases ingrained in WEIRD measures and provides increased statistical power in cross-cultural research).

recruited through Hubei University and participant recruitment groups on the WeChat and QQ platforms.

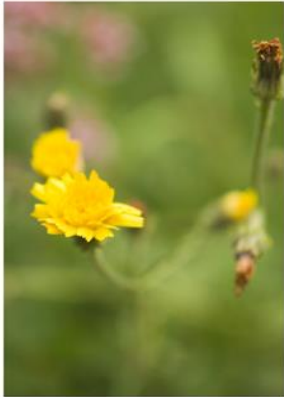
iii. Task

Our work assignment in Study 2 was an image-classification task. Participants saw images one by one and had to rate whether each image had a building or not (Figure 2). Before they started, the instructions stated that they were required to complete ten images to receive their base pay. The instructions made it clear that they would receive no extra pay for completing more than ten images. We wanted to make the task have at least some plausible purpose, unlike the transparently meaningless button-pushing task in Study 1. To give the task meaning, we told participants that we were “developing a machine learning image-classification database, and we need your help!”

Figure 2

A Sample Question in the Image-Rating Task (Study 2)

09:10



Does this image contain a building?

☐ Yes

☒ No

→

Note: This is a sample question in the image-rating task. We designed the task to be simple and avoid favoring some cultures

over others. For example, the button-pushing task in Study 1 could favor people in cultures where laptops are more prevalent. Pushing buttons quickly could be harder for participants taking the study on their phone. In contrast, the image-rating task should be more neutral.

iv. Money vs. Psychology Conditions

We randomly assigned participants to a money or psychology condition. Across all conditions, participants received a base pay for completing the ten images. In the pay-for-performance conditions, participants received extra pay for every chunk of ten images they completed. Here, we combine different samples from our original study, which included different pay conditions across rounds of data collection. One pay setup was \$1.30 base pay and an extra .05 for every ten images.

In addition to the base pay condition, we ran three psychological conditions. In the United Kingdom, China, Mexico, and some United States samples, we ran a social norm condition. In the social norm condition, participants read: “Most people try really hard on the task, assessing close to 160 images.” Participants in South Africa and some U.S. samples were assigned to conditions for competition and charity. In the charity condition, participants could earn .05 for the Red Cross for each additional ten images they completed. In the competition condition, we told participants, “This is a competition. We will tell you how well you do compared to other participants after the task.”

v. Explicit Contract

We established a clear contractual minimum for pay. Participants read these instructions:

Every ten images, you will be given an opportunity to be finished with the task. You can view as many images as you like. After the first 10 images completed, you will get the fee for your participation regardless of how many images you view or when you choose to quit the task.

To make sure participants understood the instructions, we required all participants to pass a comprehension question about the payment. We excluded participants who did not pass the comprehension question about the payment.

We also checked whether people were paying attention by asking a multiple-choice question about the purpose of the task. The correct answer was, “developing a machine-learning image-

classification database.” Ensuring that participants understood the task (and particularly the payment) was critical in understanding the differences in participants’ effort. If participants mistakenly thought they could earn more money by working harder, then it would not allow us to test differing motivations for work.

vi. Chance to Quit

After completing ten images, participants saw a page that asked them if they’d like to quit or keep working. If participants continued working, they saw that same page again every ten images. We tested for differences in (1) the percentage of workers quitting at the first possible chance (the contractual minimum) and (2) the number of total images completed.

vii. Statistical Analysis

We analyzed the data using hierarchical linear models with participants nested in cultures. For the quitting decision, we used binomial models in the GLMER function in the program R. For the number of images completed, we analyzed the data using Poisson regressions (for count data) in the GLMER function.

viii. Results

Not surprisingly, people were much more likely to quit at the first chance when we did not pay for performance (Table 3). And as we reported in our earlier paper, money made a larger difference in WEIRD cultures than non-WEIRD cultures. The homo economicus model applies better to people in WEIRD cultures than non-WEIRD cultures.

Overall, the difference between the money and psychology conditions was *larger* for women than men. In the money conditions, 13% of men quit at the first opportunity, whereas 8% of women quit at the first opportunity. Pay-for-performance cut the quit rate by 50% for men and 75% for women.

However, the pattern of gender differences was distinct across cultures (interaction between gender and non-WEIRD cultures: Table 4). In WEIRD cultures, the money advantage was larger for men than for women. Pay-for-performance cut the quit rate by 34% for men and 29% for women (Figure 3). In non-WEIRD cultures, the money advantage was larger for women than men. Pay-for-performance cut the quit rate by 6% for men and 11% for women. Thus, gender differences were not the same but smaller in non-WEIRD cultures. Instead, the direction was reversed.

Table 3

The Percentage of Workers Quitting at the First Chance After the Contractual Minimum

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Female	-0.28	0.07	-4.09	< .001***
Monetary Condition	-1.57	0.10	-15.69	< .001***
Age	-0.01	0.00	-2.58	0.010*
College Degree	0.18	0.06	2.85	0.004**
Female x Monetary Condition	-0.30	0.15	-2.04	0.041*

N = 7,269 people in 5 cultures, DV = quitting at the first chance

Note: The analysis is a hierarchical linear model with participants nested in cultures. Female is coded as 0 = male, 1 = female. Monetary is coded as 0 = non-monetary (psychological), 1 = monetary (pay-for-performance). College degree is coded as 0 = no undergraduate degree, 1 = completed or currently studying for an undergraduate degree. Participants reported their age in years.

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4

Quitting at First Chance: Gender Differences in Monetary Incentives Depend on Culture

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Female	-0.37	0.08	-4.66	< .001***
Monetary Condition	-1.78	0.11	-15.68	< .001***
Non-WEIRD	-1.82	0.38	-4.84	< .001***
Age	-0.01	0.00	-2.65	.008**
College Degree	0.18	0.06	2.80	.005**
Female x Monetary	-0.09	0.16	-0.58	0.564
Female x Non-WEIRD	0.35	0.16	2.16	.031*
Monetary x Non-WEIRD	1.05	0.23	4.50	< .001***
Female x Monetary x Non-WEIRD	-0.94	0.39	-2.43	.015*

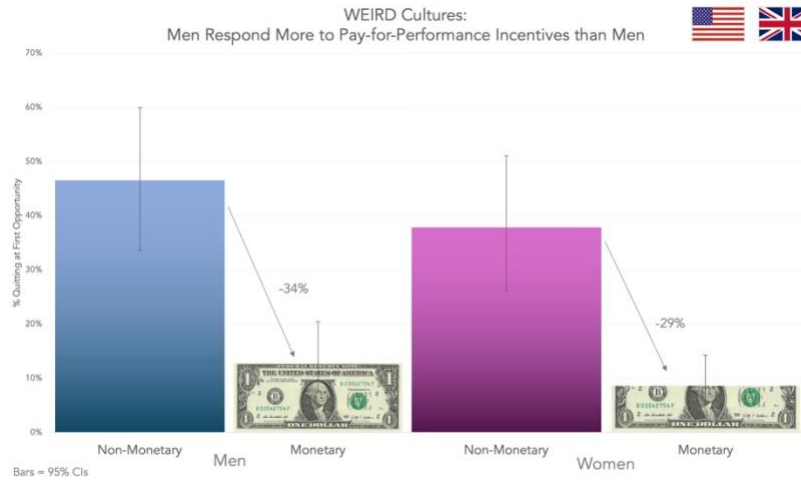
N = 7,269 people in 5 cultures, DV = quitting at first chance

Note: The analysis is a hierarchical linear model with participants nested in cultures. Non-WEIRD is coded as 0 = WEIRD, 1 = non-WEIRD. WEIRD cultures are the US and UK. Non-WEIRD cultures are China, Mexico, and South Africa. Female is coded as 0 = male, 1 = female. Monetary is coded as 0 = non-monetary (psychological), 1 = monetary (pay-for-performance). College degree is coded as 0 = no undergraduate degree, 1 = completed or currently studying for an undergraduate degree. Participants reported their age in years.

* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 3

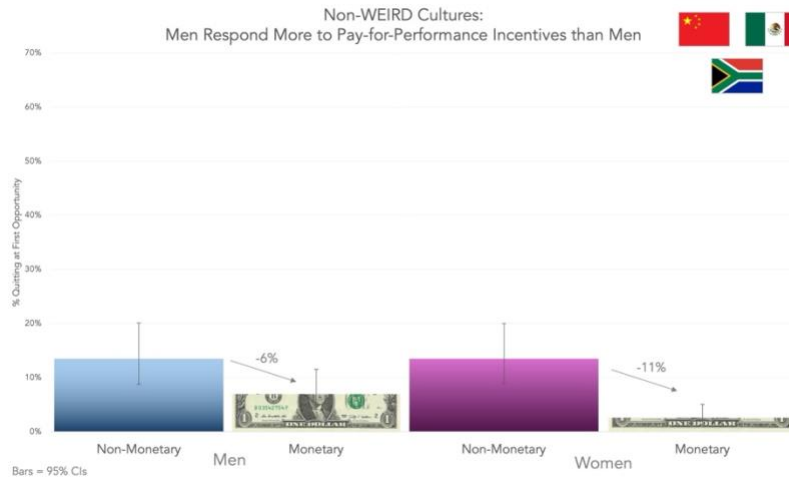
The Percentage of Workers Quitting at the Contractual Minimum in WEIRD Cultures



Note: We showed workers a screen asking if they wanted to quit the task after they completed the first ten images, which was the contractual minimum. In the monetary conditions, workers could earn more money by continuing to work past ten images. In the non-monetary conditions, workers received no extra pay for working past ten images.

Figure 4

The Percentage of Workers Quitting at the Contractual Minimum in Non-WEIRD Cultures



Note: Figures 3 and 4 present marginal means controlling for differences in education and age. The non-monetary conditions were a social norm, donation to charity, and competition.

ix. Number of Images Completed

In addition to the quit rate, we analyzed the total number of images completed. With the number of images completed, the money advantage was larger for men than women (interaction in Table 5). This is the opposite of the finding for the quit rate. It is interesting to note that the estimates from the quit rate and the number of images completed go in different directions for the overall quit rate. One reason this can happen is that the quit rate is binary, so participants have a limited range of influence on the results. But with the number of images completed, participants can range from 10 to over 200. That means an extremely motivated participant can have more influence on the averages.

Using the total number of images completed, gender differences across cultures were similar to the analysis with the quit rate (Table 6). In WEIRD cultures, pay-for-performance boosted the number of images completed 137% for men and 125% for women. In non-WEIRD cultures, pay-for-performance increased the number of images completed by 34% for men and 35% for women. Thus, the results were similar to the analysis of the quit rate. The homo

economicus model applied better to men in WEIRD cultures but not in non-WEIRD cultures.

Table 5

The Number of Images Completed in Response to Psychological and Monetary Incentives

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Female	-0.004	0.005	-0.76	0.445
Monetary Condition	0.64	0.00	132.54	< .001***
Age	-			
	0.0037	0.0002	-23.80	< .001***
College Degree	-0.12	0.004	-32.58	< .001***
Female x Monetary Condition	-0.02	0.01	-3.00	0.003**

N = 7,269 people in 5 cultures, DV = the total number of images completed

Note: The analysis is a hierarchical linear model with participants nested in cultures. Female is coded as 0 = male, 1 = female. Monetary is coded as 0 = non-monetary (psychological), 1 = monetary (pay-for-performance). College degree is coded as 0 = no undergraduate degree, 1 = completed or currently studying for an undergraduate degree. Participants reported their age in years.

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 6

Interaction Between Culture, Gender, and Pay-for-Performance Incentives

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Female	0.07	0.01	10.01	< .001***
Monetary Condition	0.90	0.01	135.10	< .001***
Non-WEIRD	0.40	0.19	2.09	0.037*
Age	-0.004	0.0002	-24.33	< .001***
College Degree	-0.11	0.0035	-31.58	< .001***
Female x Monetary Condition	-0.08	0.01	-9.47	< .001***
Female x Non-WEIRD	-0.13	0.01	-13.81	< .001***

Monetary x Non-WEIRD	-0.57	0.01	-58.23	< .001***
Female x Monetary x Non-WEIRD	0.09	0.01	6.84	< .001***

N = 7,269 people in 5 cultures, DV = number of images completed

Note: The analysis is a hierarchical linear model with participants nested in cultures. Non-WEIRD is coded as 0 = WEIRD, 1 = non-WEIRD. WEIRD cultures are the United States and the United Kingdom. Non-WEIRD cultures are China, Mexico, and South Africa. Female is coded as 0 = male, 1 = female. Monetary is coded as 0 = non-monetary (psychological), 1 = monetary (pay-for-performance). College degree is coded as 0 = no undergraduate degree, 1 = completed or currently studying for an undergraduate degree. Participants reported their age in years.

* $p < .05$, ** $p < .01$, *** $p < .001$

II. Discussion

We tested whether there are gender differences in workers' responses to psychological versus monetary incentives. In Study 1, we found that men responded more to monetary incentives than women. This suggests that the homo economicus model of work and motivation applies better to men than women. However, this effect was entirely driven by participants in the United States. In India, the pattern was flipped. Women responded more to monetary incentives than men. One limitation of Study 1 was that the sample size from India was small, which could explain why the difference across cultures was not statistically significant.

To deal with the small sample size from the non-Western culture in Study 1, we analyzed data from over 7,000 people in five cultures. With this larger sample split more evenly across WEIRD and non-WEIRD cultures, the gender differences across cultures were significant. In WEIRD cultures, men responded more to monetary incentives than women. The quit rate changed more in response to pay-for-performance more for men than women. But this pattern flipped in non-WEIRD cultures.

These findings suggest that gender differences are not hard-wired. Instead, culture shapes gender differences. In WEIRD cultures, men behave more like "economic man." But in non-WEIRD cultures, women behaved more like "economic man."

This conclusion seems to apply to WEIRD and non-WEIRD cultures outside of just a single comparison. Study 1 tested just the United States and India, which raises questions about whether the difference was specific to these two cultures. Yet Study 2 sampled people in the United States, United Kingdom, China, Mexico, and South Africa. This suggests the difference is broader than the United States and India. But we don't know whether these results apply to WEIRD and non-WEIRD cultures more broadly. As Henrich argued, "non-WEIRD" is a broad category.¹⁹ We should not assume all non-WEIRD cultures are the same as the four we tested here.

One potentially promising type of culture to look at are those cultures that farmed crops historically that did not use plows. Cultures with a history of plow use tend to have more gender inequality in the modern day.²⁰ Researchers argued this is because plows required significant upper body strength, which differs sharply between men and women. This may have led to a sharper division of labor between men and women in plow societies. In contrast, cultures that farmed non-plow crops like millet tend to have more gender equality, even after accounting for economic development. This historical difference is worth testing in future studies.

Results from an earlier study offer one way to understand why gender patterns differ across cultures. Researchers tested 80,000 participants in 76 countries around the world.²¹ They tested participants on a range of tasks, such as risk-taking, trust, and altruism. For each country, they calculated the size of gender differences. Surprisingly, gender differences were *largest* in countries with the most legal gender equality. Gender differences were also larger in economically developed countries.

One way to understand this difference is with the "seed theory." The seed theory is the idea that economic modernization gives people more freedom and resources that they can choose to use in ways that reflect some underlying group difference. One example of this comes from a study on student achievement around the world.²² Across cultures, students who reported being more

19. Joseph Henrich, *WEIRD*, OPEN ENCYC. OF COGNITIVE SCI. (July 24, 2024), <https://oecs.mit.edu/pub/spow8trw/release/1> [<https://perma.cc/BMK8-HQ5F>].

20. See Alberto Alesina, Paola Giuliano & Nathan Nunn, *On the Origins of Gender Roles: Women and the Plough*, 128 Q. J. ECON. 469 (2013).

21. See Armin Falk & Johannes Hermle, *Relationship of Gender Differences in Preferences to Economic Development and Gender Equality*, 362 SCI., Oct. 2018, at 1.

22. See Xingyu Li, Miaoze Han, Geoffrey L. Cohen & Hazel Rose Markus, *Passion Matters But Not Equally Everywhere: Predicting Achievement from Interest*,

interested in and enjoyment from learning science, math, or reading tended to score better on those tests. But that correlation varied widely across cultures. In individualistic cultures, personal interest explained 37% of the differences in test scores. But in collectivistic cultures, interest explained 16%. This offers one way to understand the pattern of gender differences around the world. More economically developed individualistic cultures may give kids more opportunities to make choices based on their preferences. But less economically developed cultures may flatten differences between kids by giving them fewer opportunities to choose based on their preferences.

Note that the seed theory does not require us to assume that gender differences are real or biological. Instead, it could be the case that people simply *believe* that there are gender differences. Then, economically developed cultures can reinforce this belief by telling stories in movies and on TV that depict gender differences. Even if no real gender differences exist in the first place, people may come to believe in gender differences and form preferences based on these beliefs. The fact that this process seems to come at the same time as legal gender equality is a surprising contradiction.

III. Supplemental Materials

A. *Conditions Included in Study 1*

When choosing which data to analyze from our earlier study, we chose to analyze data from Study 2. Study 1 was our re-analysis of the study from DellaVigna and Pope, which we already analyzed here as Study 1. We did not analyze data from Study 3 because that study used an exceptionally low monetary incentive to test the limits of the effect. We did not analyze data from Study 4 because that study randomly assigned language in India only to test the effect of culture. Thus, it would not fit with our analysis strategy in this paper.

B. *Alternatives to the Binary WEIRD vs. Non-WEIRD Binary Variable*

The WEIRD versus non-WEIRD binary variable is one way of slicing the data. However, the conclusions do not depend on this particular variable. For example, we ran robustness checks by replacing the binary WEIRD variable with continuous variables.

One alternative to the WEIRD classification is collectivism. WEIRD cultures tend to score low on collectivism, and non-WEIRD cultures tend to score high on collectivism. Thus, in one robustness check, we replaced the WEIRD variable in Table 4 with scores on the Global Collectivism Index.²³ This index uses census indicators such as the percentage of three-generation households, the percentage of people living alone, and divorce rates to estimate collectivism across cultures. Using the Global Collectivism Index, the key interaction of culture, gender, and monetary incentives predicting quitting at the first chance remained significant ($p = .012$). The full statistical output is available on the Open Science Framework page.

We also tested economic development as an alternative. One component of the WEIRD acronym is “wealthy,” and wealth is correlated with collectivism. Wealthy countries tend to score low on collectivism. Thus, we ran a model replacing the WEIRD variable in Table 4 with gross national income per capita in 2019. Again, the key interaction remained significant ($p < .001$).

In sum, these analyses suggest that the differences between cultures are robust to the method of categorizing cultures. The conclusions do not depend on the WEIRD classification. However, with only five countries, we can only draw conclusions about broad cultural differences. We do not have the fine-grained data we would need to pull apart different societal variables, such as collectivism, economic development, and rule of law.

23. See Pelham et al., *supra* note 18.

Table S1
Pay-for-Performance Versus Non-Monetary (Psychological)
Conditions from DellaVigna and Pope Included in Study 1

Pay-for-Performance Conditions	Non-Monetary (Psychological) Conditions
1 cent for 1,000 points (Self) "You will be paid an extra 1 cent for every 1,000 points."	Flat fee "Your score will not affect your payment."
1 cent for 100 points (Self) "You will be paid an extra 1 cent for every 100 points."	Unconditional Gift "In appreciation for performing this task, you will be paid a bonus of 40 cents. Your score will not affect your payment."
4 cents for 100 points (Self) "You will be paid an extra 4 cents for every 100 points."	Please Try "Please try as hard as you can."
10 cents for 100 points (Self) "You will be paid an extra 10 cents for every 100 points."	Social Comparison "We will show you how well you did relative to others."
	Social Norm "Many participants scored more than 2,000."
	1 cent for 100 points (Charity) "The Red Cross will be given 1 cent for every 100 points."
	10 cents for 100 points (Charity) "The Red Cross will be given 10 cents for every 100 points."

Note: These are the conditions from the study by DellaVigna and Pope that we analyzed for Study 1.

Table S2
Pay-for-Performance Versus Non-Monetary (Psychological)
Conditions from DellaVigna and Pope Excluded from Study 1

Excluded Condition	Reason
Risk: Low Probability of High Reward "You will have a 1% chance of an extra \$1 for every 100 points."	Bonus is not guaranteed
Risk: High Probability of Low Reward "You will have a 50% chance of an extra 2 cents for every 100 points."	Bonus is not guaranteed
2-Week Delay Discounting "You will be paid an extra 1 cent for every 100 points (2-week delay)."	Bonus is not immediate
4-Week Delay Discounting "You will be paid an extra 1 cent for every 100 points (4-week delay)."	Bonus is not immediate
40-Cent Gain "You will be paid an extra 40 cents if you score at least 2,000 points."	Bonus is not linear
40-Cent Loss "You will be paid an extra 40 cents. However, you will lose this bonus unless you score at least 2,000 points."	Bonus is not linear
80-Cent Gain "You will be paid an extra 80 cents if you score at least 2,000 points."	Bonus is not linear