

# Discussion Paper Series

IZA DP No. 18441

March 2026

## The Effects of AI Assistance on Self-Promotion

**Alexander K. Koch**

Aarhus University and IZA@LISER

**Jenny Kragl**

EBS Business School

**Sijuan Ming**

Aarhus University

**Julia Nafziger**

Aarhus University and CEPR

The IZA Discussion Paper Series (ISSN: 2365-9793) ("Series") is the primary platform for disseminating research produced within the framework of the IZA@LISER Network, an unincorporated international network of labour economists coordinated by the Luxembourg Institute of Socio-Economic Research (LISER). The Series is operated by LISER, a Luxembourg public establishment (établissement public) registered with the Luxembourg Business Registers under number J57, with its registered office at 11, Porte des Sciences, 4366 Esch-sur-Alzette, Grand Duchy of Luxembourg.

Any opinions expressed in this Series are solely those of the author(s). LISER accepts no responsibility or liability for the content of the contributions published herein. LISER adheres to the European Code of Conduct for Research Integrity. Contributions published in this Series present preliminary work intended to foster academic debate. They may be revised, are not definitive, and should be cited accordingly. Copyright remains with the author(s) unless otherwise indicated.



# The Effects of AI Assistance on Self-Promotion\*

## Abstract

Persistent gender gaps in self-promotion contribute to unequal labor market outcomes. In this study, we investigate how AI-assisted writing tools shape selfpromotion, and, as a secondary outcome, confidence and how these effects interact with gender. For this purpose, we conducted an online experiment in China in which participants wrote self-promotion texts, provided a numerical self-promotion score and stated their confidence about how they will perform in an upcoming math and logic test. We find suggestive evidence that AI assistance reduces numerical selfevaluations. Neither gender nor the interaction between gender and AI assistance is significantly related to self-promotion or confidence. We conduct a text analysis to investigate the mechanisms behind these results.

## JEL classification

C90, D03, D83, J16, M12

## Keywords

self-promotion, confidence, AI assistance, gender gaps

## Corresponding author

Alexander K. Koch

[akochoecon.au.dk](mailto:akochoecon.au.dk)

---

\* The study is pre-registered on AsPredicted (Nr. 202715) and received ethical approval from the German Association for Experimental Economics (Nr. fjC3ufY5). We are grateful to Julia Jawer, Nicola Maaser and Marina Schröder for helpful comments and discussion. Financial support from AUFF (grant number AUFF-E-2022-9-8) and DFF (grant number 10.46540/3099-00007B) is gratefully acknowledged.

---

# 1 Introduction

Despite meaningful progress towards gender equality over the past decades, persistent gaps remain across multiple domains, including wages, career advancement, political representation, and social norms, both in developed countries and globally (e.g., Eagly and Karau, 2002; Ridgeway, 2011; Duflo, 2012; Goldin, 2014; Blau and Kahn, 2017). A growing body of research seeks to understand not only explicit forms of inequality, but also implicit mechanisms, such as psychological traits and social expectations, that may reinforce existing disparities. By playing a central role in a person’s professional advancement, confidence and self-promotion are key individual-level factors contributing to unequal labor market outcomes (e.g., Rudman, 1998; Bénabou and Tirole, 2002). While promoting oneself increases individuals’ chances of being hired and their earnings, women consistently report lower confidence and are less likely to engage in self-promotion, particularly in male-typed domains such as mathematics and logic (Blau and Kahn, 2017; Exley and Kessler, 2022).

In recent years, large language models (LLMs) have increasingly been integrated as writing assistants into educational and professional settings. They offer features such as feedback, tone improvement, and motivational phrasing (e.g., Brynjolfsson et al., 2025) and they help users improve the fluency and quality of their writing (e.g., Noy and Zhang, 2023; Wiles et al., 2025). Yet, AI assistants often go beyond simple text revisions in that they can be overconfident (Sun et al., 2025a), produce outputs with a relatively positive or affirming tone (e.g., Sharma et al., 2024) and meet negativity in initial texts with softened or optimistic replies (e.g., Bardol, 2025). Such improvements in writing and increased optimism may positively affect self-promotion and confidence and, in doing so, may also affect the gender gap in self-promotion.

The aim of this paper is to investigate whether and how AI affects self-promotion and, as a secondary outcome, confidence,<sup>1</sup> and whether gender differences in self-promotion and confidence persist under the usage of AI tools. In particular, we hypothesize that AI assistance may boost user confidence and thereby enhance self-promotion efforts. We also test for the gender-related self-promotion differences documented by Exley and Kessler (2022). Finally, as women often react differently to feedback than men, we study whether AI assistance affects women differently than men.

For this purpose, we designed an online experiment that we ran in China in 2024. As most research on gender differences in various domains is based on data from the US and Western European countries, we chose China which offers a distinct socio-cultural context

---

<sup>1</sup>Confidence and self-promotion are distinct yet closely related psychological constructs that significantly influence how individuals present themselves in achievement-related contexts. Confidence generally refers to one’s belief in their own ability to succeed in a given task or domain. Self-promotion, by contrast, refers to the act of strategically presenting one’s abilities or performance to others, often in contexts where such evaluations influence real outcomes such as hiring, compensation, or promotion (Rudman, 1998). Unlike confidence, which is an internal belief about one’s abilities, self-promotion is externally directed and shaped by social and economic incentives.

(e.g., Zhang, 2013; Yi et al., 2015).

In our experiment, after having completed an initial math and logic test (a domain that is stereotypically male), participants wrote a self-promotion text about how they think they will perform on an upcoming second test and why. Then they were asked to refine that description. In the AI treatment, participants were given access to an AI tool to refine their self-promotion text; in the control, we did not explicitly encourage AI assistance. Further, participants answered three self-evaluation questions after the treatment stage: (i) a self-promotion question, where they had to indicate their agreement (on a scale of 0 to 100) with the statement, “I will perform well on the second test,” (ii) a confidence question, where they had to predict their performance on an upcoming second test, and (iii) after completing the test, participants stated their belief about the number of correct answers in the test. Finally, they completed a demographic survey. Self-promotion was incentivized: participants were explicitly informed that a potential employer would base their hiring and salary decisions on either their refined self-promotion text or their numerical self-promotion score.

Our results show that AI assistance does not significantly affect self-promotion scores, though the point estimate is negative. The confidence measures are significantly and negatively affected by AI. Thus, there is some indication that AI reduces self-evaluation scores, contrary to our hypothesis. Further, we do not find significant and consistent effects of gender or the interaction of gender with AI on self-promotion and confidence. The sign of the point estimate for the baseline gender gap in self-promotion (in the control treatment) is in line with Exley and Kessler (2022), but not significant.

We conduct a text analysis to better understand how our main results manifest in the participants’ original and refined self-promotion texts. We observe that AI assistance indeed increases text positivity and that, in the control condition, text positivity positively relates to self-promotion. Yet, when the positivity is influenced by AI, it appears to backfire and reduce confidence instead.

We discuss possible reasons for this result. First, AI-generated assessments, when perceived as external evaluations, can undermine confidence by increasing reliance on external validation rather than self-evaluation. Second, AI assistance may trigger feelings of reliance on external (technological) assistance, thereby reducing confidence.

**Literature** We contribute to the literature on confidence, self-promotion, and their interaction with AI-assisted tools. Specifically, by studying whether and how AI-generated feedback influences confidence and self-promotion and whether these effects vary across genders, we contribute insights into how generative AI may either mitigate or magnify existing gender inequalities in these domains.

Confidence influences motivation and persistence (Bandura, 1997; Bénabou and Tirole, 2002) and is, for example, a key driver for women to select out of competition (Niederle and Vesterlund, 2007; Gillen et al., 2019; Van Veldhuizen, 2022). Prior research highlights that

women, on average, report lower confidence than men, even when their actual performance is comparable (for recent studies documenting this effect, see Möbius et al., 2022; Exley and Nielsen, 2024). Such gaps persist also at the top in academic settings (Sarsons and Xu, 2021).

Related to confidence, Exley and Kessler (2022) document a gender gap in self-promotion, with women self-promoting less than equally performing men on a stereotypically male-typed task relating to math and science, even in incentivized environments. Abraham (2023) finds such effects in performance reviews, using field data. Gender differences extend beyond the numerical self-promotion or confidence score to the linguistic styles through which they are conveyed. Römer and Schröder (2025), for example, find that women tend to write more modest texts. Women also use fewer positive terms when promoting their work (Lerchenmueller et al., 2019).

Social norms and gender stereotypes are key contributing factors to these disparities in confidence and self-promotion. For example, females are characterized as more tender, shy, modest, and yielding compared to males, who are stereotypically viewed as assertive, ambitious, and dominant (e.g., Buser et al., 2014; Pan, 2019; Koch and Nafziger, 2025). These stereotypes shape both how individuals see themselves and how they are perceived by others, creating constraints on behavior and self-expression (e.g., Ellemers, 2018; Bordalo et al., 2019). As a result, females may internalize modesty norms that discourage them from overtly expressing confidence or promoting their achievements (e.g., Tradenta et al., 2025).

AI writing support can reduce skill-based inequality by leveling the playing field between more and less proficient writers (cf. Noy and Zhang, 2023; Wiles et al., 2025). In addition, as mentioned above, prior research suggests that LLMs revise texts in a positive and optimistic way. Yet, AI usage, as well as AI preference and impacts, vary by gender and context (Carvajal et al., 2024; Stöhr et al., 2024; Sun et al., 2025b), and AI-generated suggestions may sometimes feel misaligned or inauthentic (Wang et al., 2024).

Thus, despite potential advantages, concerns remain about the downsides of relying on AI tools. Sun et al. (2025a) demonstrate that although LLMs show improved accuracy on reasoning tasks, they have also more than doubled their degree of overconfidence, amplifying certainty even when the models are unsure of their answers. This tendency to produce text with a persuasive, confident tone and an authoritative style regardless of factual accuracy or personal authenticity can lead to unintended misrepresentations, especially when individuals delegate too much control over their self-descriptions to AI (Wang et al., 2024). Over-reliance on AI may also suppress the user’s original voice, leading to less nuanced and authentic narratives (Wiles and Horton, 2024).

The impact that AI has also depends on whether users are willing to trust and adopt such systems. Dargnies et al. (2024), for example, demonstrate that workers in an experiment prefer human managers rather than AI managers to make hiring decisions; and managers prefer not to delegate to AI. A possible mechanism they point to is overconfidence on the

side of the manager and beliefs about gender profiling under the AI. Normann et al. (2025) study the decision whether or not to delegate pricing decisions in a repeated Bertrand game to AI. Having the possibility to override the decision of the AI matters for delegation. Overall, trust in AI also varies across demographic lines, and gender has emerged as a key moderator of algorithm acceptance (cf. Carvajal et al., 2024).

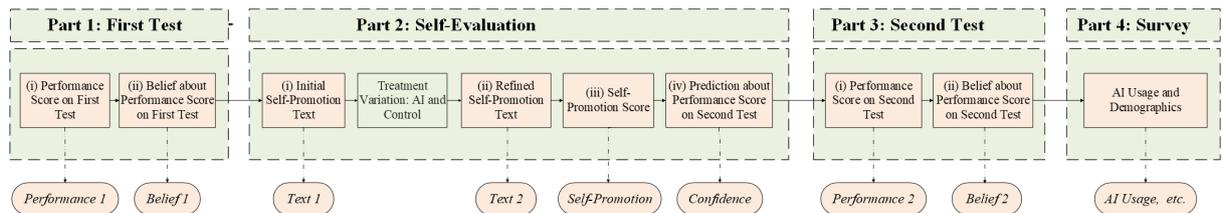
## 2 Experimental Design

We recruited 616 Chinese participants through *Credamo*.<sup>2</sup> The experiment was pre-registered (see Appendix H). Translated instructions are reproduced in Appendix G. Our experimental design comprises a main study where 600 participants engaged in self-promotion and an auxiliary employer study with 16 participants.<sup>3</sup> The employer study was used to incentivize the self-promotion measures in the main study and is described in Appendix A. To be eligible for the study, participants were required to have an approval rate of at least 80%, access the study from a Chinese IP address, and perform the study using a computer. In our analyses, we included data for all participants who completed the study.

### 2.1 Overview

Our main study comprises four parts: (1) a first test, (2) self-evaluation, (3) a second test, and (4) a follow-up survey. The flowchart in Figure 1 shows the four parts, the associated procedural steps, as well as the corresponding outcome measures at the bottom.

Figure 1: Flowchart of Main Study



*Notes:* Each dashed-line rectangle represents one of the four parts of the experiment. Each solid-line rectangle represents a procedural step with its corresponding outcome measure shown in the solid-line ovals at the bottom. Treatment variation occurs between Steps (i) and (ii) of Part 2, where participants in the AI Treatment were provided a link to an AI tool and a prompt for refining their initial self-promotion text but not participants in the control condition.

Each participant received a fixed completion fee of 9 CNY (about 1.3 USD) and a bonus for Parts 1 to 3, where one of the three parts was randomly drawn to be payoff-relevant.

<sup>2</sup>Credamo (<https://www.credamo.com/>) is a widely used platform for online surveys in China, where international platforms like Prolific or MTurk cannot be reliably accessed.

<sup>3</sup>In addition, we ran a pilot with 40 participants to create two similar math and logic tests to be used in our main study.

In the final part of the experiment, participants completed a follow-up survey collecting information on participants’ AI usage and demographics. In the following subsections, we describe the main study’s four parts in detail.

## 2.2 Part 1: First Test

Participants were asked to take a test consisting of 10 single-choice questions, with a time limit of 30 seconds per question. Participants could earn 0.2 CNY per correct answer if this part was randomly drawn as the part that counts for the bonus payment. Given the gender gaps in self-promotion that motivated our study, we selected questions related to math and logic, as gender gaps in such domains are often more pronounced in stereotypically male-typed tasks, while there are little differences in actual performance (e.g., Spencer et al., 1999; Dreber et al., 2014; Halladay and Landsman, 2022). To clearly communicate the nature of the test, we labeled it “math and logic test”. Specifically, we used Raven-like matrix problems from Matzen et al. (2010) and pre-tested the questions so that AI tools available at the time of the study were not able to solve them. The measure *Performance 1* refers to the number of questions a participant answered correctly out of the 10 questions in Step (i). After taking the test, in Step (ii), participants stated the number of questions (0-10) they believed they had answered correctly, yielding the measure *Belief 1*. To avoid possible confounds, beliefs were not incentivized.<sup>4</sup> To maintain scope for feedback from AI to have an impact on self-evaluation and to measure confidence later, participants were not informed about their actual performance at any stage of the study.

## 2.3 Part 2: Self-Evaluation

In this part of the study, participants were asked to provide different kinds of self-evaluations regarding their future performance on an upcoming second test in Part 3, which they knew would be similar to the first test conducted in Part 1. They were informed that the textual or numerical evaluations generated in the first three steps of this part would be made available to a potential “employer” in regard to hiring and salary decisions. Accordingly, these ratings reflect the participants’ willingness to self-promote themselves towards the employer.<sup>5</sup>

**Initial Self-Promotion Text (*Text 1*).** In Step (i), participants were asked to write an open-ended self-promotion text with a minimum of 60 words: “Based on your experi-

---

<sup>4</sup>Evidence suggests that non-incentivized elicitation tends to perform well (Charness et al., 2021). While incentives may reduce random noise (when more effort by the participants helps accuracy), they potentially introduce biases due to strategic misreporting of beliefs (e.g., Blanco et al., 2010; Armanier and Treich, 2013) or misunderstanding (Danz et al., 2022).

<sup>5</sup>Table F.7 shows that a higher self-promotion score indeed leads to a higher probability of being hired by the employer and a higher salary.

ence with the first test that you just took in Part 1, please describe how well you think you will perform on the second test and why.” (Translated Chinese instructions.)

**Refined Self-Promotion Text (*Text 2*) and Treatment Variation.** In Step (ii), participants were asked to refine their self-promotion texts and were told that the refined version would be treated as their final version for the employer. For this purpose, participants were provided with their initial self-promotion text and could edit it during a 5-minute period.

Participants in the CONTROL group were simply asked to submit their final text. Participants in the AI treatment did the revision with AI assistance using *DouBao*, one of the leading LLMs in China (Zhang, 2024; Chow, 2025). Specifically, they were provided with a link to DouBao and a suggested prompt for using the tool to refine their self-promotion text (see Figure G.10).

Upon completion of this step, the participants indicated whether or not they had used the suggested prompt and were asked to provide all prompts they had used while chatting with DouBao.

We use *Text 1* and *Text 2* as inputs in a text analysis, used for exploratory analysis to examine the mechanisms behind our results.

**Self-Promotion Score (*Self-Promotion*)** In Step (iii) of this part, the participants submitted a numerical self-promotion score by rating the extent to which they agreed, on a scale from 0 (entirely disagree) to 100 (entirely agree), with the statement “I will perform well on the second test.” This *Self-Promotion* measure is our main outcome variable.

Replicating the basic design features of Exley and Kessler (2022), participants were informed that either their final verbal self-promotion text or their numerical self-promotion score would be shared with another study participant, referred to as “employer”. Based on this information, the employer would then make a decision on whether to hire that participant and determine the participant’s salary if hired. That is, when making these decisions, the employer had no information on the participant’s actual performance on both tests.

The participants’ potential bonus for Part 2 was tied to the employer’s decisions. Specifically, if Part 2 was randomly selected for payment, the bonus corresponded to the salary chosen by the employer (between 0.6 CNY and 2 CNY) if the participant was hired and 0.6 CNY otherwise.

**Prediction about Performance Score on the Second Test (*Confidence*)** Finally, in Step (iv) of Part 2, the participants were asked to predict the number of questions (0-10) they will answer correctly on the second test, yielding the secondary outcome measure *Confidence*. This belief was not incentivized.

## 2.4 Part 3: Second Test

In Part 3 of the experiment, participants took a test similar to the one in Part 1, again consisting of 10 further single-choice questions, with a time limit of 30 seconds per question. The number of correct answers yields the measure *Performance 2*. Participants could earn 0.2 CNY per correct answer if this part was randomly drawn as the part that counts for the bonus payment. As before, participants were asked to provide their belief about their performance (0-10) on the second test, yielding the secondary outcome measure (*Belief 2*).

## 2.5 Part 4: Survey

In the final part of the study, we collected information on participants' demographics (including gender, education level, and age) and AI usage (e.g., asking them about their familiarity with AI platforms like DouBao).

# 3 Hypotheses and Empirical Model

This section lays out our hypotheses and the empirical model used to test them. The numerical *Self-Promotion* score is our primary outcome variable. The hypotheses for our secondary outcome variables *Confidence* and *Belief 2* are analogous. The other measures (*Belief 1*, *Performance 1* and *Performance 2*) are used for robustness and validity checks.

There is evidence that generative artificial intelligence changes sentiment when editing texts (e.g., Mak and Walasek, 2025). One reason is that model makers want users to have a positive experience and therefore tune the models to ensure that receivers perceive the output as polite and respectful (Heffner et al., 2025). But these models also reflect biases in human-generated texts that exist in the training data. Large language models can be overconfident (Sun et al., 2025a) and they often produce outputs with a relatively positive or affirming tone that align with user beliefs (e.g., Sharma et al., 2024). Negativity in user prompts is met with softened or optimistic replies, while such dampening or reversal does not usually occur for positive or neutral prompts (e.g. Bardol, 2025). In our setting, this tendency of LLMs to respond asymmetrically to the emotional tone in user prompts, could affect confidence and, in doing so, change self-promotion efforts. Thus, we expect that the AI treatment tends to provide participants with positive feedback, which could boost confidence and, in doing so, enhance self-promotion efforts.

**Hypothesis 1** *The self-promotion score is higher for participants in the AI treatment than for participants in the CONTROL treatment.*

Our second hypothesis is about replicating in our sample of Chinese participants the gender differences in self-promotion that Exley and Kessler (2022) document:

**Hypothesis 2** *Among participants in the CONTROL treatment, the self-promotion score is higher for men than for women.*

As women often react differently to feedback than men (e.g., Gill and Prowse, 2014; Berlin and Dargnies, 2016; Kogelnik, 2022), we expect that AI assistance can have a different impact on women than on men. This leads to our third hypothesis:

**Hypothesis 3** *The gender gap in self-promotion differs between participants in the CONTROL and AI treatments.*

To study the overall impact of AI assistance on the outcome variable, testing Hypothesis 1, we estimate:

$$Y_i = \alpha_0 + \alpha_1 AI_i + \gamma X_i + \epsilon_i. \quad (1)$$

$Y_i$  represents participant  $i$ 's self-promotion score (*Self-promotion*).  $AI_i$  is a binary variable that takes the value 1 if participant  $i$  belongs to the AI treatment, and 0 otherwise.  $X_i$  are controls for age, education, AI usage and employment.

To test Hypotheses 2 and 3, we extend the empirical model as follows:

$$Y_i = \theta_0 + \theta_1 Female_i + \theta_2 AI_i + \theta_3 Female_i \times AI_i + \gamma X_i + \epsilon_i. \quad (2)$$

$Female_i$  is a binary variable that takes the value 1 for female participants and 0 for men. The interaction term  $Female_i \times AI_i$  captures the difference in the gender gap between the AI and CONTROL treatments.

As a replication of Exley and Kessler (2022), we test Hypothesis 2 with the control group only. In Equation 2,  $\theta_1$  captures the gender gap in self-promotion in the control group ( $AI = 0$ ), which allows us to test Hypothesis 2 within the control condition. The impact of AI on the outcome variable is given by  $\theta_2$  for men and  $\theta_2 + \theta_3$  for women so that  $\theta_3$  reflects the differential impact of gender, testing Hypothesis 3.

Using OLS regressions, we estimate all models with and without performance-fixed effects and controls  $X_i$ , with robust standard errors.

## 4 Analysis

This section presents our quantitative analyses. We start with some descriptive statistics in Subsection 4.1 and then turn to our pre-registered main analyses, which investigate the effects of AI assistance on *Self-Promotion*, as well as on the secondary outcome variables *Confidence* and *Belief 2*.

### 4.1 Descriptive Statistics

**Sample** We stratified our sampling to achieve gender balance in both the treatment and control groups. Among the 600 participants, approximately 61% were aged 18 to 30

years, and 33% were aged 31 to 40 years. Our sample is quite educated, as over 92% of participants hold a Bachelor’s degree or higher. More than 93% of the participants reported prior experience with AI platforms, whereby the mean AI usage was significantly lower for females ( $p = .009$ ).<sup>6</sup> The sample is balanced on background characteristics and the variables recorded before Part 2, Stage (ii), when the treatment variation occurs (see Table F.1).

***Performance 1, Performance 2 and Belief 1*** Before turning to our main analysis, we examine whether there are gender gaps in beliefs and performance before the treatment took place (*Performance 1* and *Belief 1*). Specifically, pooling the data from both treatments, we find that women answered 5.37 questions correctly on average in the first test, while men answered 5.32 questions correctly, with the mean difference being statistically insignificant ( $p=.793$ , two-sided t-test). Yet, men were more optimistic about their performance than women: on average, they believed they had answered 5.89 questions correctly, whereas women believed they had answered only 5.62 questions correctly. The mean difference is borderline statistically significant ( $p = .059$ ). Finally, we observe no statistically significant effect of AI assistance on the performance score in the second test, *Performance 2* (see Table F.3). Thus, any short-term shifts in self-promotion or confidence are unlikely to stem from an effect of AI on performance.

## 4.2 Main Results

In the following, we outline our results for each hypothesis – we report them for our main outcome variable *Self-Promotion* and the secondary outcome variables *Confidence* and *Belief 2* in Table 1. The different specifications sequentially add performance-fixed effects and controls, with *Self-Promotion* as dependent variables in Panel A, *Confidence* in Panel B, and *Belief 2* in Panel C. Specifications 1-4 show the regression results based on Equation (1). Specifications 5-7 show the regression results based on Equation (2). The estimates reflect intention-to-treat effects of AI assistance, as participants in the AI treatment could modify the prompt that we provided or create their own,<sup>7</sup>

**Impact of AI on Self-Promotion (Hypothesis 1)** Contrary to Hypothesis 1, the coefficient for the AI treatment has a negative sign for Specifications 1-4 in Table 1. AI assistance has no significant effect on *Self-Promotion*, but it has a statistically significant impact on *Confidence* and *Belief 2*. Thus, overall, there is some suggestive evidence for

---

<sup>6</sup>The AI-usage variable is coded as 1 for “Not heard of”, 2 for “Heard of but not used”, 3 for “Used a few times”, 4 for “Use occasionally”, and 5 for “Use regularly”. The mean usage is 4.15 for men and 3.96 for women. See Figure E.1 for the gender-specific responses. See Table F.1 for the balance table.

<sup>7</sup>A total of 69 participants (35 females, 34 males) state to modify the given prompt or use their own. We check for the robustness of our findings by including only the participants who stuck to the prompt provided (see Table F.2).

Table 1: Regressions Results for the Three Hypotheses

	Hypothesis 1				Hypothesis 2 and 3		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Panel A: Self-Promotion as dependent variable</b>							
AI	-2.243 (1.590)	-2.210 (1.583)	-2.209 (1.582)	-2.354 (1.556)	-4.400* (2.446)	-4.182* (2.410)	-4.336* (2.370)
Female			0.109 (1.581)	0.717 (1.599)	-2.247 (2.190)	-1.865 (2.168)	-1.257 (2.172)
AI × Female					4.313 (3.180)	3.935 (3.175)	3.949 (3.175)
<b>Panel B: Confidence as dependent variable</b>							
AI	-0.250* (0.141)	-0.284** (0.140)	-0.286** (0.140)	-0.290** (0.139)	-0.140 (0.205)	-0.149 (0.204)	-0.145 (0.205)
Female			-0.197 (0.139)	-0.148 (0.140)	-0.073 (0.196)	-0.059 (0.194)	-0.003 (0.192)
AI × Female					-0.220 (0.283)	-0.274 (0.278)	-0.290 (0.276)
<b>Panel C: Belief 2 as dependent variable</b>							
AI	-0.283* (0.145)	-0.256** (0.138)	-0.255** (0.138)	-0.255** (0.137)	-0.253 (0.205)	-0.142 (0.197)	-0.136 (0.198)
Female			-0.247* (0.139)	-0.225 (0.140)	-0.220 (0.202)	-0.136 (0.199)	-0.107 (0.200)
AI × Female					-0.060 (0.289)	-0.224 (0.278)	-0.238 (0.277)
Controls				X			X
FE		X	X	X		X	X
N	600	600	600	600	600	600	600

*Notes:* Robust standard errors are reported in parentheses. \*, \*\*, \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Results are from OLS regressions based on Equation 1 (Specifications 1-4) and 2 (Specifications 5-7). *Self-Promotion* represents each participant's agreement (0-100) with the following statement "I will perform well on the second test in Part 3." *Confidence* represents the number of correct answers participants believe they can achieve on the second test in Part 3. *Belief 2* is the belief about the number of correct answers participants state after having taken the second test. *AI* is a dummy variable for the AI treatment and *Female* is a gender dummy. Performance-fixed effects for Panel A and B are dummies for each possible performance (out of the 10) on the first test, while those for Panel C are based on performance on the second test. Controls in Specifications 4 and 7 include age, AI Usage, employment, education and income.

AI to have a negative impact on self-evaluation scores.

**Gender Gap in Self-Promotion (Hypothesis 2)** For neither *Self-Promotion* (Panel A), *Confidence* (Panel B), nor *Belief 2* (Panel C) do we find a coefficient on the *Female* dummy that is statistically significant across Specifications 5-7 in Table 1. Yet, consistent with Exley and Kessler (2022), the female-indicator variable has a negative coefficient in all specifications and for all dependent variables.

**Differential Impact of AI on Self-Promotion across Genders (Hypothesis 3)** We find no statistically significant coefficient for the interaction term  $AI \times Female$  in Table 1. In addition, the signs of the coefficients on the interaction term are not consistent across the dependent variables *Self-Promotion* (Panel A), *Confidence* (Panel B), and *Belief 2* (Panel C).

## 5 Text Analysis

To better understand the results derived in the previous section, we conduct exploratory analyses of the textual self-promotion measures *Text 1* and *Text 2* (for an overview of how to conduct such text-based analyses, see, e.g., Gentzkow et al., 2019). As noted in Section 3, research has documented a positivity bias in LLMs. In addition, some evidence exists that LLM editing enhances multiple persuasive and rhetorical features of texts (Shin and Kim, 2024). To test how AI alters texts and how positivity in texts relates to our numerical self-evaluation measures, we construct three measures of text features.

### 5.1 Construction of the Text-Based Measures

First, we use a sentiment analysis of *Text 1* and *Text 2* to obtain a quantitative measure of *Positivity* of the respective text for each individual. Second, using dictionary-based methods, we count the number of *Claims* and *Justifications* contained in these texts.

**Positivity.** In Appendix B and C, we describe the technical preparation of the texts, as well as the implementation of the sentiment analysis. For our main data analysis we use the original texts written in Chinese – the results are robust to using the English translations.

To measure changes in the emotional tone of participants’ written self-evaluations before and after refinement, we applied an automated sentiment analysis pipeline based on pre-trained transformer language models (cf. Wolf et al., 2020). These models, trained on large-scale text data, assign sentiment scores (positive, neutral, and negative) to written responses. In the following, we consider being assigned a positive sentiment score to

capture the *Positivity* of self-promotion texts, where *Positivity 1 (2)* refers to the score for *Text 1 (2)*.

**Claims and Justifications.** Explicit self-claims are first-person, declarative assertions of ability or expected performance. Justificatory language consists of causal, evidential, or experiential markers used to argue for competence. The variables *Claims 1* and *Claims 2* were constructed by counting for each individual the occurrences of claims dictionary items in *Text 1* and *Text 2*, respectively. Similarly, variables *Justifications 1* and *Justifications 2* were constructed by counting the occurrences of justification dictionary items. These measures are purely linguistic and capture expressive strategies rather than underlying beliefs or true ability. Appendix D reports the details of the implementation and the complete dictionaries.

## 5.2 Analysis of Text Features

In a first step, to validate the obtained measures, we test how *Positivity*, *Claims*, and *Justifications* relate to the numerical self-assessment measures (*Self-Promotion*, *Confidence*, and *Belief 2*) in the CONTROL group. In a second step, we conduct a manipulation check – testing whether AI indeed positively shifts *Positivity*, *Claims*, and *Justifications*. In the third step, we conduct a counterfactual analysis to get at behavioral offsetting effects that could drive the main result of the empirical analysis.

**Relation between Text Features and Numerical Self-Evaluation** As shown in Table F.4, all three numerical measures (*Self-Promotion*, *Confidence*, and *Belief 2*) increase with a more positive tone in the refined self-promotion text (*Text 2*) in the CONTROL treatment. While the number of claims and the number of justifications contained in the text each have a positive coefficient, only the coefficient for claims with the dependent variable *Confidence* is significant.

**Effects of AI Assistance on Changes in Text Features** Our hypotheses were based on the premise that AI assistance tends to increase the overall positivity of revised self-descriptions. In a similar vein, we would expect that AI increases claims and justifications. As Figure E.2 illustrates, positivity, claims and justifications almost do not change when participants in the CONTROL treatment revise their self-promotion text. However, AI tends to boost both the positivity of the text as well as the number of claims and justifications. Thus, AI-generated revisions alter textual characteristics in the expected direction.

**Counterfactuals for Numerical Self-Evaluation Scores** The results that AI assistance boosts textual features and that self-evaluation scores tend to increase with textual

features in CONTROL could, in principle, imply that AI assistance boosts numerical self-evaluation scores – in contrast to our finding reported in Table 1. To understand why it does not, we conduct a counterfactual analysis. In this analysis, we predict *Self-Promotion* by multiplying *Positivity 2*, *Claims 2*, and *Justifications 2* with the respective coefficients from the baseline regression above (see Specification 4 in Table F.4). Similarly, we create counterfactuals for *Confidence*, and *Belief 2*. Table F.5 shows that *Self-Promotion*, *Confidence*, and *Belief 2* measures for the participants who had AI assistance fall significantly below the corresponding counterfactual predictions. This suggests that participants adjust their numerical self-evaluations downward relative to what textual features alone would predict, indicating a behavioral offsetting effect in the presence of AI.

Thus, summarizing, countervailing forces are at play: on the one hand, as expected, AI indeed increases positivity of texts; and positivity of texts does relate to higher self-evaluation scores in the control group. But, on the other hand, participants react negatively when positivity is generated by AI – offering a possible explanation for our main results presented in Table 1. We discuss possible explanations for this observation in the next section.

## 6 Discussion

**Results** There are several potential channels through which temporary AI assistance in writing self-promotion texts may undermine participants’ confidence in subsequent (and past) performance.

First, confidence may be reduced when participants are confronted with external, AI-generated, refined texts. In fact, a large body of psychological research argues that (even favorable) external evaluations can undermine confidence by shifting the locus of self-assessment outward and increasing reliance on external validation rather than internal self-evaluation (Ryan, 1982; Mueller and Dweck, 1998; Kluger and DeNisi, 1996; Henderlong and Lepper, 2002). This argument aligns with the observation of a higher number of claims and justifications in participants’ refined self-evaluations in the treatment condition. Participants may get suspicious about these AI-generated additions regarding their performance and become less confident accordingly.

Second, the temporary availability of AI assistance may trigger feelings of reliance on external or technological assistance. In this regard, psychological and behavioral-economics research shows that reliance on external aids or advice can undermine confidence when such support is unavailable, not because of skill loss but because success is attributed externally, reducing self-efficacy and confidence in one’s own judgment and capabilities (Bandura, 1997).

**Limitations** When interpreting the results of the study it is important to be aware of its limitations. First, as the test consists of only 10 questions, participants have relatively limited scope to meaningfully reflect on and describe their (future) performance. With such a short test, participants may have struggled to generate detailed or confident self-assessments, reducing the variability and richness of the descriptive data. As a result, the interpretation of self-evaluation scores may be constrained by the narrow scope of the task. Future research could employ longer or more complex tasks to enable participants to generate richer, more confident self-assessments.

Second, we cannot prevent the usage of AI in the control condition <sup>8</sup>. Thus, comparing the treatment and control groups corresponds to an intention-to-treat effect, but potentially not to the true treatment effect. Our study was powered to detect a 8.7pp difference in numerical self-promotion scores ( $\alpha = 0.05$ ,  $\beta = 0.8$ ), which might be too low when the treatment only captures the intention-to-treat effect.

Third, cultural factors specific to China may limit generalizability, as gender norms, acceptance of AI, and perceptions of self-promotion vary significantly across cultural contexts. That we cannot replicate the findings of Exley and Kessler (2022) in this cultural context however is interesting. Thus, expanding this study across diverse cultural and organizational contexts could help determine the cross-cultural robustness and boundary conditions of the findings.

## 7 Conclusion

This study investigates how gender and AI assistance influence self-promotion and confidence in both written self-descriptions and numerical responses, using data from an online experiment conducted in China. We find little support for self-promotion and confidence scores to differ between gender; there is some suggestive evidence for them to be affected by AI, but the effect of AI does not differ by gender. To examine possible mechanisms behind the latter result, we conduct a text analysis on the self-promotion texts. We observe that AI increases the positivity, the number of explicit claims, and to a lesser extent also the number of justifications of the revised self-promotion text. While these features are positively related to numerical self-evaluations without AI, they appear to undermine these evaluations when “imposed” by the AI because they may no longer reflect participants’ own self-assessment or preferred self-presentation strategy – creating uncertainty and in doing so undermining confidence.

---

<sup>8</sup>We tracked tab-switching behavior at each step of the experiment. In the control group, 243 participants did not switch tabs during Step (ii) (Refined Self-Promotion), while 33 participants switched tabs once (out of 300 participants in total).

## References

- ABRAHAM, L. (2023): “The Gender Gap in Performance Reviews,” *Journal of Economic Behavior & Organization*, 214, 459–492.
- ARMANTIER, O. AND N. TREICH (2013): “Eliciting Beliefs: Proper Scoring Rules, Incentives, Stakes and Hedging,” *European Economic Review*, 62, 17–40.
- BANDURA, A. (1997): *Self-Efficacy: The Exercise of Control*, New York: W. H. Freeman and Company.
- BARDOL, F. (2025): “ChatGPT Reads Your Tone and Responds Accordingly – Until It Does Not – Emotional Framing Induces Bias in LLM Outputs,” Tech. rep., arXiv.
- BÉNABOU, R. AND J. TIROLE (2002): “Self-Confidence and Personal Motivation,” *The Quarterly Journal of Economics*, 117, 871–915.
- BERLIN, N. AND M.-P. DARGNIES (2016): “Gender Differences in Reactions to Feedback and Willingness to Compete,” *Journal of Economic Behavior & Organization*, 130, 320–336.
- BLANCO, M., D. ENGELMANN, A. K. KOCH, AND H.-T. NORMANN (2010): “Belief elicitation in experiments: is there a hedging problem?” *Experimental Economics*, 13, 412–438.
- BLAU, F. D. AND L. M. KAHN (2017): “The Gender Wage Gap: Extent, Trends, and Explanations,” *Journal of Economic Literature*, 55, 789–865.
- BORDALO, P., K. COFFMAN, N. GENNAIOLI, AND A. SHLEIFER (2019): “Beliefs about Gender,” *American Economic Review*, 109, 739–773.
- BRYNJOLFSSON, E., D. LI, AND L. RAYMOND (2025): “Generative AI at Work,” *The Quarterly Journal of Economics*, qjae044.
- BUSER, T., M. NIEDERLE, AND H. OOSTERBEEK (2014): “Gender, Competitiveness, and Career Choices,” *The Quarterly Journal of Economics*, 129, 1409–1447.
- CARVAJAL, D., C. FRANCO, AND S. ISAKSSON (2024): “Will Artificial Intelligence Get in the Way of Achieving Gender Equality?” *NHH Dept. of Economics Discussion Paper*.
- CHARNESS, G., U. GNEEZY, AND V. RASOCHA (2021): “Experimental Methods: Eliciting Beliefs,” *Journal of Economic Behavior & Organization*, 189, 234–256.
- CHOW, V. (2025): “ByteDance chatbot Doubao still China’s most popular AI app as rival DeepSeek loses users,” *South China Morning Post*, accessed: 2025-09-17.

- DANZ, D., L. VESTERLUND, AND A. J. WILSON (2022): “Belief Elicitation and Behavioral Incentive Compatibility,” *American Economic Review*, 112, 2851–2883.
- DARGNIES, M.-P., R. HAKIMOV, AND D. KÜBLER (2024): “Aversion to hiring algorithms: Transparency, gender profiling, and self-confidence,” *Management Science*.
- DREBER, A., E. VON ESSEN, AND E. RANEHILL (2014): “Gender and competition in adolescence: task matters,” *Experimental Economics*, 17, 154–172.
- DUFLO, E. (2012): “Women empowerment and economic development,” *Journal of Economic literature*, 50, 1051–1079.
- EAGLY, A. H. AND S. J. KARAU (2002): “Role Congruity Theory of Prejudice toward Female Leaders,” *Psychological Review*, 109, 573.
- ELLEMERS, N. (2018): “Gender Stereotypes,” *Annual Review of Psychology*, 69, 275–298.
- EXLEY, C. L. AND J. B. KESSLER (2022): “The Gender Gap in Self-Promotion,” *The Quarterly Journal of Economics*, 137, 1345–1381.
- EXLEY, C. L. AND K. NIELSEN (2024): “The Gender Gap in Confidence: Expected but not Accounted for,” *American Economic Review*, 114, 851–885.
- GENTZKOW, M., B. KELLY, AND M. TADDY (2019): “Text as data,” *Journal of Economic Literature*, 57, 535–574.
- GILL, D. AND V. PROWSE (2014): “Gender differences and dynamics in competition: The role of luck,” *Quantitative Economics*, 5, 351–376.
- GILLEN, B., E. SNOWBERG, AND L. YARIV (2019): “Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study,” *Journal of Political Economy*, 127, 1826–1863.
- GOLDIN, C. (2014): “A Grand Gender Convergence: Its Last Chapter,” *American Economic Review*, 104, 1091–1119.
- HALLADAY, B. AND R. LANDSMAN (2022): “Perception matters: The role of task gender stereotype on confidence and tournament selection,” *Journal of Economic Behavior & Organization*, 199, 35–43.
- HEFFNER, J., C. QIN, M. CHADWICK, C. KNUTSEN, C. SUMMERFIELD, Z. KURTH-NELSON, AND R. B. RUTLEDGE (2025): “Increasing happiness through conversations with artificial intelligence,” *arXiv preprint arXiv:2504.02091*.
- HENDERLONG, J. AND M. R. LEPPER (2002): “The Effects of Praise on Children’s Intrinsic Motivation: A Review and Synthesis,” *Psychological Bulletin*, 128, 774–795.

- KLUGER, A. N. AND A. DENISI (1996): “The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory,” *Psychological Bulletin*, 119, 254–284.
- KOCH, A. K. AND J. NAFZIGER (2025): “Gender Norms, Stereotypical Beliefs, and Competitiveness,” .
- KOGELNIK, M. (2022): “Performance Feedback and Gender Differences in Persistence,” *Available at SSRN 4313281*.
- LERCHENMUELLER, M. J., O. SORENSON, AND A. B. JENA (2019): “Gender Differences in How Scientists Present the Importance of Their Research: Observational Study,” *BMJ*, 367.
- LEWIS, M., Y. LIU, N. GOYAL, M. GHAZVININEJAD, A. MOHAMED, O. LEVY, V. STOYANOV, AND L. ZETTLEMOYER (2019): “BART: Denoising Sequence-to-Sequence Pre-Training for Natural Language Generation, Translation, and Comprehension,” *arXiv preprint arXiv:1910.13461*.
- MAK, M. H. AND L. WALASEK (2025): “Style, sentiment, and quality of undergraduate writing in the AI era: A cross-sectional and longitudinal analysis of 4,820 authentic empirical reports,” *Computers and Education: Artificial Intelligence*, 100507.
- MATZEN, L. E., Z. O. BENZ, K. R. DIXON, J. POSEY, J. K. KROGER, AND A. E. SPEED (2010): “Recreating Raven’s: Software for Systematically Generating Large Numbers of Raven-Like Matrix Problems with Normed Properties,” *Behavior Research Methods*, 42, 525–541.
- MÖBIUS, M. M., M. NIEDERLE, P. NIEHAUS, AND T. S. ROSENBLAT (2022): “Managing Self-Confidence: Theory and Experimental Evidence,” *Management Science*, 68, 7793–7817.
- MUELLER, C. M. AND C. S. DWECK (1998): “Praise for Intelligence Can Undermine Children’s Motivation and Performance,” *Journal of Personality and Social Psychology*, 75, 33–52.
- NIEDERLE, M. AND L. VESTERLUND (2007): “Do Women Shy Away from Competition? Do Men Compete Too Much?” *The Quarterly Journal of Economics*, 122, 1067–1101.
- NORMANN, H.-T., N. RULIÉ, O. STYPA, AND T. WERNER (2025): “Delegate Pricing Decisions to an Algorithm? Experimental Evidence,” *arXiv preprint arXiv:2510.27636*.
- NOY, S. AND W. ZHANG (2023): “Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence,” *Science*, 381, 187–192.

- PAN, S. (2019): “The Instability of Matching with Overconfident Agents,” *Games and Economic Behavior*, 113, 396–415.
- RIDGEWAY, C. L. (2011): *Framed by Gender: How Gender Inequality Persists in the Modern World*, Oxford University Press.
- RÖMER, N. AND M. SCHRÖDER (2025): “Effects of Written Self-Promotion on Gender Bias and Decision Quality,” Tech. rep., Leibniz-Universität Hannover, Wirtschaftswissenschaftliche Fakultät.
- RUDMAN, L. A. (1998): “Self-Promotion as a Risk Factor for Women: The Costs and Benefits of Counterstereotypical Impression Management,” *Journal of Personality and Social Psychology*, 74, 629.
- RYAN, R. M. (1982): “Control and Information in the Intrapersonal Sphere: An Extension of Cognitive Evaluation Theory,” *Journal of Personality and Social Psychology*, 43, 450–461.
- SARSONS, H. AND G. XU (2021): “Confidence Men? Evidence on Confidence and Gender among Top Economists,” in *AEA Papers and Proceedings*, American Economic Association, vol. 111, 65–68.
- SHARMA, M., M. TONG, T. KORBAK, D. DUVENAUD, A. ASKELL, S. R. BOWMAN, E. DURMUS, Z. HATFIELD-DODDS, S. R. JOHNSTON, S. M. KRAVEC, T. MAXWELL, S. MCCANDLISH, K. NDOUSSE, O. RAUSCH, N. SCHIEFER, D. YAN, M. ZHANG, AND E. PEREZ (2024): “Towards Understanding Sycophancy in Language Models,” in *The Twelfth International Conference on Learning Representations*.
- SHIN, M. AND J. KIM (2024): “Large Language Models Can Enhance Persuasion through Linguistic Feature Alignment,” *Available at SSRN 4725351*.
- SPENCER, S. J., C. M. STEELE, AND D. M. QUINN (1999): “Stereotype Threat and Women’s Math Performance,” *Journal of Experimental Social Psychology*, 35, 4–28.
- STÖHR, C., A. W. OU, AND H. MALMSTRÖM (2024): “Perceptions and Usage of AI Chatbots among Students in Higher Education across Genders, Academic Levels and Fields of Study,” *Computers and Education: Artificial Intelligence*, 7, 100259.
- SUN, F., N. LI, K. WANG, AND L. GOETTE (2025a): “Large Language Models Are Overconfident and Amplify Human Bias,” *arXiv preprint arXiv:2505.02151*.
- SUN, S., Z. A. LI, M.-D. FOO, J. ZHOU, AND J. G. LU (2025b): “How and for Whom Using Generative AI Affects Creativity: A Field Experiment,” *Journal of Applied Psychology*, 110, 1561–1573.

- TRADENTA, J. M., A. NEELIM, AND J. VECCI (2025): “Gender Differences in the Self-promotion of Prosocial Behaviour: Exploring the Female Modesty Constraint,” *Experimental Economics*, 28, 1–21.
- VAN VELDHUIZEN, R. (2022): “Gender Differences in Tournament Choices: Risk Preferences, Overconfidence, or Competitiveness?” *Journal of the European Economic Association*, 20, 1595–1618.
- WANG, Q., C. L. ANYI, V. D. SWAIN, AND A. K. GOEL (2024): “Navigating AI Fallibility: Examining People’s Reactions and Perceptions of AI after Encountering Personality Misrepresentations,” *arXiv preprint arXiv:2405.16355*.
- WILES, E. AND J. J. HORTON (2024): “More, but Worse: The Impact of AI Writing Assistance on the Supply and Quality of Job Posts,” NBER Working Paper, National Bureau of Economic Research, Cambridge, MA.
- WILES, E., Z. MUNYIKWA, AND J. HORTON (2025): “Algorithmic Writing Assistance on Jobseekers’ Resumes Increases Hires,” *Management Science*.
- WOLF, T., L. DEBUT, V. SANH, J. CHAUMOND, C. DELANGUE, A. MOI, P. CISTAC, T. RAULT, R. LOUF, M. FUNTOWICZ, ET AL. (2020): “Transformers: State-of-the-Art Natural Language Processing,” in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 38–45.
- YI, X., B. RIBBENS, L. FU, AND W. CHENG (2015): “Variation in career and workplace attitudes by generation, gender, and culture differences in career perceptions in the United States and China,” *Employee Relations*, 37, 66–82.
- YUAN, L. (2021): “distilbert-base-multilingual-cased-sentiments-student,” <https://huggingface.co/lxyuan/distilbert-base-multilingual-cased-sentiments-student>, accessed: 2025-05-04.
- ZHANG, J. (2024): “TikTok Parent ByteDance Now Has China’s Most Popular AI Chatbot,” *Bloomberg*, accessed: 2024-05-20.
- ZHANG, Y. J. (2013): “Gender Differences in Competitiveness: Evidence from China,” *Economic Journal*, 123, 119–159.

# Appendix

## A Employer Experiment

In addition to the main study, we subsequently conducted an auxiliary experiment to incentivize decisions in the main experiment. Each participant acted as an “employer” and was required to make hiring and salary decisions for 20 different self-promotion scores (SP, 0-100), as submitted in Part 2(iii) of the main study. In particular, for each given score, the employer needed to decide whether to hire the respective “employee” and, if hired, decide on his or her salary.<sup>9</sup>

The data collected from this experiment were used to calculate the bonus payments for those participants for whom Part 2 was randomly selected as their bonus-determining part. In particular, if the employer chose not to hire the employee for some given *Self-Promotion* score, the employer received 2 CNY and the employee earned a basic bonus of 0.6 CNY. In case of hiring, the employer chose a salary for that employee ranging from 0.6 CNY to 2 CNY, in increments of 0.2 CNY. In that case, the employer obtained a payment, depending on both the employee’s salary and actual performance on the second test;  $(2 - salary + 0.2 \times Performance)$  CNY.

## B Data Pre-processing for the Text Analysis

To prepare text data for subsequent analysis, we implement different preprocessing pipelines for the original Chinese texts and their translated English versions, considering the linguistic characteristics and structural differences between the two languages.<sup>10</sup>

Regarding the Chinese texts, we implement a multi-step preprocessing pipeline focused on stopword removal and tokenization using *jieba*, a standard Chinese word-segmentation toolkit. First, we compile a comprehensive stopword list by merging three widely used Chinese stopword dictionaries, Baidu, Harbin Institute of Technology (HIT), and Sichuan University (SCU),<sup>11</sup> ensuring all entries are unique. This step filters out function words and other semantically uninformative tokens, improving the effectiveness of subsequent vectorization (e.g., TF-IDF). Next, we define a custom tokenizer function using *jieba.lcut*, a widely used library for Chinese word segmentation. The tokenizer removes any tokens that appear in the combined stopword set and filters out blank or trivial strings. This preprocessing routine ensures that only informative content words are retained for further

---

<sup>9</sup>This experiment is run to provide participants in Part 2 with exogenous incentives for self-promotion. In fact, Table F.7 in Appendix F verifies that higher *self-promotion* scores are significantly positively related to hiring probabilities and salary levels ( $p < .01$ ).

<sup>10</sup>Prior to preprocessing, 12 observations were manually excluded because their refined descriptions are invalid, for example, they repeat the provided prompt, are entirely irrelevant to the task, or are left blank.

<sup>11</sup>[https://github.com/xiyusullos/chinese\\_stopwords](https://github.com/xiyusullos/chinese_stopwords)

feature extraction, thus enhancing the semantic relevance of the resulting token set. Regarding the translated English texts, we use a pipeline that includes stopword removal, part-of-speech (POS) tagging, and lemmatization. In addition to the standard Natural Language Toolkit (NLTK) stopword set, we incorporate a set of common function words frequently omitted in social science applications (e.g., “just,” “also,” “very”). This augmented stopword list aims to eliminate uninformative tokens that do not contribute meaningfully to the content of the text. Next, we lowercase all text and remove punctuation using regular expressions. The cleaned text is then tokenized and POS-tagged using NLTK’s ‘pos\_tag’ function. To retain semantically meaningful forms of each token, we apply lemmatization using WordNet, mapping each word to its base form based on its POS category (e.g., verbs, nouns, adjectives). This step helps unify inflectional variants (e.g., “running” → “run”, “better” → “good”), improving feature consistency. The final output of this function is a space-separated string of lemmatized, content-rich tokens, which can be used in downstream text analysis.

## C Implementation of Sentiment Analysis

We implemented the pipeline in Python, using the HuggingFace Transformers library (Wolf et al., 2020), along with pandas for data handling and tqdm for progress tracking. For the sentiment analysis of the Chinese texts, we use the lxyuan/distilbert-base-multilingual-cased-sentiments-student model developed by Yuan (2021). This multilingual transformer is fine-tuned for three-class sentiment classification: positive, neutral, and negative. We apply the model using the TextClassificationPipeline from HuggingFace, which returns probability scores for each sentiment category. We extract and store the sentiment scores for each description. This analysis covers both the initial and refined versions of the participant descriptions.

For the English-translated texts, we apply a zero-shot classification approach using the facebook/bart-large-mnli model, developed by Facebook AI (Lewis et al., 2019). This model classifies text based on how closely it matches candidate sentiment labels, “positive,” “neutral,” and “negative”, using natural language inference. We extract and store these sentiment scores in the same format as the Chinese analysis. We choose this model for its strong performance on general language tasks and its ability to handle nuanced emotional tone without task-specific fine-tuning.

## D Construction and Implementation of Claims and Justifications

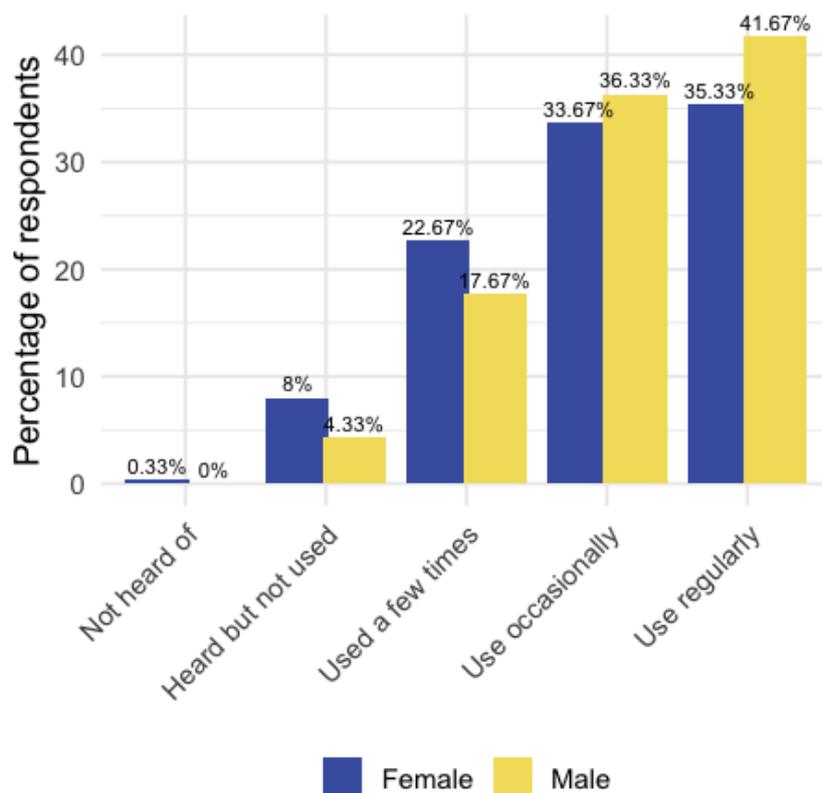
We construct dictionaries that consist of fixed lists of Chinese words and short token sequences associated with (i) explicit self-claims or (ii) justificatory framing. These dic-

tionaries were constructed in two steps. First, we used an LLM (ChatGPT 5.2) to develop an initial set of claim-related and justification-related terms based on a random sample of 100 original Chinese self-promotion texts (50 initial texts and 50 revised texts). Second, we used an LLM with an independent prompt and a separate user account to generate additional candidate words and token sequences associated with explicit self-claims and justificatory language in the context of self-promotion. These candidates were reviewed for conceptual consistency and merged with the lists constructed in the first step to obtain the final dictionaries.

Following dictionary construction, all texts were processed using a rule-based pipeline implemented in Python. Because participants frequently relied on comma-separated structures rather than sentence-final punctuation, we segmented texts into clause-like units based on Chinese punctuation and selected discourse markers. We then applied deterministic string-matching rules to identify occurrences of dictionary items within each clause. For each text, we counted the frequency of claim-related and justification-related markers. We computed within-participant changes in claim and justificatory language as the difference between the revised and initial texts (i.e., *Text 1* and *Text 2*), which serves as the key explanatory variable in subsequent analyses of *Self-Promotion*, *Confidence*, and *Belief 2*.

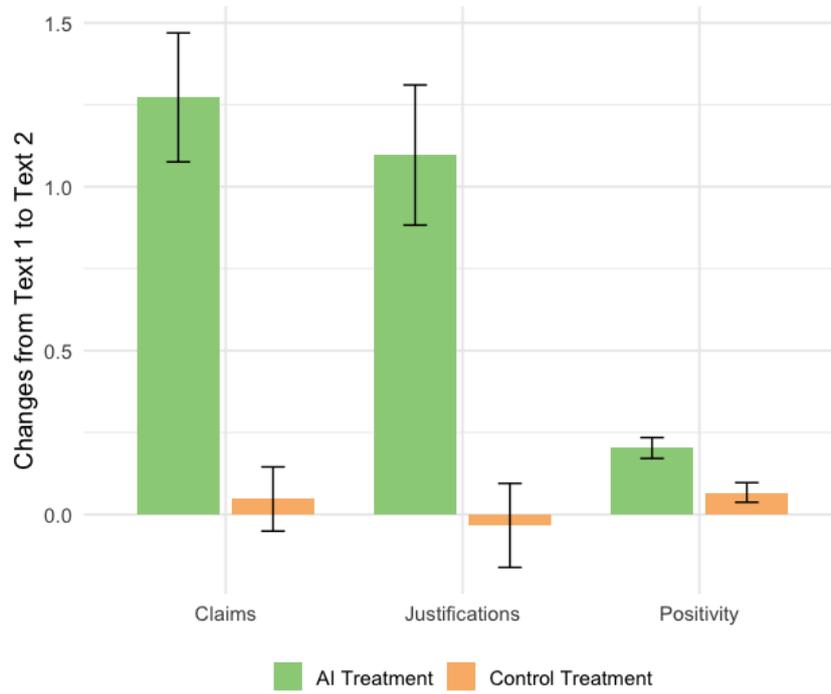
## E Additional Figures

Figure E.1: AI Usage Among Females and Males



*Notes:* The figure shows the percentage of females and males choosing each answer option to the question “How familiar are you with ChatGPT, DouBao or other similar AI platforms?” Within gender the percentages across categories add up to 100%. Data are from both treatment and control groups.

Figure E.2: Changes in Text Characteristics by Treatment Group



*Notes:* The figure shows the changes in text characteristics from *Text 1* to *Text 2* by treatment groups. *Positivity* represents the positive sentiment expressed in the self-promotion texts; *Claims* measures the number of Claims in the self-promotion texts; and *Justifications* represents the number of Justifications in the self-promotion texts. Data are from both treatment and control groups, excluding 12 participants whose *Text 2* are either empty or nonsense.

## F Additional Tables

Table F.1: Balance Table

	AI treatment	CONTROL treatment	Difference	p-value
<b>Panel A: Background characteristics</b>				
AI Usage	4.11	4.00	0.120	0.143
Age category	2.38	2.45	0.097	0.235
Female	0.50	0.50	<0.001	1.000
Edu	5.19	5.18	0.005	0.951
Employed	0.78	0.81	0.066	0.421
Income	2.31	2.32	0.008	0.920
<b>Panel B: Outcomes before treatment variation</b>				
Performance 1	5.43	5.26	0.084	0.302
Belief 1	5.74	5.77	0.017	0.838
<b>Panel C: Outcomes after treatment variation</b>				
Performance 2	5.59	5.70	0.056	0.490
Self-Promotion	80.17	82.42	0.115	0.159
Confidence	6.94	7.19	0.144	0.078
Belief 2	6.31	6.59	0.160	0.051
Text 1 Positivity	0.62	0.64	0.041	0.616
Text 1 Clauses	2.49	2.49	<0.001	0.998
Text 1 Claims	1.49	1.40	0.101	0.222
Text 1 Justifications	1.75	1.68	0.074	0.371

*Notes:* Treatment variation occurs in Part 2, Stage (ii) of the main study (see Figure 1). AI Usage categories: (1) "Not heard of", (2) "Heard of but not used", (3) "Used a few times", (4) "Use occasionally", (5) "Use regularly". Age range categories: (1) 18 – 20, (2) 21 – 30, (3) 31 – 40, (4) 41 – 50, (5) 51 – 60, (6) Above 60. Education categories: (1) "Primary school and below", (2) "Junior middle school", (3) "General high school / technical secondary school / technical school / vocational school", (4) "Training school", (5) "Bachelor degree", (6) "Master degree", (7) "Doctor degree". Employed categories: (1) "Have full-time or part-time job", (0) otherwise. Income categories (CNY): (1) < 60k, (2) 60k – 120k(exl.), (3) 120k – 180k(exl.), (4) 180k – 240k(exl.), (5) 240k – 300k(exl.), (6)  $\geq$  300k.

Table F.2: Regression Results for the Hypotheses (Restricted Sample)

	Hypothesis 1				Hypothesis 2 and 3		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Panel A: Self-Promotion as dependent variable</b>							
AI	-1.066 (1.679)	-1.054 (1.674)	-1.056 (1.674)	-1.005 (1.643)	-3.057 (2.587)	-2.849 (2.557)	-2.680 (2.515)
Female			-0.216 (1.663)	0.062 (1.672)	-2.247 (2.191)	-1.777 (2.164)	-1.387 (2.179)
AI × Female					3.990 (3.357)	3.578 (3.349)	3.339 (3.375)
<b>Panel B: Confidence as dependent variable</b>							
AI	-0.181 (0.150)	-0.217 (0.147)	-0.220 (0.146)	-0.209 (0.146)	0.041 (0.215)	0.048 (0.214)	0.078 (0.215)
Female			-0.294** (0.146)	-0.263* (0.146)	-0.073 (0.196)	-0.061 (0.194)	-0.015 (0.192)
AI × Female					-0.446 (0.298)	-0.534* (0.292)	-0.573** (0.290)
<b>Panel C: Belief 2 as dependent variable</b>							
AI	-0.261* (0.154)	-0.237* (0.148)	-0.237* (0.147)	-0.228 (0.146)	-0.157 (0.218)	-0.058 (0.211)	-0.021 (0.210)
Female			-0.305** (0.148)	-0.304** (0.148)	-0.220 (0.202)	-0.151 (0.200)	-0.126 (0.201)
AI × Female					-0.210 (0.307)	-0.356 (0.296)	-0.412 (0.293)
Controls				X			X
FE		X	X	X		X	X
N	531	531	531	531	531	531	531

*Notes:* The restricted sample drops 69 participants from the AI treatment who did not use the provided prompt for revising the self-promotion text. Robust standard errors are reported in parentheses. \*, \*\*, \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Results are from OLS regressions based on Equation 1 (Specifications 1-4) and 2 (Specifications 5-7). *Self-Promotion* represents each participant's agreement (0-100) with the following statement "I will perform well on the second test in Part 3." *Confidence* represents the number of correct answers participants believe they can achieve on the second test in Part 3. *Belief 2* is the belief about the number of correct answers participants state after having taken the second test. *AI* is a dummy variable for the AI treatment and *Female* is a gender dummy. Performance-fixed effects for Panel A and B are dummies for each possible performance (out of the 10) on the first test, while those for Panel C are based on performance on the second test. Controls in Specifications 4 and 7 include age, AI Usage, employment, education and income.

Table F.3: Regressions Performance 2 as Dependent Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AI	-0.11 (0.159)	-0.183 (0.142)	-0.184 (0.142)	-0.179 (0.142)	-0.267 (0.223)	-0.267 (0.192)	-0.233 (0.193)
Female			-0.133 (0.140)	-0.163 (0.141)	-0.253 (0.218)	-0.215 (0.188)	-0.216 (0.189)
AI*Female					0.313 (0.319)	0.165 (0.279)	0.106 (0.282)
Controls				X			X
FE		X	X	X		X	X
N	600	600	600	600	600	600	600

*Notes:* Robust standard errors are reported in parentheses. \*, \*\*, \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Performance 2 is the performance score 1 – 10 on the second test. *AI* is a dummy variable for the AI treatment and *Female* is a gender dummy. Performance-fixed effects are dummies for each performance score 1 – 10 on the first test (with 0 as the reference group). Controls in Specifications 4 and 7 include age, AI Usage, employment, education and income.

Table F.4: Relation Between Text Features and Self-Promotion, Confidence, and Belief 2 in CONTROL

	Self-Promotion				Confidence				Belief 2			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Text 2_Positivity	18.191*** (4.447)			18.852*** (4.641)	1.691*** (0.359)			1.688*** (0.387)	1.641*** (0.376)			1.597*** (0.423)
Text 2_Num_Claims		1.096 (0.833)		-0.948 (1.055)		0.141* (0.083)		-0.013 (0.100)		0.106 (0.084)		0.050 (0.102)
Text 2_Num_Justifications			1.108 (0.825)	1.414 (1.007)			0.092 (0.085)	0.080 (0.097)			-0.003 (0.085)	-0.044 (0.096)
Female	-1.846 (2.043)	-1.211 (2.083)	-1.253 (2.086)	-1.858 (2.048)	0.009 (0.185)	0.074 (0.191)	0.063 (0.191)	0.015 (0.185)	-0.068 (0.195)	0.003 (0.202)	-0.014 (0.204)	-0.176 (0.188)
Controls	X	X	X	X	X	X	X	X	X	X	X	X
FE	X	X	X	X	X	X	X	X	X	X	X	X
N	298	298	298	298	298	298	298	298	298	298	298	298

*Notes:* Robust standard errors are reported in parentheses. \*, \*\*, \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Results are from OLS regressions of the respective dependent variable (*Self-promotion*, *Confidence* or *Beliefs 2*) on the independent variable *Positivity*, *Claims* or *Justifications* as well as dummy for Female. *Self-Promotion* represents each participant's agreement (0-100) with the following statement "I will perform well on the second test in Part 3." *Confidence* represents the number of correct answers participants believe they can achieve on the second test in Part 3. *Belief 2* is the number of questions a participant believes he or she answered correctly in the second test.  $Female_i$  is the binary gender indicator. Performance-fixed effects for *Self-Promotion* and *Confidence* columns are dummies for each possible performance (out of the 10) on the first test, while those for the *Belief 2* columns are based on performance on the second test. Data are from control groups, excluding 2 participants whose *Text 2* are either empty or nonsense. Controls include Age, AI Usage, Employment, Education and Income.

Table F.5: Observed vs Counterfactual Outcomes in the AI treatment

Outcome	Observed	Predicted	Difference	p-value	N
Self-Promotion	80.424	85.896	-5.472	0.000***	290
Confidence	6.959	7.488	-0.529	0.000***	290
Belief 2	6.324	6.839	-0.515	0.000***	290

*Notes:* \*, \*\*, \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. *Self-Promotion* represents each participant's agreement (0100) with the following statement "I will perform well on the second test in Part 3." *Confidence* represents the number of correct answers participants believe they can achieve on the second test in Part 3. *Belief 2* is the number of questions a participant believes he or she answer correctly in the second test. The "Observed" column reports outcome variables from the experimental data. The "Predicted" column reports outcome variables predicted using a model estimated from the control group. Data are from treatment group only, excluding 10 participants whose *Text 2* are either empty or nonsense.

Table F.6: Effects of AI on Changes in Text Features

	Delta_Positivity			Delta_Claims			Delta_Justifications		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
AI	0.136*** (0.023)	0.136*** (0.023)	0.138*** (0.023)	1.213*** (0.112)	1.215*** (0.112)	1.218*** (0.112)	1.101*** (0.128)	1.101*** (0.128)	1.090*** (0.127)
Female		0.033 (0.022)	0.031 (0.023)		0.129 (0.111)	0.104 (0.112)		0.022 (0.127)	0.007 (0.127)
Controls			X			X			X
FE	X	X	X	X	X	X	X	X	X
N	588	588	588	588	588	588	588	588	588

*Notes:* Robust standard errors are reported in parentheses. \*, \*\*, \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Results are from OLS regressions based on Equation 1 with *Delta\_Positivity*, *Delta\_Claims*, and *Delta\_Justifications* as dependent variable, respectively. *Delta\_Positivity* represents the difference of positive sentiment between *Text 1* and *Text 2*. *Delta\_Claims* represents the difference of the number of Claims between *Text 1* and *Text 2*. *Delta\_Justifications* represents the difference of the number of Justifications between *Text 1* and *Text 2*.  $Female_i$  is the binary gender indicator, and  $AI_i$  is the binary AI treatment indicator. Performance-fixed effects are dummies for each possible performance (out of the 10) on the first test. Data are from both treatment and control groups, excluding 12 participants whose *Text 2* are either empty or nonsense. Controls in (3) include Age, AI Usage, Employment, Education and Income.

Table F.7: Probability of Being Hired and Salary

	Prob. of being hired	Salary
Self-promotion	0.009*** (0.001)	0.006*** (0.001)
N	320	320

*Notes:* Robust standard errors are reported in parentheses. \*, \*\*, \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. *Self-Promotion* represents each participant's agreement (0-100) with the following statement "I will perform well on the second test in Part 3." "Prob. of being hired" is a binary variable that takes the value 1 if the employer chose to hire a candidate with the specific *Self-Promotion* score, and 0 otherwise. "Salary" represents the payment determined by the employer in case of hiring. The results are from OLS regressions of the "Prob. of being hired" or the "Salary." Data are from the employer study.

## G Experimental Instructions

We reproduce the instructions from the main study in Appendix G.1 and the employer study in Appendix G.2. Both are translated from the Chinese original instructions.

### G.1 Main study

The main study consists of four parts: first test, self-evaluation, second test and a brief follow-up survey.

The experiment begins with a consent form providing an overview of the experiment, as shown in Figure G.3.

Figure G.3: Consent Form

#### Informed Consent Form

We would first like to express our sincere gratitude for your participation in this study. To ensure the accuracy and voluntariness of the research, there are some details you need to understand and agree to before participating.

**Please read these instructions carefully and complete the study without any assistance.**

Unlike in traditional surveys where questions are hypothetical, every person mentioned here is real and every decision has real consequences.

The study lasts about 22 minutes. It has three parts and a short questionnaire. You will receive a minimum payment of 9 CNY as a fixed completion fee. Apart from the completion fee, you can earn a performance bonus based on one of three parts. That means one of three parts will be selected randomly to determine your bonus, so you should pay equal attention to all parts.

Note that the bonus calculation varies by parts, you can find the details in the instruction for each part.

You must complete the entire study to receive any payments.

The information you provide will be used solely for research purposes. Your data will be processed anonymously and will only be made public anonymously.

By answering and submitting this survey, you understand and agree to the above content, and voluntarily participate. If you do not wish to participate in this study, you may exit the survey at any time.

---

Do you consent to participate this study?

- Yes, I consent.
- No, I do not consent.

After providing their consent to participate in the experiment, participants are shown the instructions for part 1. They are informed that one of the first three parts (i.e., first test, self-evaluation, and second test) will be randomly selected to determine their bonus, emphasizing the need to pay equal attention to these parts. The instructions for part 1 are presented in Figure G.4, and an example question from the first test is shown in Figure G.5.

Figure G.4: Instruction for Part 1

### Instruction for part 1

Please recall that one of three parts will be randomly selected as the part that counts for determining your bonus payment.

**In this part, you complete a test with 10 questions evaluating your math and logic skills.**

You will be asked to solve problems commonly used as a measure of cognitive ability by employers and researchers.

After completing the entire study, you will receive a completion fee plus a bonus. If this part is selected for your bonus calculation, your bonus will be 0.2 CNY **per correct answer** in this part. Unanswered questions count as wrong answers.

You have up to 30 seconds to answer each question. When the time is out, the page will be submitted automatically and move to the next page.

How much time do you have for each question?

- 20 minutes
- 20 seconds
- 5 minutes
- 30 seconds

---

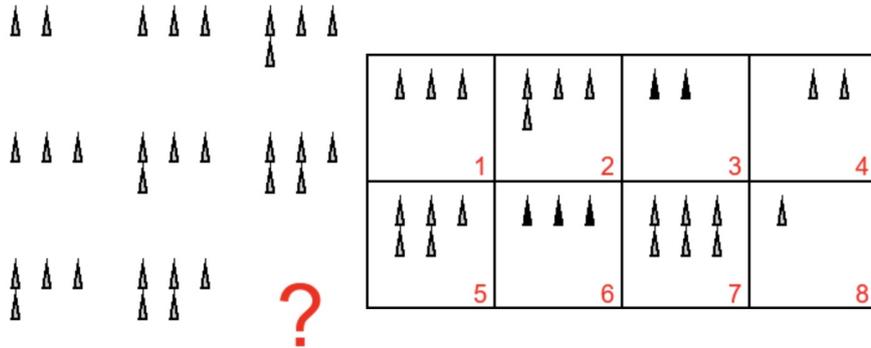
Please answer this understanding question before you start the test.

My bonus for this part ...

- is independent of my answers on the test
- is 0.2 CNY per correct answer on the test

Figure G.5: Example Question of First Test

In the image on the right, there are eight figures numbered 1 to 8. Which one should logically be placed in the spot marked with a red question mark "?" in the image on the left?



- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

After completing the first test in part 1, but before proceeding to part 2, participants are asked to report their belief about their performance on the first test, as shown in Figure G.6.

Figure G.6: Belief Performance on First Test (*Belief 1*)

Congrats! You have now completed the first test.  
Before proceeding to the next parts, please answer the following question.

Out of the 10 questions on the test you took in part1, how many questions do you think you answered correctly?



Next, participants are presented with the instructions for part 2, as shown in Figure G.7. To proceed, they must correctly answer the understanding question to ensure they understand the instructions.

Figure G.7: Instruction for Part 2

**Instruction for part 2**

Please recall that one of the three parts will be randomly selected as the part that counts for determining your bonus payment.

In part 2, you will be asked to answer two questions. First, you will be asked to write a statement related to your performance on the first test in part 1 and how you believe you will perform on an upcoming second test (which will have similar question types and the same number of questions as the first test) in part 3. Then you will be asked to provide a summary statement on a numerical scale.

---

Page Break

---

If part 2 is selected for your bonus calculation, your bonus will be determined as follows.

Your answer to one of two questions will be shown to your "employer", who will be one participant from another study on Credamo. The "employer" can decide whether to hire you and, if so, how much to pay you. For the "employer", it is important how well you will perform on the upcoming second test in part 3.

Specifically, the bonus given to your "employer" and to you will be determined as follows:

- If your "employer" chooses NOT to hire you, your bonus will equal 0.6 CNY and your "employer's" bonus will equal 2 CNY.
- If your "employer" chooses to hire you, your bonus will equal how much the "employer" chooses to pay you. And your "employer's" bonus will equal (i) 2 CNY minus how much the "employer" chooses to pay you, plus (ii) 0.2 CNY times the number of questions you answered correctly on the second test in part 3. The "employer" can choose to pay you any amount in increments of 0.2 CNY, ranging from 0.6 CNY to 2 CNY.

When deciding whether to hire you and, if so, how much to pay you, your "employer" :

- will only know your answer to one of the two questions that you will answer on the next pages in this part.
- will NOT know how many questions you answered correctly on the two tests.

Note that being "hired" does not come with extra work. The hiring decision of the employer only determines how much you and the "employer" are paid.

Understanding question:  
Your bonus for this part will be:

- Equal 0.6 CNY
- Equal 0.2 CNY times the number of questions you answered correctly on the test in part 2.
- Equal 0.6 CNY if you are not hired and the amount your "employer" chooses to pay you if you are not hired.

Step 1: Write an initial self-promotion text. The question is displayed in Figure G.8.

Figure G.8: Initial Self-promotion Text

Based on your experience with the first test that you just took in part 1, please describe how well you think you will perform on the second test and why (**at least 60 words**).

Note that

- the second test in part 3 is similar to the first test in part 1.
- your "employer" will see your final answer to one of the two questions. Based on your answer, the "employer" will decide whether to hire you and, if so, how much to pay you. When making these decisions, your "employer" will NOT know how many questions you answered correctly on the two tests.

Step 2: Refine the text. Participants are provided with the initial self-promotion texts they wrote in the previous step. Instructions for treatment group are displayed in Figure G.9 and Figure G.10. Instruction for control group is displayed in Figure G.11.

Figure G.9: Instruction for Refining Text with AI

Congrats! You have answered the first question.

[Edit HTML](#)

---

Now, you have the opportunity to refine your statement. When you move to the next page, a 5-minute time limit starts, and you can't proceed to the next part until 2 minutes have passed.

Please refine your answer to the first question by chatting with DouBao. Below is the link for DouBao. You can copy the link, then open it in your browser. Do NOT close the current Credamo website. Please use a separate tab or window to open DouBao now.  
Link: <https://www.doubao.com/chat/>

When you are ready move to the next page to see and revise your previous answer to the first question.

Figure G.10: Refining Text with AI

The 5-minute timer starts now.

[Edit HTML](#)

---

Please refine your answer to the first question by chatting with DouBao. Do NOT close the current Credamo website. Please use a separate tab or window to open DouBao.  
Link: <https://www.doubao.com/chat/>

Note that:

- Do not to close the DouBao page until you have completed this study. You will need to copy and paste all prompts you have used when chatting with DouBao on the next page.
- Your "employer" will see your final answer to one of the two questions. **And the answer on this page will be used as your final answer to the first question.** Based on your answer, the "employer" will decide whether to hire you and, if so, how much to pay you. When making these decisions, your "employer" will NOT know how many questions you answered correctly on the two tests.

Below is a suggested prompt for requesting refinement. You may use it directly or create your own.

---

**Suggested Prompt:**  
I am conducting a study on Credamo. As part of the study, I completed a 10-question math and logic test and wrote a text describing how I will perform on a similar 10-question test to follow. A second group of people will then play the role of the employer. They will decide whether to hire me and determine the amount of salary to be paid based on my self-description, which is directly related to the compensation I can get for participating in this study. Below is my self-description. Please help me improve it with an objective attitude so that it can more clearly show my abilities and potential, so that I can get the due rewards in this study.  
**Here is the text I have written before:**

[Edit HTML](#)

---

Please write/copy and paste your refined description in the blank text box below.

Figure G.11: Refining Text without AI

Congrats! You have answered the first question.

[Edit HTML](#)

---

Now, you have the opportunity to refine your statement within a 5-minute time limit, and you can't proceed to next part until 2 minutes have passed.

Tip: You can view your previous statement in the lower section of this page.

Note that your "employer" will see your final answer to one of the two questions. **And the answer on this part will be used as your final answer to the first question.** Based on your answer, the "employer" will decide whether to hire you and, if so, how much to pay you. When making these decisions, your "employer" will NOT know how many questions you answered correctly on the two tests.

Please write your refined description in the blank text box below.

---

Your previous statement:

[Edit HTML](#)

Step 3&4: Provide the self-promotion score and prediction about performance score on second Test, as shown in Figure G.12

Figure G.12: *Self-Promotion and Confidence*

Below is the second question.

Indicate the extent to which you agree with the following statement: "I will perform well on the second test," using a scale from 0 (entirely disagree) to 100 (entirely agree).

Note that your "employer" will see your final answer to one of the two questions. Based on your answer, the "employer" will decide whether to hire you and, if so, how much to pay you. When making these decisions, your "employer" will NOT know how many questions you answered correctly on the two tests.

	Entirely disagree 0	Neither disagree nor agree 50	Entirely agree 100
--	------------------------	-------------------------------------	-----------------------

Click to edit 

---

Page Break

---

Congrats! You have now completed part 1 and part 2.  
Before proceeding to the next part, please answer the following question.

Based on your experience on the first test you just took in part 1, how many questions do you think you will answer correctly in the upcoming second test in part 3 (out of 10)?

Number of Correct Answers 

Next, participants are provided with the instructions for part 3, which involves a test similar to the one in part 1. The instructions are shown in Figure G.13, and an example question from the second test is displayed in Figure G.14.

Figure G.13: Instruction for Part 3

**Instruction for part 3**

Please recall that one of three parts will be randomly selected as the part that counts for determining your bonus payment.

**In part 3, you complete a similar test as in part 1, consisting of 10 questions evaluating your math and logic skills.**

You will be asked to solve problems commonly used as a measure of cognitive ability by employers and researchers.

After completing the entire study, you will receive a completion fee plus a bonus. If this part is selected for your bonus calculation, your bonus will be 0.2 CNY **per correct answer** in this part. Unanswered questions count as wrong answers.

You have up to 30 seconds to answer each question. When the time is out, the page will be submitted automatically and move to the next page.

How much time do you have for each question?

20 minutes  
 20 seconds  
 5 minutes  
 30 seconds

---

Next, begin the test by clicking "NEXT PAGE" button.

Figure G.14: Example Question for Test 2

In the image on the right, there are eight figures numbered 1 to 8. Which one should logically be placed in the spot marked with a red question mark "?" in the image on the left?

The puzzle consists of a 3x3 grid of ovals on the left and a 2x4 grid of numbered options on the right. The 3x3 grid has a red question mark in the bottom-right cell. The 2x4 grid contains eight options numbered 1 to 8, each with a different combination of shading and size.

Options 1-8:

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

After completing the second test, but before moving on to the final survey, participants are asked to report their belief about their performance on the second test, as shown in Figure G.15.

Figure G.15: Belief Performance on Second Test (*Belief 2*)

Congrats! You have now completed the second test.  
Before proceeding to the next part, Please answer the following question.

Out of the 10 questions on the test you took in test 2, how many questions do you think you answered correctly?

Number of Correct Answers  10

Finally, participants are asked to answer demographic questions and AI usage (shown in Figures G.16 - G.18) .

Figure G.16: AI Usage

Congrats! You have now completed all tests.

In the final part of this study we ask you to answer a few questions about your background. Please answer these questions based on your actual information and experience.

---

----- Page Break -----

How familiar are you with ChatGPT, DouBao or other similar AI platforms?

- I have not heard of them before this study.
- I have heard of them but have not used any of them myself before this study.
- I used some of them a few times.
- I use some of them occasionally.
- I use some of them regularly

Figure G.17: Age, Gender, and Education

What is your age group?

- 0 to 20 years old
- 21 to 30 years old
- 31 to 40 years old
- 41 to 50 years old
- 51 to 60 years old
- Over 60 years old

---

----- Page Break -----

What is your gender?

- Male
- Female

---

----- Page Break -----

What is your highest education?

- Primary school and below
- Junior middle school
- General high school / technical secondary school / technical school / vocational school
- Training school
- Bachelor degree
- Master degree
- Doctor degree

Figure G.18: Employment and Income

What is your employment status?

- Full-time work
- Part-time work
- Unemployed (and job seeking)
- Not in paid work (e.g. students, homemaker, retired or disabled)
- Other

---

----- Page Break -----

What is your personal income per year (after tax) ?

- Less than 60,000 CNY
- 60,000 CNY - 119,999 CNY
- 120,000 CNY - 179,999 CNY
- 180,000 CNY - 239,999 CNY
- 240,000 CNY - 299,999 CNY
- More than or equal 300,000 CNY

---

----- Page Break -----

How many studies have you previously completed on Credamo?

- 0 - 10
- 11 - 100
- More than 100

## G.2 Employer Study

The employer experiment begins by a consent form, as shown in Figure G.19.

Figure G.19: Consent Form

By clicking the button below, I give permission to use my data for scientific purposes.

I acknowledge:

that I had enough time to decide whether I want to participate. that participation is voluntary and that I can decide at any time to withdraw from the study. I do not have to provide any justification for such a decision. To receive any payments, I must complete the entire study. that I am informed that this study involves no deception. Any questions have been answered sufficiently. that I am informed that the data will be processed anonymously and will only be made public anonymously.

Do you consent to

- Yes, I consent.
- No, I do not consent

After providing informed consent, participants are presented with the study instructions (see Figure G.20). They are informed that they will make 20 hiring decisions based on results from another study (i.e., the main study). Participants receive a fixed payment of 1 CNY, plus an additional performance-dependent bonus.

Figure G.20: Instruction 1

**Please read these instructions carefully and complete the study without any assistance.**

Unlike in traditional surveys where questions are hypothetical, every person mentioned here is real and every decision has real consequences. We do not deceive participants.

The study lasts about 6 minutes. The study consists of 20 decisions and some comprehension questions. You will receive a minimum payment of 1 CNY as a fixed payment. In addition, you can earn a bonus ranging from 2 CNY up to several CNY.

Which of these 20 decisions actually count for your bonus payment also depends on the decisions of participants from another study. Therefore, if you earn a bonus, the bonus payment may arrive at your account at a later point in time than the fixed payment.

You must complete the entire study to receive any payments.

Understanding question:

How many decisions will you make in this study?

- 1
- 10
- 20
- 25

Participants are then introduced to another study, including descriptions of the two tests, two example questions (see Figure G.21), and the self-promotion scores. They are subsequently informed about how their extra bonus is determined based on their decisions, followed by two comprehension questions (see Figure G.22). Participants can proceed to the next page only after correctly answering both comprehension questions.

Figure G.21: Introduction about the Main Study I

In another study, Credamo participants from China complete a math and logic test. We call these participants job applicants.

Each "job applicants" will finish two similar tests. **Each test has 10 questions evaluating the math and logic skills of the applicants.** The questions are used as a measure of cognitive ability by employers and researchers.

Please look at the two examples on the following screens. Applicants have up to 30 seconds to answer each question (The question page will be automatically submitted when the 30-seconds time limit is up).

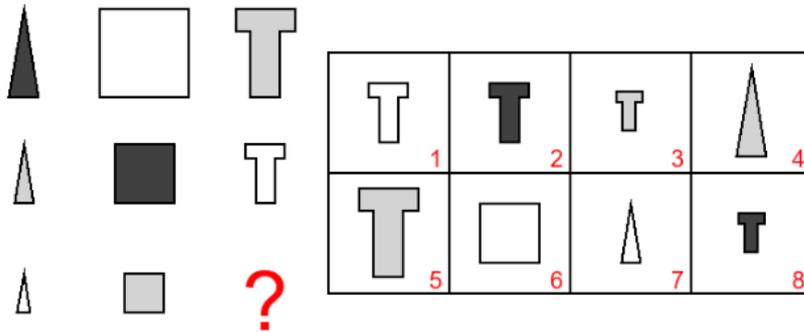
To give you a sense of this, you will be automatically moved to the next page after 30 seconds. You do not have to submit an answer yourself.

----- Page Break -----

右图中有编号1到8的八个图形，哪一个在逻辑上应放入左图中标有红色问号“?”的地方？

In the image on the right, there are eight figures numbered 1 to 8. Which one should logically be placed in the spot marked with a red question mark "?" in the image on the left?

(You can find the correct answer below the picture. Job applicants won't see correct answers when they complete the tests.)



Correct answer: 8

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

## Figure G.22: Introduction about the Main Study II

After completing the first test, the **job applicants answer a question in which they indicate the extent to which they agree, on a scale from 0 ('entirely disagree') to 100 ('entirely agree'), with the statement 'I will perform well on the second test.'** Then the job applicants will complete the second test.

Note that the second test contains the same number of questions and similar types of questions as the first test.

**The applicants know that a person like you will see their answer and use it to decide whether to 'hire' the applicant and how much to pay the applicant if hired.**

You will see 20 possible answers that applicants may provide, **ordered from the lowest to the highest answer.** For each of these 20 answers, you must decide whether to hire the applicant and if so, how much to pay the applicant. Any of these 20 answers can be actual answers of an applicant and hence can count for determining your bonus and the bonus of the applicant. Therefore, you should carefully consider each answer.

It is possible that several applicants paired with you give the same answer. In that case, your decision applies to all of them, and you will receive a bonus for each individual.

---

Your bonus and the bonus of the job applicants will be determined as follows:

- If you do NOT to hire the job applicant, your bonus will equal 2 CNY and the bonus of job applicant will equal 0.6 CNY.
- If you HIRE the job applicant, you choose an amount between 0.6 CNY and 2 CNY to pay the job applicant. This is the bonus of job applicant. Your bonus will equal (i) 0.2 CNY times the number of questions(out of 10) the applicant answered correctly on the second test, plus (ii) 2 CNY minus how much you choose to pay the applicant. That is, your earnings are reduced by the amount that you choose to pay the applicant.

You can choose to pay the applicant any amount in increments of 0.2 CNY, ranging from 0.6 CNY to 2 CNY, i.e., you can choose an amount to pay the applicant from 60cents, 80cents, 100cents, 120cents, 140cents, 160cents, 180cents, or 200cents. Note that 100cents=1CNY.

The applicant receives the same information as you about the hiring decision and the bonus.

Understanding question:

If you hire a job applicant, your bonus is...

- 2 CNY.
- 0.2 CNY times the number of questions(out of 10) the applicant answered correctly on the second test, plus 2 CNY minus how much you choose to pay the applicant.
- 0.2 CNY times the number of questions(out of 10) the applicant answered correctly on the second test.

---

Understanding question:If you do NOT hire the job applicant, your bonus is...

- 2 CNY.
- 0.2 CNY times the number of questions(out of 10) the applicant answered correctly on the second test, plus 2 CNY minus how much you choose to pay the applicant.
- 0.2 CNY times the number of questions(out of 10) the applicant answered correctly on the second test

Then, participants enter the decision stage. They are first shown an overview of the 20 self-promotion scores, followed by 20 individual decision pages, with one hiring decision presented per page (see Figure G.23).

Figure G.23: Decision Page

Below are the 20 possible answers to the scale question, ordered from lowest to highest.  
2, 3, 6, 9, 13, 14, 16, 20, 34, 41, 41, 47, 55, 64, 70, 72, 78, 88, 89, 100.

You need to make 20 hiring decisions based these answers.  
Click "NEXT PAGE" to start.

---

----- Page Break -----

---

Decision 3:  
On a scale from 0 ('entirely disagree') to 100 ('entirely agree'), the job applicant chooses **2** with the statement 'I will perform well on the second test.'

What would you like to do?

Don't hire the job applicant.  
 Hire the applicant (choose below how much to pay the applicant).

---

Show:

Q11 Decision 3: On a scale from 0 ('entir... Hire the applicant (choose below how muc... Selected

---

How much would you like to pay the applicant?  
Note that 100cents=1CNY.

Cents 60 cents 200cents 0

Finally, participants are asked to provide consent to receive a follow-up survey in order to receive the extra bonus (see Figure G.24).

Figure G.24: Consent Form for Bonus

Congrats!

You have now made all decisions. You will receive the completion fee after submitting this study. And we will calculate your bonus within three weeks.

If you consent, you will receive a follow-up survey titled "Bonus for decision making". The follow-up survey will only display the bonus calculation details for the purpose of distributing the extra bonus, without any additional questions or tests.

Do you consent to receive the follow-up survey through Credamo?

Yes, I consent  
 No, I do not consent.

## H Pre-Analysis Plan and Ethics

## AI assistance and self-promotion (#202715)

### Author(s)

This pre-registration is currently anonymous to enable blind peer-review.  
It has 4 authors.

Pre-registered on: 12/04/2024 12:58 AM (PT)

### 1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

### 2) What's the main question being asked or hypothesis being tested in this study?

This experiment investigates the impact that artificial intelligence (AI) has on job applications by studying two questions. First, whether AI assistance boosts self-promotion by providing positive feedback and inspiration for how to write a successful job application. Second, whether AI affects the gender gap in self-promotion.

H1: The self-promotion score is higher for participants in the AI treatment than for participants in the NoAI treatment.

H2: Among participants in the NoAI treatment, the self-promotion score is higher for men than for women. (Replication of the gender gap in self-promotion in Exley & Kessler, 2022, QJE.)

H3: The gender gap in self-promotion is different in the AI treatment than in the NoAI treatment.

### 3) Describe the key dependent variable(s) specifying how they will be measured.

Self-promotion score (SP): Measured by the response to the statement, "I will perform well on the second test," rated on a scale from 0 (entirely disagree) to 100 (entirely agree).

Other outcome variables:

1) Actual performance 1 and 2 (AP1 and AP2): Measured by the number of correct answers that participants achieve on test 1 and 2, respectively.

2) Belief about test performance 1 and 2 (BP1 and BP2): Measured by the response to the question "Out of the 10 questions on the test you took, how many questions do you think you answered correctly?" For test 1 and 2, respectively.

3) Predicted test performance (PP2): Measured by the response to the question "Based on your experience on the first test you just took in part 1, how many questions do you think you will answer correctly in the upcoming second test in part 3 (out of 10)?"

### 4) How many and which conditions will participants be assigned to?

Two treatments: Treatment AI: participants refine their written self-promotion statement by chatting with AI (using a provided prompt). Treatment NoAI: participants refine their written self-promotion statement by themselves.

### 5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

We estimate OLS regression models with SP as the dependent variable and robust standard errors. Our main specifications include performance fixed effects (indicator variables for each possible performance on test 1). Additionally, we will report specifications that add the following control variables: dummies for age range, education level, and usage of AI.

To test H1, we regress SP on a treatment indicator variable AI (that is equal to 1 for the AI treatment and 0 otherwise), using the data from both treatments.

To test H2, we regress SP on a gender indicator variable FEMALE (that is equal to 1 for female participants and 0 otherwise), using the data from the NoAI treatment only.

To test H3, we regress SP on AI, FEMALE, and the interaction AI x FEMALE, using the data from both treatments.

### 6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

We will include data for all participants who complete the study.

### 7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

Target sample size: 600 (300 men and 300 women). Estimated minimal detectable difference in SP: 8.7 (power 0.08, alpha 0.05).

### 8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

In further analysis, we will investigate potential mechanisms. First, we will consider whether confidence (measured by PP2) is affected by the AI treatment. Second, we will consider whether performance on the second test is affected by the AI treatment. Depending on the results, we will do further exploratory analyses. For example, investigating whether women are as likely as men to incorporate feedback from AI in their final self-promotion statement.

Auxiliary study: We conduct a separate study with 16 participants who act as employers. The data will not be used for the main analysis.



Gesellschaft für experimentelle Wirtschaftsforschung e.V.  
German Association for Experimental Economic Research e.V.

## Institutional Review Board Certificate

No. fjC3ufY5

<https://gfew.de/ethik/fjC3ufY5>

Date: 12/04/2024

Certificate Expires: 12/04/2026

Status: Approved

Review Level: Expedite

### Details of Experiment:

#### Titel

*AI assistance and self-promotion*

#### Authors

*Koch, Alexander; Kragl, Jenny; Nafziger, Julia; Ming, Sijuan*

#### Short description of the experimental study

*This experiment investigates the impact that artificial intelligence (AI) has on job applications by studying two questions. First, whether AI assistance boosts self-promotion by providing positive feedback and inspiration for how to write a successful job application. Second, whether AI affects the gender gap in self-promotion.*

*The experiment is based on the self-promotion paradigm of Exley & Kessler (2022, QJE). It consists of a main study and an auxiliary study.*

*In the main study, participants first complete a math and logic test. Then they are informed that they will perform a second test (that has the same number of questions and a similar level of difficulty as the first test), and that their performance on this test will be relevant for an employer. Participants are informed that the employer matched with them will decide whether to hire the participant and how much to pay the participant based on the answer that the participant gives to one of the following questions. The first question asks participants to write a statement related to their performance on the first test and how they believe they will perform on the upcoming second test. The second question asks participants to indicate their agreement with the statement "I will perform well on the second test" on a scale from 0 (entirely disagree) to 100 (entirely agree).*

*There are two treatments. In the AI treatment, participants refine their written self-promotion statement by chatting with AI (using a provided prompt). In the NoAI treatment, participants refine their written self-promotion statement by themselves. After the revision stage, participants answer the second question that provides a numerical self-promotion score. This score is the main outcome variable in this study.*

*In the final stage of the experiment, participants complete the second math and logic test.*

*In the auxiliary study, participants are assigned the role of employers. They receive information about the math and logic tests in the main study and how the test performance of a participant in the main study affects an employer's pay. Thereafter, employers are presented with several possible*

---

Gesellschaft für experimentelle Wirtschaftsforschung e.V.

German Association for Experimental Economic Research e.V.

University of Passau | School of Business, Economics and Information Systems

Innstrasse 27 | 94032 Passau | Germany | [info@gfew.de](mailto:info@gfew.de) | [www.gfew.de](http://www.gfew.de)



Gesellschaft für experimentelle Wirtschaftsforschung e.V.  
German Association for Experimental Economic Research e.V.

## Institutional Review Board Certificate

No. fjC3ufY5

<https://gfew.de/ethik/fjC3ufY5>

*self-promotion scores that participants in the main study may generate. For each score, employers decide whether to hire a participant with that score and how much to pay the participant.*

*Each participant in the main study is matched with an employer in the auxiliary study based on the self-promotion score of the participant. The participant in the main study is paid based on the employer's hiring and payment decision. In case the participant from the main study is hired, the employer's pay additionally depends on the performance of the participant on the second math and logic test.*

*The experiment will be conducted online in Chinese, using the research platform Credamo. The target sample size is 600 participants (300 men and 300 women).*

The German Association for Experimental Economic Research (Gesellschaft für experimentelle Wirtschaftsforschung - GfeW) provides this certificate of approval in an expedited review process according to the rules and procedures approved by the members of the association on September 25, 2024. The criteria for approving an ethics certificate can be retrieved here: <https://gfew.de/criteria>

The approval is based on the agreement of the proposers to abide by the legal requirements of the European Union and by the policies and procedures of the GfeW. Furthermore, the approval is only valid, if (i) the proposers have provided all necessary information completely and truthfully, (ii) no part of the planned experimental project has been conducted before the approval was granted, and (iii) the experimental project is conducted as described.

This approval does not supercede or replace decisions by other statutory bodies. Furthermore, it is not a substitute for certifications or approvals that are required by law, by the statutes of the involved research institutes, or by the codes of conduct that are applicable to the participating researchers, including those from other fields of research.

The expiration date of the approval is noted above. To continue the proposed experimental study beyond that date, a new approval must be requested, even if the experimental protocol remains unchanged.

Please also note that prior to implementing any changes or amendments in the protocol of the proposed experimental study, a new approval must be requested, even if the new experimental study including the changes or amendments to the protocol is expected to be governed by the GfeW's rules for the expedited review process.

A copy of this letter should be kept in the records together with all documents associated with this experimental study.

---

Gesellschaft für experimentelle Wirtschaftsforschung e.V.

German Association for Experimental Economic Research e.V.

University of Passau | School of Business, Economics and Information Systems

Innstrasse 27 | 94032 Passau | Germany | [info@gfew.de](mailto:info@gfew.de) | [www.gfew.de](http://www.gfew.de)



Gesellschaft für experimentelle Wirtschaftsforschung e.V.  
German Association for Experimental Economic Research e.V.

## Institutional Review Board Certificate

No. fjC3ufY5

<https://gfew.de/ethik/fjC3ufY5>

Prof. Dr. Johann Graf Lambsdorff  
President of the GfeW

Prof. Dr. Björn Frank  
Vice-President of the GfeW