Rank versus Inequality—Does Gender Composition Matter?*

Duk Gyoo Kim† Max Riegel‡

September 13, 2024

Abstract

This study investigates the influence of gender composition on allocation decisions involving a rank–inequality tradeoff. In a laboratory experiment, participants chose to either alleviate inequality by relinquishing their current relative rank or exacerbate inequality while maintaining their current rank. Two essential features of the experiment are: 1) participants' relative rank is the outcome of their real-effort performance and luck; 2) participants' genders are naturally revealed by gender-specific nicknames. We found that female participants are more reluctant to relinquish their current relative rank when the persons ranked below and above them are of the opposite gender. This tendency was less pronounced in the male participants.

JEL Classification: C91, D63

Keywords: Gender composition, Positional concerns, Preferences for Redistribution, Last-Place Aversion, Perception of luck

*We thank Puneet Arora, Syngjoo Choi, Lawrence Choo, Hans Peter Grüner, SeEun Jung, Andrea Martinangeli, Masao Ogaki, Hannes Rau, Hyelim Son, participants at Sungkyunkwan University, ZEW and University of Mannheim Experimental Seminar, KBER Summer Workshop in Behavioral and Experimental Economics, HeiKaMaxY Workshop, Virtual East Asia Experimental and Behavioral Economics Seminar, and European Winter Meeting of the Econometric Society for their comments.

†Department of Economics, Sungkyunkwan University. Email: kim.dukgyoo@skku.edu

‡Department of Economics, University of Mannheim. Email: mariegel@mail.uni-mannheim.de

1 Introduction

Imagine a situation where you have to choose to either: 1) relinquish your current relative rank or position to reduce economic inequalities around you, or 2) maintain your current rank, which exacerbates inequalities; your absolute monetary payoffs are unaffected in both cases. Although drastically simplified, this situation captures the tradeoff relationship between concerns for relative rank and social preferences for distributional outcomes. For instance, such a tradeoff could affect the political support for redistribution schemes: Supporting more redistribution than the status quo may lead some supporters to relinquish their current relative rank while reducing economic inequalities. Kuziemko et al. (2014) provide survey evidence that some workers whose current hourly wage is slightly above the minimum wage stand against the idea of increasing the minimum wage. Increasing the minimum wage can reduce economic inequalities without making any harms to them, but it involves the workers relinquishing their current relative rank from ‘above the minimum wage’ to ‘at the minimum wage’. Our research questions regarding this hypothetical situation are as follows: How is your decision affected if you know the gender of the persons affected? Would your decisions differ if you obtain a relatively lower/higher rank? To answer our research questions, we examined the effect of gender composition on participants’ allocation decisions in a plausible environment where both individual performance and luck jointly determine payoff-relevant rank.[1]

Our question stems from two actively studied research agendas: utility from a relative rank and preferences for redistribution. Although muted in most economic theories, concerns for relative rank drive our daily decisions and affect our well-being to a non-negligible extent.[2] In the two-dimensional space of incomes and favors toward progressive tax–redistribution policies, the standard economic models à la Meltzer and Richard (1981) predict that the preferences for redistribution should be negatively aligned with income levels. However, we sometimes observe "off-diagonal" preferences for redistribution: Some low-income citizens vote against redistribution, while some high-income citizens vote for it.

[1]Another straightforward question is whether men and women act differently when trading rank versus inequality. Kuziemko et al. (2014) as well as Martinangeli and Windsteiger (2021) however do not report statistically significant differences in that regard.

[2]The concerns for relative standing were recognized by early philosophers such as Veblen (1899), Keynes (1930), and Russell (1930). Following Duesenberry (1949), some theoretical studies focus on the resource allocation to positional goods (Frank, 1985; Hirsch, 1995; Hopkins and Kornienko, 2004) and discuss possible policy interventions (Frank, 2008) and optimal taxation (Aronsson and Johansson-Stenman, 2010, 2018). For a broader review of the literature on relative rank and happiness, see Clark et al. (2008).

Among several plausible ways of attempting to answer "why the poor vote for the rich,"[3] we are particularly interested in the role of utility from a relative rank.

The broad question regarding the relationship between concerns for relative position and preferences for redistribution has been examined in various ways. People may be averse to redistribution even when it is costless to themselves because they want to maintain their relative standing (Kuziemko et al., 2014), because they prefer to avoid decisions that reverse the existing rankings (Xie et al., 2017), or because they take into account the potential loss from other dimensions associated with changes in the redistribution policy (Kim, 2019; Gallice and Grillo, 2019). However, to the best of our knowledge, little is known about the role of gender composition of persons adjacent in rank to the decision maker. Gender composition could affect the decision maker's well-being via preferences for rank and redistribution. On the one hand, individuals could be habituated to care more about their intra-gender rank, as in many real-life reference groups consisting of either men or women only, such as sports teams, circles of friends, and professional networks in sectors dominated by one gender. In addition, partnering or marriage decisions involve only intra-gender comparisons. Given this context, people may care more about their relative position within the same-gender group. On the other hand, evidence from dictator games (Ben-Ner et al., 2004; Sharma, 2015) suggests that preferences for redistribution might depend on the gender of the decision maker's counterparts. We contribute to the literature by experimentally investigating the effect of gender composition on allocation decisions when facing a rank–inequality trade-off.

We faced two empirical challenges, which was the reasoning behind using a laboratory[4] experiment to answer our research questions. One challenge was to make gender information salient without imposing an emphasis on it. The second challenge was to allow for within-participant variation in rank over rounds while ensuring that ranks should not be

[3] Low-income citizens with the prospect of upward mobility may believe that redistribution will prevent them from prospering in the future (Benabou, 2000). Biased perceptions in individuals' evaluations of their own relative positions (Cruces et al., 2013) or heterogeneous beliefs about intergenerational mobility (Alesina et al., 2018) and about immigrants (Alesina et al., 2022) could affect preferences for redistribution. The public may be sensitive to the media coverage of household-income inequality, rather than the inequality itself (Kelly and Enns, 2010). Doubts about the political system and the government's efficiency may prevent the poor from supporting liberal parties or policies (Houtman et al., 2008). Economic inequality could have an informational value for social decisions (Corneo and Grüner, 2000). The social and cultural conservatism applied to a specific region (Frank, 2004) or religion (Guiso et al., 2003) may drive opposition to the political party which advocates for economic redistribution. Meanwhile, some high-income citizens' support for redistribution can be explained when values are viewed as luxury goods (Enke et al., 2023).

[4] To be more specific, we conducted a real-time online experiment with subjects recruited from the laboratory pool.

entirely random. In our experimental setting, the gender information of other participants was revealed innocuously by displaying the (gender-identifiable) nicknames self-chosen for anonymity, and the luck factor and the participant's performance co-determined the relative rank of the participant in approximately equal weights. Considering that both luck and effort affect an individual's relative standing in diverse real-life circumstances, the inclusion of the luck factor in the experimental setup adds a significant dimension. We believe that both experimental features contribute to the methods in experimental economics, as the revelation of innocuous gender information can be used in many other experimental studies,[5] and the adjustment of the random scaling parameter can allow us to examine another underexplored aspect regarding the preferences for redistribution—the (perceived) attribution of luck and effort for economic achievement.[6]

The experiment consisted of two main parts. Prior to the first part, the participants were asked to provide basic demographic information, including gender. In the first part, participants individually worked on a real-effort task, knowing that better performance would more likely lead to a higher payoff. The second part consisted of five decision rounds. In each decision round, the participants were randomly divided into groups of six and informed about their relative rank within the group. Their rank was based on their performance in the first part of the experiment, scaled by a luck factor (dice roll of a computer). The participants decide to allocate additional money to either a participant one rank above or one rank below while knowing the gender (via the gender-identifiable nicknames) of all group members. After each decision round, the participants were randomly reshuffled to form new groups, chose new nicknames, and were informed about new ranks. After the five rounds of the second part, the participants filled out the post-experiment survey, which included their perceptions about the role of luck and their performance.

We found the female participants more reluctant to relinquish their current relative rank (and alleviate the inequality by doing so) when the rank-below and rank-above persons were male. A similar pattern was observed for male participants, but it was not as strong as that

[5] Making gender salient without an unnecessary emphasis is easier said than done, as there is a tradeoff relationship. Sometimes concerns for demand effects make the gender salience weak: For example, Baranski et al. (2023) use colors (pink for female participants and blue for male participants) as cues for gender information in a three-person bargaining experiment where no explanation was given about what the colors meant. Croson et al. (2008) composed the single-gender sessions by recruiting subjects from fraternities and sororities, while the both-gender sessions were recruited from a service fraternity and a business fraternity. The opposite case also exists: For example, Solnick (2001) uses real first names (and tries to exclude participants with unrevealing first names) to examine gender differences in the ultimatum game.

[6] By adjusting the support of the luck scaling factor, we could build different institutions varying from a complete meritocracy (when luck plays no role) to a complete ex-ante egalitarian (when luck dominates effort).

of females. We did not find support for last-place aversion (LPA, Kuziemko et al., 2014), the reluctance to allocate downward when ranked second from the bottom. Combined with a well-established finding that low-income females are more supportive to redistributive policies (Mengel and Weidenholzer, 2023), our findings suggest that the gender composition of people at adjacent ranks and the perceptions thereof matter for the support of redistribution.

The remainder of this paper is organized as follows. The following subsection provides an overview of the relevant literature. Section 2 outlines the experimental design, hypotheses, and procedures. Section 3 reports the results of the experiment, and Section 4 concludes.

1.1 Literature Review

Our experimental design adopts the encryption task used in Erkal et al. (2011) and the money allocation task used in Kuziemko et al. (2014, henceforth KBRN). The encryption task involves repeatedly decoding a random sequence of letters into corresponding numbers using a decoding table during a prespecified period. Our experimental setup was similar to theirs in the sense that an individual's earnings were determined by the relative rank of the task performance, rather than by the absolute level of performance. The primary difference was that we considered a luck factor adjustment so that each participant's rank could vary substantially by decision round, while their relative level of the absolute performance would vary less.

Although our research is not focused on reexamining KBRN's LPA, it is worth discussing as our experimental design allows us to examine it in the second part. Motivated by the fact that some workers whose hourly wage is just above the minimum wage rate do not support the idea of raising the minimum wage, KBRN conducted an experiment where each participant in a group of six was randomly assigned a payoff-relevant rank and asked to decide whether to allocate additional money to either a person ranked just above or a person ranked just below. KBRN found that participants with the second rank from the bottom are distinctively reluctant to allocate money to the person in the last place, calling this effect LPA. A replication of their experiment (Camerer et al., 2016) reported the opposite effect, questioning the validity of the observations. Martinangeli and Windsteiger (2021) replicated and extended KBRN, claiming that the way of assigning rank could matter. They found that when ranks were randomly updated in each decision round, as in KBRN, LPA is rejected. However, when considering a treatment where the randomly assigned ranks are maintained over the entire decision rounds, they find support for LPA. This suggests that randomly updated ranks may not be perceived as participants' actual ranks. We appreciate the way of assigning the same rank for the entire experiment, as it renders the participants some sense that the rank is endowed, but it does not allow for the observations of how the participants' decisions vary with rank. Our experimental design embraces the advantages of the design considered in KBRN and Martinangeli and Windsteiger (2021): By taking the performance in the real-effort task as a base, each participant's rank is determined by a rank of the adjusted performance that is scaled by a luck factor. In this way, we can provide the sense that the participant's "average" rank is endowed, yet the participant's rank varies by decision round.

Laboratory experiments on gender differences form another important strand of the related literature.[7] We focus on reviewing a subset of studies that help us form conjectures on the effect of gender composition on redistribution. Overall, the results are mixed: Some studies suggest that women tend to allocate money to men at a lower rank, while others suggest the opposite. Ben-Ner et al. (2004) reported that in a dictator game context, women give less to other women when the gender of participants is known. Dalmia and Filiz-Ozbay (2021) reported that behindness aversion is distinct among women, and Dasgupta et al. (2019) also found evidence of stronger behindness aversion among women. Regarding our experimental setting, these findings suggest that a female participant might allocate less to a woman of a lower rank. On the other hand, women give more to others in a dictator game (Eckel and Grossman, 2001; Bilén et al., 2021) and in a variant of the dictator game called the solidarity game (Selten and Ockenfels, 1998). In addition, women tend to favor redistributive policies more than men do (Alesina and Giuliano, 2011; Funk and Gathmann, 2015). These observations may indicate that women are more inequality-averse than men; hence, a female participant would allocate more to a participant at a lower rank, regardless of gender.

More mixed yet, our experimental setting involves noisy feedback in performance, so gender differences in the perception of the feedback may affect the allocation decision. Shastry et al. (2020) examined gender differences in tournament entry upon receiving noisy feedback on relative standing. They found that women tend to attribute negative feedback to a lack of ability, even when it was actually due to bad luck. In our experiment, if the female participants holding a low rank feel that they "deserve" their rank (and the rank-related payoffs), it may suggest that they could be more lenient in allocating downward. Mean-

[7] For an extensive review on gender differences in economics experiments, consult Croson and Gneezy (2009) and Niederle (2017).

while, women perform competitive tasks better in a single-gender environment (Gneezy et al., 2003) and exert more effort in contests when competing against other women (Mago and Razzolini, 2019). These results are not directly applicable because, in our experimental setting, the participants individually performed the real-effort task without knowing how exactly the individual performance would be translated into monetary payoffs. However, if women's competitive behaviors when surrounded by other women are driven by concerns for intra-gender relative rank, a female participant in our experiment may be reluctant to allocate to a woman at a lower rank.

2 Experimental Design, Hypotheses, and Procedure

2.1 Design

We adopted a within-subject design, where the primary source of variation was the gender composition of the adjacent participants. The aim of the experiment was to observe: 1) how gender composition affects the allocation decision; 2) how such effects vary in the participant's current rank; and 3) how the perception of luck affects the allocation decision.[8]

Prior to the main experiment, the participants were asked to provide basic information, such as gender, age, field of study, and previous experience of online experiments, without leaving any questions unanswered. Since our primary manipulation was to make the group members' genders salient, while avoiding the experimenter demand effect (Zizzo, 2010) for the participants to become aware that the revelation of the gender is crucial, we did not emphasize any part of the questionnaire.[9]

The experiment consisted of two parts. In Part 1, the participants worked on an encryption task (Erkal et al., 2011) for five minutes. Per trial, each participant was given a sequence of five random letters and a decoding table showing how letters map into numbers. Their task was to decode the sequence of letters into corresponding numbers in the table. For example, if the text sequence YDDYO is given, and a decoding table states that D, O, and Y correspond to 3, 6, and 8, respectively, then the correct answer is 83386. A new decoding table was provided for each text sequence. The participants were told that the more correct answers they gave, the more money they were likely to earn. We also informed the participants that final earnings would be based on their decisions, others' decisions, and

[8]The complete experimental instructions can be found in Appendix A.

[9]In addition, since the first part of the experiment—the individual real-effort task—had nothing to do with gender, we believe that the experimenter demand effect, if any, is negligible.

some luck, but we only provided further details of how their earnings would be determined in Part 2. Previous studies reported no significant gender differences in encryption task performance (Erkal et al., 2011).

Part 2 of the experiment comprised five decision-making rounds of the money allocation task à la Kuziemko et al. (2014). Figure 1 illustrates how Part 2 proceeded.

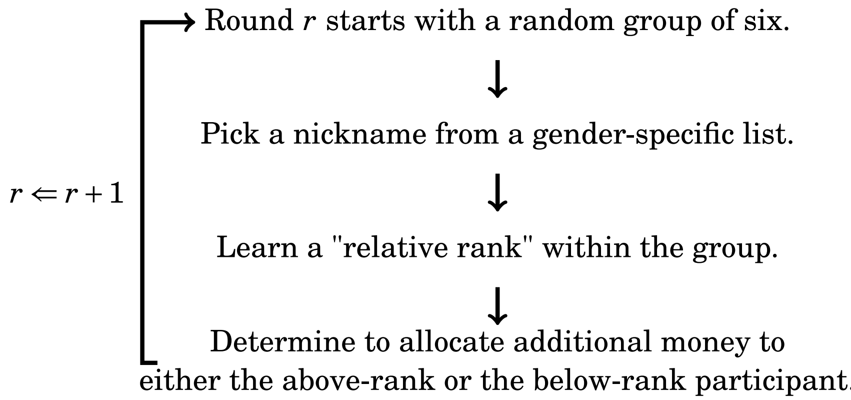


Figure 1: Flowchart of Part 2

At the beginning of each round, the participants were randomly assigned to a group of six, and each picked their nickname from a provided gender-specific list of alternatives. We explained that picking a nickname is a way to make every participant remain anonymous, yet clearly distinguish other participants in the group by their nicknames. After they selected their nicknames, a server computer calculated and displayed their *relative rank* within the group. The rank was determined by their performance in Part 1, scaled by the outcomes of the computerized dice rolls. Specifically, the rank of participant i, R_i, is determined by $\#\{j \in G_i | P_j \times [D_j + 2] \geq P_i \times [D_i + 2]\}$, where G_i is the group for participant i, P_j is the performance measure (i.e., the number of correct answers in Part 1) of participant j, and D_j is the random outcome of a computerized dice roll for participant j. In case of a tie, the participant whose participant number (a unique number internally assigned to each participant) was smaller received a better rank. We instructed the participants that the computer randomly rolled a dice for each participant, calculated the "adjusted performance," which is the number of correct answers in Part 1 multiplied by [dice roll+2], and showed the relative rank of the adjusted performances within the group. For example, if participant 1 gave 10 correct answers in Part 1, her adjusted performance could vary from 30 ($= 10 \times [1+2]$) to 80 ($= 10 \times [6+2]$). The participants received different amounts of money according to the ranking: The person obtaining rank r, $r \in \{1, 2, \ldots, 6\}$, received $17 - 2r$ euros. As the participants were only informed about their relative rank, they could infer neither the realization of the dice roll nor the other group members' actual performance measures.

After every participant observed the rank of all group members, hence the initial money allocations, they were asked to allocate two additional euros to one of the two adjacent participants in ranking.[10] For example, a third-ranked participant was required to allocate two additional euros to either a person ranked second or fourth. For notational simplicity, we refer to the person who is at one rank above the participant as Rank-Above and the person who is at one rank below the participant as Rank-Below. The participant ranked first (last) had to allocate money to the Rank-Below (Rank-Above). Since the base payoff of each rank increases by two euros, allocating the additional two euros to the Rank-Below implies that the participant chooses to sacrifice her relative rank. Participants were informed that the allocation decision of only one randomly selected participant would be implemented. Final payoffs were not displayed during Part 2, such that other participants' decisions remained unknown.

For every new round, the participants repeated the procedures mentioned above. All participants were reshuffled and assigned to a new group of six. The computer rolled new dice for the new round, independent of previous outcomes. The new ranking was then determined based on their performance in Part 1, their new group members' performance, and new dice rolls. Once all five rounds ended, the computer randomly selected one round and one group member's allocation decision for the actual payment. Until the end of the experiment, the participants did not receive any feedback on the selected outcome.

There are several points worth discussing. We made the participants' gender salient by asking them to select one fictitious name from a list of popular names.[11] Thus, it was crucial to make the participants unambiguously grasp the gender of other participants. For this purpose, we carefully selected a list of first names. First, we collected 100 boy names and 100 girl names that were popular for babies born in the 1990s and 2000s, excluding known unisex names. Most participants were born in the 1990s and 2000s, so the name list contains contemporary names from their perspectives.[12] We then asked 80 Amazon Mechanical

[10]Participants were not allowed to decline the allocation of the additional two euros. Our experimental design aimed to capture a rank-inequality tradeoff, and with the inclusion of the option for no allocation, efficiency concerns also could become relevant for participants.

[11]Chakraborty and Serra (2023) also asked experiment participants to choose a nickname from a list of gender-specific names. However, their primary purpose was to prevent confounding bias of race, nationality, or ethnicity associated with the participant's actual name, so they used distinctively white-sound names taken from Bertrand and Mullainathan (2004), and the chosen nickname was used throughout the entire experiment.

[12]An undesirable side effect is that participants might attribute good or bad personal impressions to certain first names. However, we claim that such name effects are orthogonal to rank and gender composition and

Turk (mTurk) participants to classify the 200 names as male, female, or unsure/unisex and used 60 names for each gender on which the survey participants formed a near-consensus as boy and girl names.[13] We filtered survey respondents to obtain 29 German-speaking, 37 European, and 14 Asian respondents. This procedure aimed to match the proportion of survey respondents with that of participants in the other experiments that used the same participant pool.[14] Also, to avoid potential confusion, every experiment participant received different name options per round; therefore, the participants were not concerned about possible duplication within a round.

We carefully calibrated the range of the random scaling factor from 3 to 8 (dice roll+2), which determined the adjusted performance. If the random scaling factor varied too much compared to the performance variance among participants, then the actual performance would not matter for the relative rank. In contrast, if the random scaling factor varied too little, the relative rank across rounds would also vary too little. Ideally, we wanted to ensure that an increase in the actual performance rank would on average raise the adjusted performance rank by 0.5.[15]

The main experiment was followed by a post-experiment survey, which included the participant's guess of the actual rank in performance, perceptions about the luck factor, and a general view about the inequality caused by luck.[16]

2.2 Hypotheses

We had three testable null hypotheses regarding how (perception of) ranks and gender composition affect allocation decisions. When previous findings from related studies are mixed, we hypothesize that there are no effects.

---

thus do not affect our primary observations in aggregate.

[13]The following boy names are used: Aaron, Adam, Albert, Anthony, Arthur, Christopher, David, Derek, Edward, Ethan, Frank, Fred, Hugo, James, Jason, Patrick, Peter, Richard, Robert, Ronald, Steven, Thomas, Walter, Xavier, Benjamin, Brandon, Bruno, Charles, Christian, Daniel, Diego, Eric, Gregor, Isaac, Ivan, Jacob, John, Jonathan, Kevin, Leonard, Manuel, Marco, Matthew, Nicholas, Norman, Paul, Phillip, Ralph, Samuel, Scott, Stanley, Victor, Vincent, William, Alexander, Andrew, Brian, George, Henry, and Oscar. The following girl names are used: Alyssa, Elena, Helena, Isabella, Jasmine, Jennifer, Lisa, Luisa, Monica, Sarah, Sylvia, Alicia, Angelica, Christina, Jessica, Julia, Katherina, Lara, Laura, Marisa, Olivia, Rebecca, Regina, Samantha, Sandra, Selina, Teresa, Claudia, Cynthia, Elizabeth, Ella, Eva, Lena, Linda, Lydia, Melissa, Mia, Michaella, Natasha, Rose, Sabrina, Sophia, Ariana, Bianca, Cassandra, Clara, Daniela, Diana, Emily, Emma, Erica, Lily, Maya, Vanessa, Alina, Amanda, Gloria, Maria, Stella, and Victoria.

[14]See Online Appendix Section A for mTurk instructions and other details.

[15]See Online Appendix Section B for more details.

[16]Before answering the survey, participants learned their adjusted performance rank in five different groups. This noisy feedback on performance is thus incorporated in their perceptions.

Hypothesis 1. *There is no last-place aversion. Fifth-ranked participants are equally as likely to allocate downward as participants holding ranks 2, 3, and 4.*

Hypothesis 1 was based on the previous findings on LPA. Using random assignment of ranks in every round, KBRN provided supporting evidence for LPA, rejecting this hypothesis, while Camerer et al. (2016) and Martinangeli and Windsteiger (2021) replicated the same experimental setup and provided opposing evidence for LPA. Our experimental settings were not directly comparable: The existence of the performance factor in the determination of ranks is very consequential on individual behavior. It potentially affects the system of attitudes towards ranks and preferences for alterations of the distribution. However, we believe that the environment under consideration that provides noisy feedback (rank) based on performance and luck is more aligned to the relevant situation for LPA than a mere random assignment of ranks.

Hypothesis 2. *Information about the genders of the Rank-Above and Rank-Below does not affect an individual's tendency for redistribution.*

Hypothesis 2 states our null hypothesis regarding the effect of gender composition. As described in the literature review section, our conjectures regarding the effect of gender information on allocation decisions are mixed.

Hypothesis 3. *Participants who believe their performance-based rank was higher than the announced rank tend to distribute downward less often.*

Hypothesis 3 states how participants' sense of "deservedness" affects their allocation decisions. When a participant feels that they are "under-placed," that is, their performance-based rank is higher than the announced rank, they believe that the luck factor negatively affects their payoff. This sense, in turn, might embolden the participant's concern for their relative rank and drive them to allocate upward to keep their current (still not satisfying) relative rank.

2.3 Procedures

The experimental sessions were conducted in November 2021 and January 2022 and included participants recruited from the Mannheim Laboratory for Experimental Economics (mLab) of the University of Mannheim. Invitations were sent to members of the participant pool who had already provided gender information in the recruitment system.[17] In-

[17] Using that information, we invited only subjects who identify as either male or female.

structions were provided in English. Owing to COVID-19 restrictions, we did not bring participants to the laboratory. Instead, we invited them to join an online meeting to receive instructions from the experimenter, distributed a link for participating in the online experiment, and paid them via either PayPal or bank transfer. Six sessions with 18 participants (nine females and nine males) were conducted, hence a total of 108 participants participated in the experiment. We used an interactive online platform called LIONESS (Live Interactive Online Experimental Server Software, Arechar et al., 2018). After the participants joined an online meeting, the experimenter disabled their profile photos and asked them to turn off their webcam. Furthermore, the participants were asked to edit their displayed names to two letters they arbitrarily chose so that their identities, and hence decisions, remained anonymous to the experimenter as well as the other participants. The participants were asked to carefully read the instructions for each part of the experiment and pass a comprehension quiz.

The average payment per participant was 10.33 euros. Payments were made via online transfers after receiving the personal payment code that was generated at the end of the experiment. Each session lasted less than 40 minutes.

3 Results

In Part 1, on average, each participant correctly decoded 24.79 text sequences (SD 3.93). As the minimum and maximum performances were 15 and 36, respectively, there is a nonzero probability that the highest-performing participant could be ranked lower than the lowest-performing participant. The correlation between the ranks based on adjusted performance and those based on actual performance was 0.4065 ($p < 0.001$), which means that the luck factors perturbed the relative rank of the group members to some extent, while the rank of actual performance significantly affected the relative rank.

As intended, there were no differences based on gender in the participants' performance. The average number of correctly decoded text sequences was 24.76 (SD 3.78) for females and 24.81 (SD 4.11) for males, and the difference was not statistically significant (Kolmogorov-Smirnov (KS) test, $p = 0.60$; Mann-Whitney (MW) test, $p = 0.90$). The distribution of ranks based only on actual performance was not statistically different by gender (KS, $p = 0.90$; MW, $p = 0.81$) either. Thus, we claim that our findings are not driven by gender differences in the ability to perform in Part 1.

The randomly determined groups of six in Part 2 were fairly balanced in terms of the overall gender composition: 37.78% of groups consisted of exactly three males and females. 90% of groups were composed of at least two and at most four females.[18]

3.1 The effects of rank and gender composition

Figure 2 shows the proportion of decisions allocating additional money to the Rank-Below, collated by rank. Trivial decisions from participants holding rank 1 and rank 6 were excluded, as they had only one option to allocate money. For all ranks, the majority of participants (80.83%) preferred to sacrifice their relative rank by allocating money to the Rank-Below.[19] Thus, participants tended to prioritize inequality over rank concerns. This tendency was slightly more pronounced for females, as 83.05% of female participants allocated downward, and 78.69% of males did so. However, the difference was not statistically significant ($p = 0.51$). Another notable finding was that we did not observe LPA. Although statistically insignificant, we observed the opposite pattern: The participants at rank 5 allocated additional money to the Rank-Below more often than participants at any other rank. These results support our Hypothesis 1.

Result 1. *Fifth-ranked participants allocated downward as much as other participants.*

Unlike rank, we find significant effects of the gender composition of the Rank-Above and Rank-Below on allocation decisions. From the perspective of the member holding rank $r \in \{2,3,4,5\}$, when both members at rank $r-1$ and $r+1$ are of the same gender as member r, we call it the *same-gender* situation (SG, 75 observations). When both Rank-Above and Rank-Below members are of the opposite gender from member r, we call it the *mixed-gender* situation (MG, 88 observations). When the Rank-Above is of the same gender as member r, and the Rank-Below is of the different gender from member r, we call it the *same-above* situation (SA, 99 observations). The remaining situation is analogously called the *same-below* situation (SB, 98 observations).[20]

[18] Simulations (random permutation of 9 males and 9 females to form three groups of six for 1,000,000 iterations) show that the empirical CDF for the number of females per group is close to the true CDF under random allocation. In particular, the probability of a group with three females is 37.97%, and that with two to four females is 86.88%.

[19] We do not find any order effects: The share of subjects allocating downward in the first round (84.72%) is almost the same as in the last round (83.33%), and the estimated coefficient of the round number is statistically insignificant for all regressions we performed.

[20] Since the participants are asked to consider two adjacent participants, we believe that it is appropriate not to differentiate the situations further with respect to the gender composition of participants at remote ranks. Although the contexts are somewhat different, Fisman et al. (2020) reported that experimental participants are locally competitive, preferring to decrease the income level directly above them, but not two positions

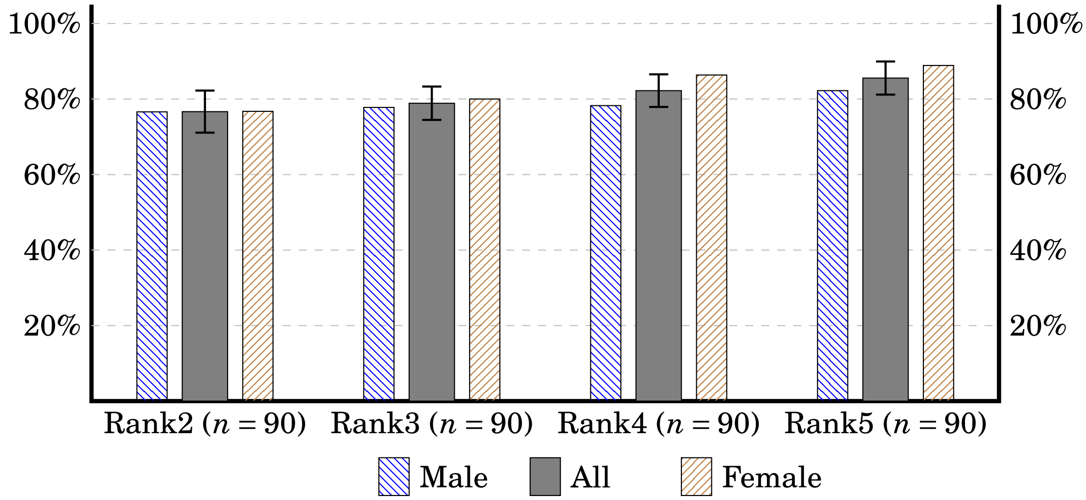


Figure 2: Proportions of allocating downward, by rank
Error bars represent standard errors from an OLS regression (no controls) clustered at the individual level.
Results of OLS and Probit regressions can be found in Appendix B, Table B1.

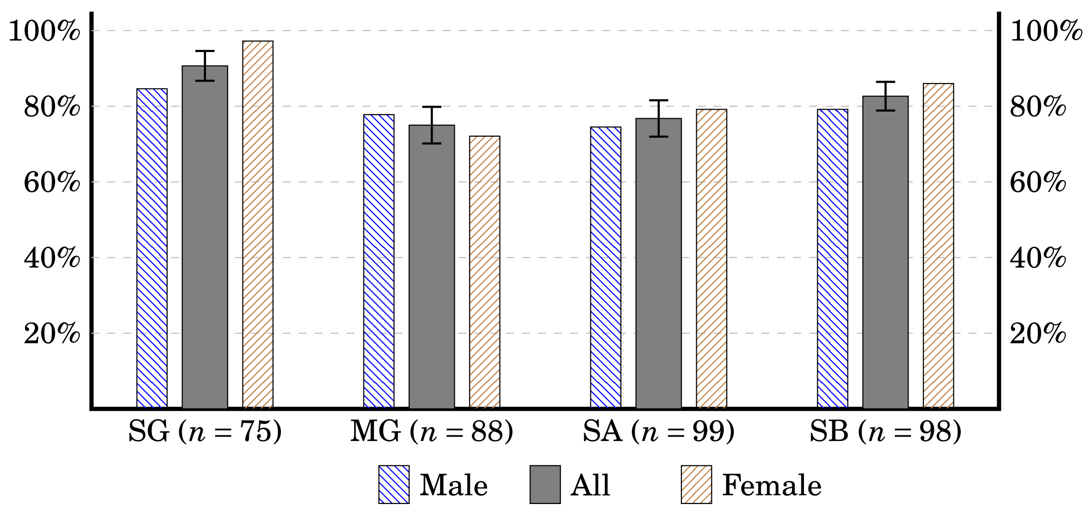


Figure 3: Proportions of allocating downward, by gender composition
Error bars represent standard errors from an OLS regression (no controls) clustered at the individual level.
SG/MG=rank $r-1$ and $r+1$ have the same/different gender. SA/SB=the rank above/below has same gender.
Results of OLS and Probit regressions can be found in Appendix B, Table B2.

Figure 3 shows the proportion of decisions allocating additional money to the Rank-Below, collated by gender composition. Compared to other situations, there was a distinct tendency to allocate money to the Rank-Below in the SG (t-test[21], $p = 0.005$). The proportion of downward allocation in the MG was the lowest, and it was not statistically different from that in the SA ($p = 0.895$). The proportion of downward allocation in the SB was slightly larger than that in the MG, but the difference was not statistically significant ($p = 0.151$) either. These observations imply that the participants were more inclined to allocate the money to the Rank-Below, especially when the recipient of the money and the Rank-Above shared their gender. Moreover, this tendency of allocating more to the Rank-Below was driven by the behavior of female participants. In the SG, almost every female (97%) allocated money to the female Rank-Below, while 85% of males allocated it to the male Rank-Below ($p = 0.126$).[22]

Result 2. *Participants allocated downward more in the same-gender situation than in the mixed-gender situation. This tendency was driven by females.*

To further examine the significant differences between the SG and MG, we checked the proportions of downward allocation by rank only for these two situations. We found suggestive evidence that participants at ranks 2 and 5 allocated downward less often in the MG. Figure 4 shows the proportions of downward allocation in the SG and MG, collated by rank. The participants holding ranks 2 and 5 allocated additional money to the Rank-Below significantly more frequently when the adjacent participants were of the same gender ($p = 0.001$). This difference was more distinctive for females ($p = 0.014$) than for males ($p = 0.048$),[23] and many second-ranked and fifth-ranked female participants were reluctant to bring the male

above. This evidence suggests that the gender composition of the participants two or more ranks away would not lead to noticeable differences. Indeed, the number of females in a group has no statistically significant effect (p=0.35) on allocation behavior when controlling for the genders of the decision maker and the adjacent players. This holds also true when only considering female (p=0.82) or male (p=0.76) decision makers.

[21] Unless otherwise stated, we have reported the p value of an estimated coefficient of a linear regression model controlling for ranks. The standard errors for the test statistics are clustered at the individual level. For regression results, including Probit regression results that are not reported in the main context, see Appendix B.

[22] This observation might be due to in-group bias based on gender: Rudman and Goodwin (2004) reported that women tend to be more favorable to other women than men to other men. Conversely, given that the patterns of the overall observations are weak but similar for men, this observation might indicate that women are more sensitive to the experimental context, as reported in Croson and Gneezy (2009) and Miller and Ubeda (2012).

[23] When separately examining the effect for: 1) the male participants ranked fifth, 2) the male participants ranked second, 3) the female participants ranked fifth, and 4) the female participants ranked second, the effect is statistically significant only for the last case ($p = 0.027$). However, there are only a few (16–20) observations per case, so we pooled the observations for ranks 2 and 5.

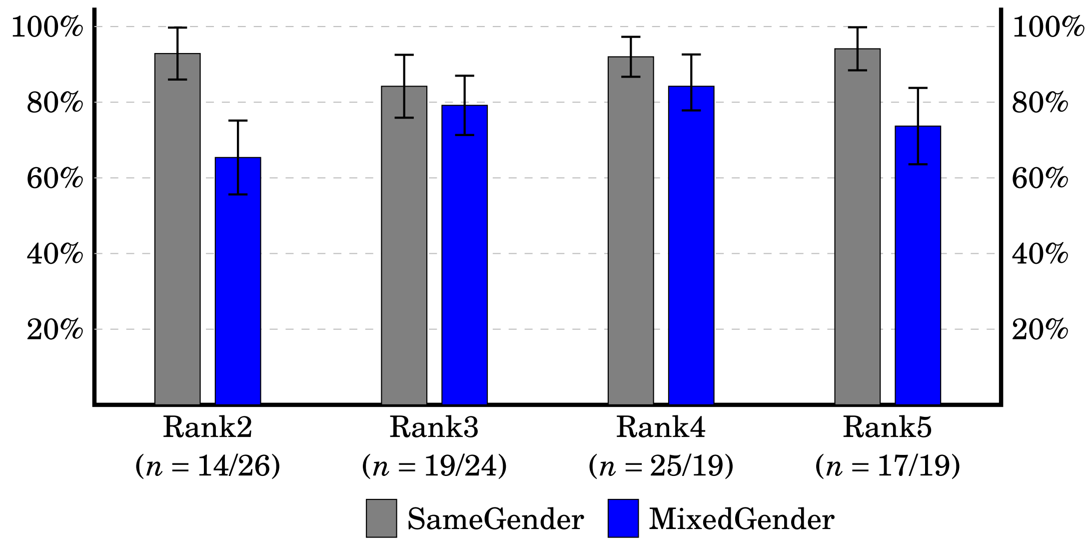


Figure 4: Proportions of decisions allocating to the Rank-Below in the SG and MG, by rank
Error bars represent standard errors from an OLS regression (no controls) clustered at the individual level. Results of OLS and Probit regressions can be found in Appendix B, Table B4.

Rank-Below up to the same rank. Our preferred interpretation for participants ranked fifth is that their reluctance to allocate downward is related to mixed-gender LPA. The reluctance of participants ranked second to allocate downward is also observed by Martinangeli and Windsteiger (2021) in their treatment involving fixed ranks and no rank reversals.[24] A possible explanation for this pattern in the MG is that participants want to avoid rankings that are dominated by other genders. If a second-ranked female allocated downward, two men would end up in the first place and (a shared) second place.[25]

3.2 Perceptions of Performance and Allocation Decisions

Participants were generally overconfident about their actual performance rank in their post-experiment survey, with the average difference between the guessed performance rank and the actual performance rank among the 18 participants per session being equal to 1.68 (SD

[24]Xie et al. (2017) show that even third-party dictators have an aversion to change a pre-existing ranking. Martinangeli and Windsteiger (2021) thus argue that allocations involving rank reversals would not allow for a clean test of last-place aversion.

[25]Some may argue that the second-ranked participants' reluctance to allocate downward might be related to first-place loving (Gill et al., 2019), but neither mixed-gender first-place loving nor within-gender first-place loving can clearly explain our observation in the MG. If a female participant cares for a relative standing with respect to male participants, she ranks second, and her decision is irrelevant to the first-place loving. If a female participant cares for a relative standing among female participants, she ranks first, but neither of her decision affects her within-gender rank.

5.58). This difference was statistically significant (t-test, $p = 0.002$). Male participants were more overconfident than female participants regarding their actual performance rank.[26] Figure 5 shows the empirical cumulative distribution functions of the actual performance rank by session and perceived performance rank. Male participants' guesses first-order stochastically dominate their actual ranks, implying that the men were overconfident, regardless of their actual performance. Female participants were also overconfident, but the magnitude was much smaller, with some high-performing females less confident about their relative performance rank. On average, male participants guessed that their relative performance rank was 2.5 ranks (SD 5.40) better than their actual rank, while female participants guessed that their rank was 0.85 ranks (SD 5.67) better. Although statistically insignificant, the average overconfidence of males is greater than females ($p = 0.054$, controlling for actual performance rank).

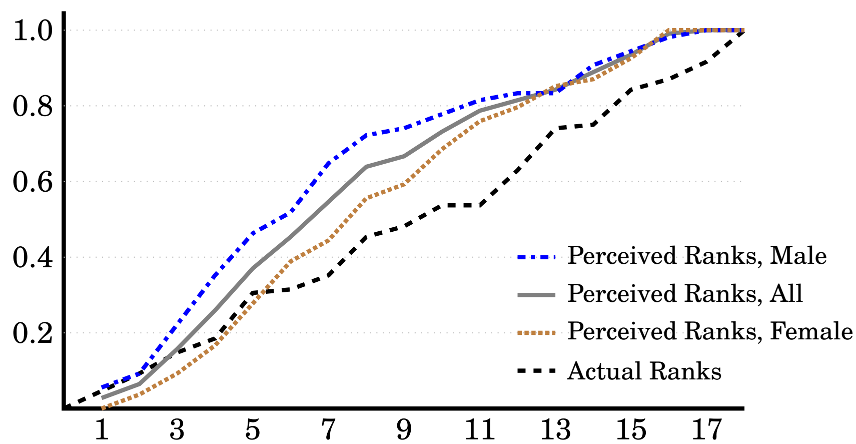


Figure 5: Empirical cumulative distribution functions for relative ranks

Moreover, we also notice that the participants whose performance was worse tend to overestimate their relative performance in a more significant manner: When actual rank increases by 1,[27] the gap between actual rank and guessed rank increases by 0.74, on average ($p < 0.001$). Combined with the gender differences in the perception of their relative performance rank, this observation suggests that low-performing male participants were more likely to believe that their rank, and accordingly, the allocation of money was worse

[26]This finding is in line with Niederle and Vesterlund (2007), who found that men are more overconfident about their relative performance on a math task, although there is no significant gender difference in actual performances.

[27]Since the participant with the best performance is ranked first (1) and the one with the worst performance is ranked last (18), an increase in rank means that the relative rank falls. We are aware of this potential confusion, so we clarified its meaning to the experiment participants.

than they deserved. This conjecture leads us to examine LPA for males only: If a male participant learns that his rank is five out of six, he might feel that the rank is significantly worse than his performance rank, driving him not to allocate money to his Rank-Below. It is indeed the case that the fifth-ranked male participants in the MG allocated to the Rank-Below less often than the males holding ranks 2, 3, and 4. However, this difference was not statistically significant ($p = 0.261$).

3.3 Perceptions of Luck and Allocation Decisions

During the post-experiment survey, we asked the participants to evaluate the following three statements on a 1–5 Likert scale:

1. "I feel that, overall, the random adjustment of the performance places me worse ranks than my actual ranks."
2. "I feel that the effect of random adjustment on the relative rank is stronger than that of actual performance."
3. "A society should accept inequality caused by differences in luck."

The first question was used to check the validity of the survey answers. If the participants strongly agreed with the statement, they would report their performance rank better than the rank from Part 2; thus, the difference between the actual performance rank and the perceived performance rank would increase. Confirming this narrative, we found a strong and positive correlation between them (ρ=0.2160, $p < 0.001$).

The second question asked about the perceived importance of the luck factor. We examined whether participants who felt that the luck factor's impact was strong behaved differently from those who did not. We found that male participants agreed that the luck factor's impact was stronger than the actual performance (KS, $p = 0.017$) more often than female participants, echoing the above-reported higher overconfidence of males. Moreover, we found that the perceived importance of the luck factor does not significantly affect allocation decisions in any way.[28]

Regarding the last question, we found that the participants who strongly agreed with the statement tended to allocate additional money to the Rank-Below less often ($p = 0.061$). From the perspective of the decision maker, allocating downward involves giving up one's

[28]This also holds true when controlling for the interaction of the perceived importance of the luck factor with the rank obtained.

rank to mitigate inequality. Our preferred interpretation for this observation is that those who embrace luck-driven inequality would be less inequality-averse and put relatively more weight on their concern for relative rank.

Result 3. *The participants' beliefs about how much their performance-based rank was higher than the announced rank did not affect their allocation decisions. Participants with a general attitude of accepting inequality caused by differences in luck tended to allocate downward less often.*

4 Conclusions

In our experiment, we added a gender composition factor to the existing studies on relative rank and distributional outcomes. We hypothesized that decision makers' weightings between concerns for their relative position and equality of outcomes might depend on the gender of the potentially affected persons. Thus, the gender composition of the reference groups could affect allocation decisions. However, there is no clear indication of whether people are relatively more or less sensitive to rank concerns when affected persons are of the same or opposite gender. This uncertainty motivated the present experiment.

We let the participants encounter different payoff-relevant ranks, innocuously provided the gender information of others, and asked them to allocate additional money to a person at one rank below or above. The crucial variation was whether the Rank-Below was of the same gender. When the Rank-Below was of the different gender, allocating downward did not alter the decision maker's within-gender rank. The experiment participants allocated downward more often when the Rank-Below was of the same gender. This tendency is more distinctive when the Rank-Above also was of the same gender. As this occurred primarily among women, we interpreted that solidarity is higher among women. Allocating money to a female Rank-Below minimizes inequality within the group of women, whereas allocating money to a female Rank-Above exacerbates income differences.

Meanwhile, we found no support for LPA. The participants at the fifth rank (one rank above the bottom) allocated money downward as much as the participants at other ranks. Combined with the findings on same-gender solidarity, these results indicate that gender composition is pertinent in allocation decisions that involve a rank-inequality tradeoff, potentially more than the overall rank.

Future research is needed to investigate the many underexplored questions on the tradeoff between concerns for relative rank and distributional outcomes. For the current study, the participant's performance and luck were similarly attributed to determine the observed ranks, but it would be worth investigating the effects of different rank-determining schemes, ranging from pure luck to pure performance. The effects of group size, variation in ranks over rounds, and varying degrees of income inequality relative to the magnitude of redistribution could also be investigated. Another possible direction of future research is to utilize our experimental design—using nicknames to innocuously reveal some characteristic information about the subjects—for investigating the effect of individual characteristics other than gender, e.g., immigration status and ethnicity, on the preferences for redistribution.

References

Alesina, Alberto and Paola Giuliano, "Preferences for redistribution," in "Handbook of Social Economics," Vol. 1, Elsevier, 2011, pp. 93–131.

__ , Armando Miano, and Stefanie Stantcheva, "Immigration and Redistribution," *The Review of Economic Studies*, 03 2022, *90* (1), 1–39.

__ , Stefanie Stantcheva, and Edoardo Teso, "Intergenerational Mobility and Preferences for Redistribution," *American Economic Review*, February 2018, *108* (2), 521–554.

Arechar, Antonio A., Simon Gächter, and Lucas Molleman, "Conducting interactive experiments online," *Experimental Economics*, 2018, *21*, 99–131.

Aronsson, Thomas and Olof Johansson-Stenman, "Positional Concerns in an OLG Model: Optimal Labor and Capital Income Taxation," *International Economic Review*, 2010, *51* (4), 1071–1095.

__ and __ , "Paternalism against Veblen: Optimal Taxation and Non-respected Preferences for Social Comparisons," *American Economic Journal: Economic Policy*, February 2018, *10* (1), 39–76.

Baranski, Andrzej, Diogo Geraldes, Ada Kovaliukaite, and James Tremewan, "An experiment on gender representation in majoritarian bargaining," *Management Science*, 2023.

Ben-Ner, Avner, Fanmin Kong, and Louis Putterman, "Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving," *Journal of Economic Psychology*, 2004, *25* (5), 581–589.

Benabou, Roland, "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, 2000, *90* (1), 96–129.

Bertrand, Marianne and Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, September 2004, *94* (4), 991–1013.

Bilén, David, Anna Dreber, and Magnus Johannesson, "Are women more generous than men? A meta-analysis," *Journal of the Economic Science Association*, 2021, *7* (1), 1–18.

Camerer, Colin F., Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, Emma Heikensten, Felix Holzmeister, Taisuke Imai, Siri Isaksson, Gideon Nave, Thomas Pfeiffer, Michael Razen, and Hang Wu, "Evaluating replicability of laboratory experiments in economics," *Science*, 2016, *351* (6280), 1433–1436.

Chakraborty, Priyanka and Danila Serra, "Gender and Leadership in Organisations: the Threat of Backlash," *The Economic Journal*, 12 2023, p. uead110.

Clark, Andrew E., Paul Frijters, and Michael A. Shields, "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles," *Journal of Economic Literature*, 2008, *46* (1), 95–144.

Corneo, Giacomo and Hans Peter Grüner, "Social limits to redistribution," *American Economic Review*, 2000, *90* (5), 1491–1507.

Croson, Rachel and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature*, June 2009, *47* (2), 448–474.

_, Melanie Marks, and Jessica Snyder, "Groups Work for Women: Gender and Group Identity in Social Dilemmas," *Negotiation Journal*, 2008, *24* (4), 411–427.

Cruces, Guillermo, Ricardo Perez-Truglia, and Martin Tetaz, "Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment," *Journal of Public Economics*, 2013, *98*, 100–112.

Dalmia, Prateik and Emel Filiz-Ozbay, "Your success is my motivation," *Journal of Economic Behavior & Organization*, 2021, *181*, 49–85.

Dasgupta, Utteeyo, Subha Mani, Smriti Sharma, and Saurabh Singhal, "Can gender differences in distributional preferences explain gender gaps in competition?," *Journal of Economic Psychology*, 2019, *70*, 1–11.

Duesenberry, James S., *Income, saving, and the theory of consumer behavior*, Cambridge: Harvard University Press, 1949.

Eckel, Catherine C. and Philip J. Grossman, "Are Women Less Selfish Than Men?: Evidence From Dictator Experiments," *The Economic Journal*, 12 2001, *108* (448), 726–735.

Enke, Benjamin, Mattias Polborn, and Alex A. Wu, "Values as Luxury Goods and Political Behavior," 2023. Working Paper.

Erkal, Nisvan, Lata Gangadharan, and Nikos Nikiforakis, "Relative Earnings and Giving in a Real-Effort Experiment," *American Economic Review*, December 2011, *101* (7), 3330–3348.

Fisman, Raymond, Ilyana Kuziemko, and Silvia Vannutelli, "Distributional Preferences in Larger Groups: Keeping up with the Joneses and Keeping Track of the Tails," *Journal of the European Economic Association*, 09 2020, *19* (2), 1407–1438.

Frank, Robert H., "The Demand for Unobservable and Other Nonpositional Goods," *The American Economic Review*, 1985, *75* (1), 101–116.

__ , "Should Public Policy Respond to Positional Externalities?," *Journal of Public Economics*, 2008, *92* (8-9), 1777–1786. Special Issue: Happiness and Public Economics.

Frank, Thomas, *What's the Matter with Kansas?: How Conservatives Won the Heart of America*, Henry Holt and Company, 2004.

Funk, Patricia and Christina Gathmann, "Gender gaps in policy making: Evidence from direct democracy in Switzerland," *Economic Policy*, 2015, *30* (81), 141–181.

Gallice, Andrea and Edoardo Grillo, "Economic and Social-Class Voting in a Model of Redistribution with Social Concerns," *Journal of the European Economic Association*, 11 2019, *18* (6), 3140–3172.

Gill, David, Zdenka Kissová, Jaesun Lee, and Victoria Prowse, "First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision," *Management Science*, 2019, *65* (2), 494–507.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini, "Performance in Competitive Environments: Gender Differences," *The Quarterly Journal of Economics*, 2003, *118* (3), 1049–1074.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "People's Opium? Religion and Economic Attitudes," *Journal of Monetary Economics*, 2003, *50* (1), 225–282.

Hirsch, Fred, *Social Limits to Growth*, Taylor & Francis, 1995.

Hopkins, Ed and Tatiana Kornienko, "Running to Keep in the Same Place: Consumer Choice as a Game of Status," *American Economic Review*, September 2004, *94* (4), 1085–1107.

Houtman, Dick, Peter Achterberg, and Anton Derks, *Farewell to the Leftist Working Class*, Transaction Publishers, 2008.

Kelly, Nathan J. and Peter K. Enns, "Inequality and the Dynamics of Public Opinion: The Self-Reinforcing Link Between Economic Inequality and Mass Preferences," *American Journal of Political Science*, 2010, *54* (4), 855–870.

Keynes, John M., "Economic Possibilities for Our Grandchildren," in "Essays in Persuasion," New York: Harcourt Brace (1932), 1930, pp. 358–373.

Kim, Duk Gyoo, "Positional concern and low demand for redistribution of the poor," *European Journal of Political Economy*, 2019, *56*, 27–38.

Kuziemko, Ilyana, Ryan W. Buell, Taly Reich, and Michael I. Norton, "Last-Place Aversion: Evidence and Redistributive Implications," *The Quarterly Journal of Economics*, 2014, *129* (1), 105–149.

Mago, Shakun D. and Laura Razzolini, "Best-of-five contest: An experiment on gender differences," *Journal of Economic Behavior & Organization*, 2019, *162*, 164–187.

Martinangeli, Andrea and Lisa Windsteiger, "Last Word not Yet Spoken: Last Place and Rank Reversal Aversion," *Experimental Economics*, September 2021, *24*, 800–821.

Meltzer, Allan H. and Scott F. Richard, "A Rational Theory of the Size of Government," *Journal of Political Economy*, 1981, *89* (5), 914–927.

Mengel, Friederike and Elke Weidenholzer, "Preferences for redistribution," *Journal of Economic Surveys*, 2023, *37* (5), 1660–1677.

Miller, Luis and Paloma Ubeda, "Are women more sensitive to the decision-making context?," *Journal of Economic Behavior & Organization*, 2012, *83* (1), 98–104. Gender Differences in Risk Aversion and Competition.

Niederle, Muriel, "8. Gender," in John H. Kagel and Alvin E. Roth, eds., *The Handbook of Experimental Economics, Volume 2*, Princeton: Princeton University Press, 2017, pp. 481–562.

— **and Lise Vesterlund**, "Do women shy away from competition? Do men compete too much?," *The Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.

Rudman, Laurie A. and Stephanie A. Goodwin, "Gender differences in automatic ingroup bias: why do women like women more than men like men?," *Journal of personality and social psychology*, 2004, *87* (4), 494–509.

Russell, Bertrand., *The Conquest of Happiness*, Allen & Unwin, 1930.

Selten, Reinhard and Axel Ockenfels, "An experimental solidarity game," *Journal of Economic Behavior & Organization*, 1998, *34* (4), 517–539.

Sharma, Smriti, "Gender and distributional preferences: Experimental evidence from India," *Journal of Economic Psychology*, 2015, *50*, 113–123.

Shastry, Gauri Kartini, Olga Shurchkov, and Lingjun Lotus Xia, "Luck or skill: How women and men react to noisy feedback," *Journal of Behavioral and Experimental Economics*, 2020, *88*, 101592.

Solnick, Sara J., "Gender differences in the ultimatum game," *Economic Inquiry*, 2001, *39* (2), 189–200.

Veblen, Thorstein, *The Theory of the Leisure Class: An Economic Study of Institutions.*, New York: The Macmillan Company, 1899.

Xie, Wenwen, Benjamin Ho, Stephan Meier, and Xinyue Zhou, "Rank reversal aversion inhibits redistribution across societies," *Nature Human Behaviour*, 2017, *1* (8), 1–5.

Zizzo, Daniel J., "Experimenter demand effects in economic experiments," *Experimental Economics*, 2010, *13*, 75–98.

A Appendix - Experimental Instructions

Thank you for your participation. Before moving on to the main experiment, please complete a short survey.

[Ask participants' gender, age, a field of study, and experience of online experiments.]

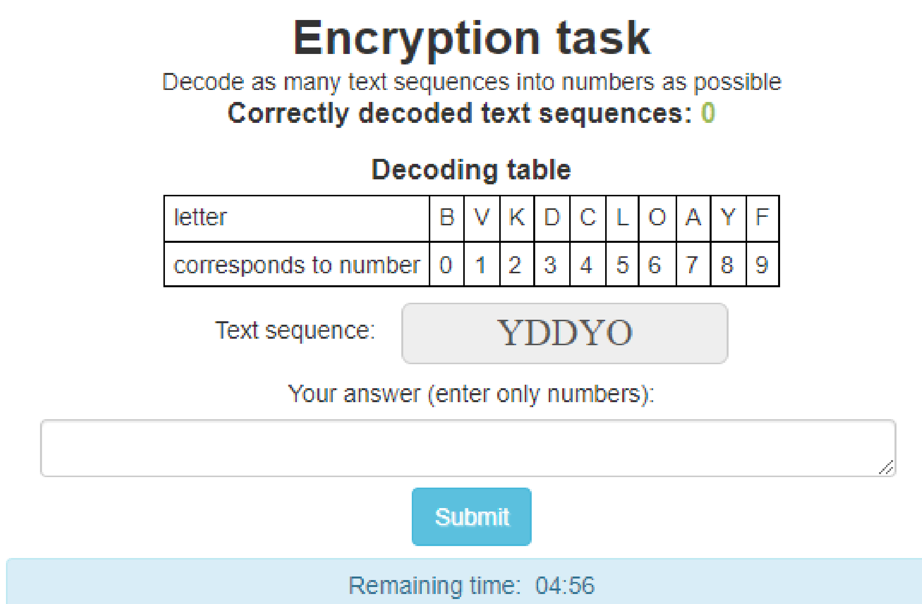
In this experiment, your earnings will be based on your decisions, others' decisions, and some luck. It is important that you fully understand the following instructions. Please read them carefully.

Overview

This experiment mainly consists of two parts. In Part 1 of the experiment, which we call WORKING, you will work on a task for 5 minutes. In Part 2 of the experiment, which we call DECISIONS, you will report your decisions under a given situation. Your earnings in this experiment will depend on both—your task performance and random factors. More details follow.

Part 1: WORKING

You will perform a so-called Encryption Task. Specifically, you will read sequences of random letters. These have to be **decoded into numbers** with the given decoding table. For each text sequence, the decoding table changes. Below is an example of such a sequence.



In the example, you can see the text sequence *YDDYO*. The decoding table tells you that B=0, V=1, and so on. This means that you have to decode *YDDYO* into *83386* and enter this numeric value into the answer field.

You will have **5 minutes** to correctly decode as many text sequences into numbers as possible. Note that each letter must be decoded correctly. After entering the decoded text sequence, hit the submit button. Subsequently, irrespective of whether the text sequence was decoded correctly or not, a new text sequence and a new decoding table will appear. You will see the remaining time in the upper right corner of the screen. **The more correct answers, the more money you will earn.** We will provide you more detail in the second part of the experiment.

Quiz Please answer the following two questions. If you answer one or more questions wrongly, you have to re-take the quiz.

Q1 What is the aim of the task? (A) Write meaningful words from the given letters. (B) Decode the letters into numbers. (C) Reorder the letters according to the alphabet. (D) Assign a value to each letter and calculate the total value.

Q2 Which of the following statements is **false**? (A) For each text sequence, the decoding table changes. (B) The answer only contains numbers. (C) Regardless of whether you decode the sequence correctly, a new text sequence will appear. (D) You have as much time as you want to decode the text sequences.

[participants perform the real-effort task for 5 minutes.]

Part 2: DECISIONS

This part of the experiment consists of 5 rounds. At the beginning of each decision round, we want you to pick your nickname among the list of alternatives. This way, you and others are completely anonymous, yet you can clearly distinguish other participants by their nicknames. (Everyone receives different options, so do not worry about possible duplication within a decision round. If you see the same names in other rounds, it is the result of pure coincidence, not implying the same person picking the same name twice.)

In each round, after you pick your nickname, a computer randomly assigns you to a group of six and displays your and other group members' **relative rank.** The rank is determined by your performance within Part 1 (WORKING) **and the unknown outcomes of a dice roll**. Specifically, the rank is determined as follows.

1. After the computer rolls a dice, the "adjusted performance" is determined. For simplicity, let N denote the number of correct answers in Part 1.

    - If your dice roll is 1, then your adjusted performance is $3\times N$.

- If your dice roll is 2, then your adjusted performance is $4 \times N$.
- If your dice roll is 3, then your adjusted performance is $5 \times N$.
- If your dice roll is 4, then your adjusted performance is $6 \times N$.
- If your dice roll is 5, then your adjusted performance is $7 \times N$.
- If your dice roll is 6, then your adjusted performance is $8 \times N$.

So, for example, if you have 10 correct answers in Part 1, your adjusted performance could vary from 30 to 80.

2. The relative rank shows the rank of the adjusted performance, not the rank of the actual performance.

3. You will not learn the realizations of dice rolls.

For illustration, suppose that your rank is higher than the rank of another member. A higher rank could mean (1) that your performance in Part 1 (WORKING) was better than her, or (2) that your dice roll outcome was higher than her. When your adjusted performance happens to be the same as another member's, a computer randomly determines who receives the higher rank.

In each round, you and your group members will receive different amounts of money according to the ranking. The person ranked first will receive 15€, the person ranked second will receive 13€, the person ranked third will receive 11€, and so on.

After you observed the initial money amounts, everyone will be asked to choose who should get more money. The additional money comes from an **external source** and does not take away from the amount of money you have. The choices you make are private and will not be shown to anyone playing the game.

Once everyone in your group has decided, the computer will randomly select one group member and award the additional money according to the selected member's decision. At that point, the computer records everyone's updated earnings, but it will not show your final earnings from the round.

For every new round, all participants are reshuffled, and you are randomly assigned to a new group of six. The computer rolls new dices for the new round, which are independent of the previous outcomes. The new ranking is then determined based on your performance within Part 1 (Working), your new group members' performance, and the new dice rolls.

Payment

At the end of the experiment, the server computer will randomly select one of the five rounds with equal probability. Your final earnings in that selected round will be your actual payment. Since each round is equally likely selected, it is in your best interest to take every round equally seriously.

Quiz

Please answer the following two questions. If you answer one or more questions wrongly, you have to re-take the quiz.

Q1 Suppose Adam and Bea are in the same group. Adam solved 10 tasks and Bea solved 11 tasks within Part 1 (WORKING). Which of the following statements is **false**? (A) The "adjusted performance" of Adam could range from 30 to 80. (B) Bea will hold a better rank than Adam in any case. (C) If their dice rolls are equal, Bea will hold a better rank than Adam. (D) If Bea holds a better rank than Adam, she will initially receive a higher amount of money.

Q2 Which of the following statements is **true**? (A) Your nickname will remain the same during all five rounds. (B) At the end of each round, the choices of all group members are implemented. (C) Your group members will remain the same during all five rounds. (D) The additional money you are asked to give to another group member comes from an external source.

[Participants make redistribution decisions for 5 rounds. See Figure A1 for illustration.]

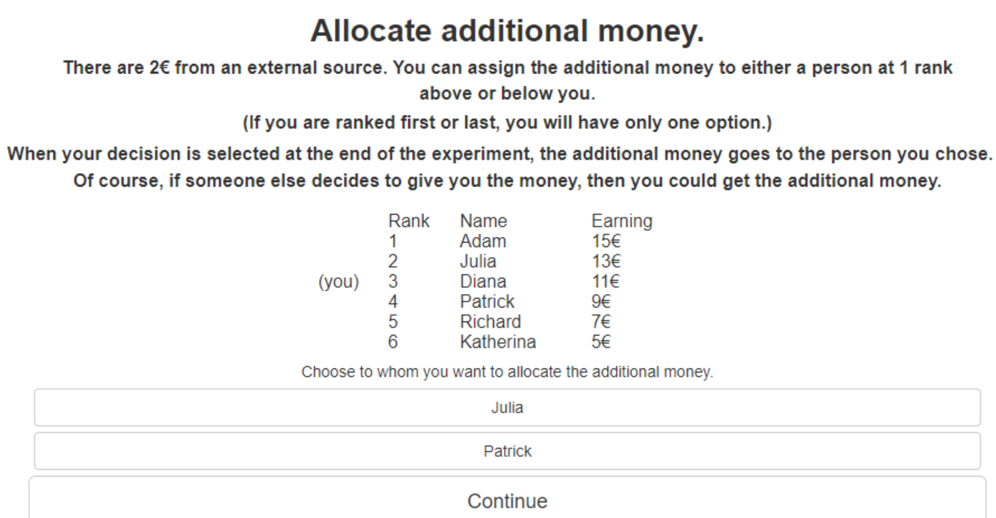


Figure A1: Screenshot of Part 2

[participant filled out post-experimental survey questions.]

Survey

1. There were 18 (including you) in the experiment you just finished. Please give your best guess about your actual rank of the performance. (Write an integer between 1 and 18.)
2. Please specify how you agree on the following statement: "I feel that, overall, the random adjustment of the performance places me worse ranks than my actual ranks."
3. Please specify how you agree on the following statement: "I feel that the effect of random adjustment on the relative rank is stronger than that of actual performance."
4. Please specify how you agree on the following statement: "A society should accept inequality caused by differences in luck."

B Appendix - Regression results

	OLS			Probit		
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled	Females	Males	Pooled	Females	Males
rank2	-0.0889	-0.121	-0.0563	-0.333	-0.490	-0.198
	(0.0669)	(0.101)	(0.0919)	(0.249)	(0.393)	(0.324)
rank3	-0.0667	-0.0889	-0.0444	-0.258	-0.379	-0.159
	(0.0482)	(0.0740)	(0.0649)	(0.192)	(0.323)	(0.237)
rank4	-0.0333	-0.0253	-0.0396	-0.137	-0.124	-0.143
	(0.0542)	(0.0604)	(0.0884)	(0.223)	(0.296)	(0.321)
_cons	0.856^{***}	0.889^{***}	0.822^{***}	1.061^{***}	1.221^{***}	0.924^{**}
	(0.0467)	(0.0531)	(0.0778)	(0.204)	(0.278)	(0.296)
N	360	177	183	360	177	183

The dependent variable is a dummy indicating whether a participant allocated downward. Observations holding ranks 1 and 6 are omitted. Standard errors clustered at the individual level are in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table B1: Rank effects

	OLS			Probit		
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled	Females	Males	Pooled	Females	Males
SG	0.157^{**}	0.251^{**}	0.0684	0.646^{*}	1.329^{*}	0.255
	(0.0568)	(0.0909)	(0.0663)	(0.259)	(0.516)	(0.266)
SA	0.0177	0.0707	-0.0327	0.0567	0.227	-0.106
	(0.065)	(0.109)	(0.0755)	(0.208)	(0.344)	(0.24)
BS	0.0765	0.139	0.0139	0.266	0.495	0.0475
	(0.0533)	(0.0795)	(0.0742)	(0.182)	(0.26)	(0.253)
_cons	0.750^{***}	0.721^{***}	0.778^{***}	0.674^{***}	0.586^{*}	0.765^{***}
	(0.0513)	(0.0847)	(0.0601)	(0.161)	(0.25)	(0.2)
N	360	177	183	360	177	183

The dependent variable is a dummy indicating whether a participant allocated downward. Observations holding ranks 1 and 6 are omitted. Standard errors clustered at the individual level are in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table B2: Effects of gender composition

	OLS			Probit		
	(1) Pooled	(2) Females	(3) Males	(4) Pooled	(5) Females	(6) Males
rank2	-0.0849	-0.109	-0.0561	-0.322	-0.468	-0.194
	(0.0632)	(0.0921)	(0.0910)	(0.239)	(0.377)	(0.319)
rank3	-0.0737	-0.0884	-0.0527	-0.299	-0.429	-0.188
	(0.0486)	(0.0744)	(0.0676)	(0.197)	(0.329)	(0.249)
rank4	-0.0474	-0.0401	-0.0505	-0.202	-0.191	-0.182
	(0.0556)	(0.0612)	(0.0911)	(0.231)	(0.305)	(0.327)
SG	0.152^{**}	0.240^{**}	0.0685	0.636^{*}	1.313^{**}	0.263
	(0.0551)	(0.0876)	(0.0660)	(0.257)	(0.501)	(0.271)
SA	0.00924	0.0557	-0.0364	0.0289	0.161	-0.110
	(0.0652)	(0.110)	(0.0764)	(0.209)	(0.353)	(0.238)
SB	0.0746	0.129	0.0168	0.262	0.460	0.0643
	(0.0518)	(0.0759)	(0.0736)	(0.175)	(0.251)	(0.243)
_cons	0.805^{***}	0.789^{***}	0.818^{***}	0.898^{***}	0.907^{**}	0.904^{**}
	(0.0552)	(0.0786)	(0.0812)	(0.208)	(0.297)	(0.290)
N	360	177	183	360	177	183

The dependent variable is a dummy indicating whether a participant allocated downward. SG, SB, and SA respectively refer to a *same-gender*, *same-above*, and *same-below* situation. Observations holding ranks 1 and 6 are omitted. Standard errors clustered at the individual level are in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table B3: Effects of rank and gender composition

	OLS		Probit	
	(1)	(2)	(3)	(4)
	SG	MG	SG	MG
rank2	-0.0126	-0.0830	-0.0995	-0.238
	(0.0920)	(0.148)	(0.707)	(0.422)
rank3	-0.0991	0.0548	-0.562	0.179
	(0.0754)	(0.131)	(0.436)	(0.415)
rank4	-0.0212	0.105	-0.160	0.370
	(0.0815)	(0.141)	(0.617)	(0.493)
_cons	0.941^{***}	0.737^{***}	1.565^{**}	0.634^{*}
	(0.0594)	(0.104)	(0.496)	(0.312)
N	75	88	75	88

The dependent variable is a dummy indicating whether a participant allocated downward. SG and MG respectively refer to a *same-gender* and *mixed-gender* situation. Observations holding ranks 1 and 6 are omitted. Standard errors clustered at the individual level are in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table B4: Rank effects in SG and MG

	OLS			Probit		
	(1)	(2)	(3)	(4)	(5)	(6)
rank2	-0.0888	-0.0871	-0.0948	-0.333	-0.319	-0.363
	(0.0672)	(0.0657)	(0.0676)	(0.249)	(0.244)	(0.257)
rank3	-0.0667	-0.0612	-0.0737	-0.258	-0.232	-0.281
	(0.0483)	(0.0471)	(0.0490)	(0.192)	(0.186)	(0.198)
rank4	-0.0334	-0.0272	-0.0241	-0.137	-0.105	-0.103
	(0.0537)	(0.0517)	(0.0531)	(0.221)	(0.213)	(0.221)
PL	-0.00123			-0.00599		
	(0.0250)			(0.0942)		
IL		0.0325			0.116	
		(0.0297)			(0.109)	
SL			-0.0487			-0.183
			(0.0258)			(0.0940)
_cons	0.860***	0.734***	0.982***	1.081**	0.628	1.558***
	(0.0890)	(0.107)	(0.0706)	(0.353)	(0.376)	(0.321)
N	360	360	360	360	360	360

The dependent variable is a dummy indicating whether a participant allocated downward. PL (Perception of Luck), IL (Importance of Luck), and SL (Society embracing Luck) respectively refer to the Likert-scale answers on survey questions 2–4. Observations holding ranks 1 and 6 are omitted. Standard errors clustered at the individual level are in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table B5: Effects of post-experimental perceptions

Online Appendix for
Rank versus Inequality—Does Gender Composition Matter?

Duk Gyoo Kim, Sungkyunkwan University
Max Riegel, University of Mannheim

This online appendix consists of two sections. Section A describes the script of a survey (and the results) conducted at Amazon Mechanical Turk (MTurk). Based on the responses to this survey, we construct the list of gender-specific names. Section B elaborates on how we calibrated the range of the random scaling factor that determined the adjusted performance.

A Mechanical Turk survey for unambiguous first names

1. **Preliminary Information**: This survey is part of a research project by Duk Gyoo Kim and Max Riegel at the University of Mannheim. Your participation in this research study is voluntary. If you decide to participate in this research survey, you may withdraw at any time. The survey will take approximately 10 minutes. Your responses will be confidential, and we do not collect identifying information such as your name, email address, or IP address. All data is stored in a password-protected electronic format. The results of this study will be used for scholarly purposes only.

Important Disclaimer: The survey will contain two attention-check questions. You will be asked to solve two simple mathematical tasks. You will only be paid if you answer those questions correctly! Please avoid random answers. Since we spread 20 questions whose answers of previous participants were 100% identical, we can cross-validate your answers.

2. Attention Check: Choose the lowest number. (10 options are given.)

3. What is your gender?

Response	Frequency
male	52
female	28

4. Where do you live?

Response	Frequency
Germany, Austria or Switzerland	29
Other European country	37
Asia	14

5. How old are you?

Response	Frequency
younger than 18	0
18–29	36
30–39	21
40–49	18
50 or older than 50	5

6. Please classify the following names as male, female, or unsure/unisex. (Names were displayed in random order.)

Name	#M	#F	#U	Name	#M	#F	#U	Name	#M	#F	#U	Name	#M	#F	#U
Aaron	80	0	0	Adam	80	0	0	Alyssa	0	80	0	Elena	0	80	0
Albert	80	0	0	Anthony	80	0	0	Helena	0	80	0	Isabella	0	80	0
Arthur	80	0	0	Christopher	80	0	0	Jasmine	0	80	0	Jennifer	0	80	0
David	80	0	0	Derek	80	0	0	Lisa	0	80	0	Luisa	0	80	0
Edward	80	0	0	Ethan	80	0	0	Monica	0	80	0	Sarah	0	80	0
Frank	80	0	0	Fred	80	0	0	Sylvia	0	80	0	Alicia	1	79	0
Hugo	80	0	0	James	80	0	0	Angelica	1	79	0	Christina	1	79	0
Jason	80	0	0	Patrick	80	0	0	Jessica	1	79	0	Julia	1	79	0
Peter	80	0	0	Richard	80	0	0	Katherina	1	79	0	Lara	1	79	0
Robert	80	0	0	Ronald	80	0	0	Laura	1	79	0	Marisa	1	79	0
Steven	80	0	0	Thomas	80	0	0	Olivia	1	79	0	Rebecca	1	79	0
Walter	80	0	0	Xavier	80	0	0	Regina	1	79	0	Samantha	1	79	0
Benjamin	79	1	0	Brandon	79	0	1	Sandra	1	79	0	Selina	1	79	0
Bruno	79	1	0	Charles	79	1	0	Teresa	1	79	0	Claudia	2	78	0
Christian	79	0	1	Daniel	79	0	1	Cynthia	2	78	0	Elizabeth	2	78	0
Diego	79	0	1	Eric	79	1	0	Ella	2	78	0	Eva	2	78	0
Gregor	79	1	0	Isaac	79	1	0	Lena	1	78	1	Linda	2	78	0
Ivan	79	1	0	Jacob	79	1	0	Lydia	2	78	0	Melissa	2	78	0
John	79	1	0	Jonathan	79	0	1	Mia	2	78	0	Michaella	2	78	0
Kevin	79	1	0	Leonard	79	1	0	Natasha	2	78	0	Rose	2	78	0
Manuel	79	0	1	Marco	79	0	1	Sabrina	2	78	0	Sophia	2	78	0
Matthew	79	1	0	Nicholas	79	1	0	Ariana	3	77	0	Bianca	3	77	0
Norman	79	0	1	Paul	79	1	0	Cassandra	3	77	0	Clara	3	77	0
Phillip	79	1	0	Ralph	79	1	0	Daniela	3	77	0	Diana	3	77	0
Samuel	79	1	0	Scott	79	1	0	Emily	3	77	0	Emma	3	77	0
Stanley	79	0	1	Victor	79	1	0	Erica	3	77	0	Lily	2	77	1
Vincent	79	0	1	William	79	1	0	Maya	2	77	1	Vanessa	3	77	0
Alexander	78	1	1	Andrew	78	0	2	Alina	3	76	1	Amanda	4	76	0
Brian	78	2	0	George	78	0	2	Gina	3	76	1	Gloria	4	76	0
Henry	78	2	0	Oscar	78	1	1	Jacqueline	3	76	1	Madeleine	4	76	0
Owen	78	1	1	Roman	78	1	1	Maria	3	76	1	Miriam	3	76	1
Ryan	78	1	1	Stuart	78	1	1	Stella	4	76	0	Victoria	4	76	0
Alan	77	2	1	Bill	77	1	2	Anna	5	75	0	Caroline	5	75	0
Fabian	77	2	1	Harry	77	2	1	Chiara	3	75	2	Joanna	5	75	0
Jackson	77	1	2	Jeffrey	77	3	0	Kathryn	5	75	0	Melanie	5	75	0
Marvin	77	1	2	Oliver	77	1	2	Miranda	4	75	1	Nina	4	75	1
Ruben	77	1	2	Sebastian	77	2	1	Patricia	4	75	1	Freya	1	74	5
Simon	77	2	1	Todd	77	0	3	Josephine	4	74	2	Veronica	6	74	0
Carl	76	4	0	Joseph	76	3	1	Amy	5	73	2	Claire	8	72	0
Lucas	76	3	1	Michael	76	1	3	Natalie	8	72	0	Cindy	9	71	0
Felix	75	1	4	Timothy	75	3	2	Hannah	9	71	0	Ida	3	71	6
Finn	74	2	4	Gary	74	5	1	Kira	6	71	3	Evelyn	10	70	0
Magnus	74	0	6	Max	74	1	5	Stephanie	10	70	0	Vera	8	70	2
Sean	74	1	5	Dennis	73	3	4	Margaret	11	69	0	Alexandra	12	67	1
Liam	73	3	4	Warren	73	1	6	Rachel	10	67	3	Charlotte	12	66	2
Jasper	71	5	4	Jeremy	70	8	2	Leah	11	66	3	Megan	14	66	0
Noah	70	4	6	Marcel	68	6	6	Chloe	14	65	1	Judith	14	65	1
Nathaniel	68	7	5	Joshua	67	10	3	Kimberly	11	65	4	Michelle	15	64	1
Julian	64	11	5	Louis	62	12	6	Bea	10	63	7	Nicole	18	61	1
Valentin	60	9	11	Maurice	57	10	13	Carmen	16	58	6	Zoe	16	56	8

7. Attention Check 2: Choose the largest number. (10 options are given.)

B Calibration of the luck factor

This section elaborates on how we determine the size of the luck factor for the adjusted performance. A subject in a group learns the relative rank of the adjusted performance, which is a number of correct answers in the Encryption Task multiplied by a random (luck) factor.

It is worth noting that the governance of the luck factor increases with the variance of it. If the luck factor varies too much, then the relative rank of the adjusted performance would be merely determined by the luck factor. Meanwhile, if the luck factor varies too little, then the relative rank would be merely determined by the actual performance. That is, in terms of the linear relationship between a subject's actual rank in performance and the rank in luck-adjusted performance, as shown in Figure 1, must be between no correlation (Pure Luck) and perfect correlation (No Luck). We aimed to construct the distribution of the random variable (the luck scaling factor,) to satisfy two purposes: (1) The linear relationship between ranks in actual and adjusted performance lies in a halfway. (2) The random variable is easy enough to explain to laypeople without using math.

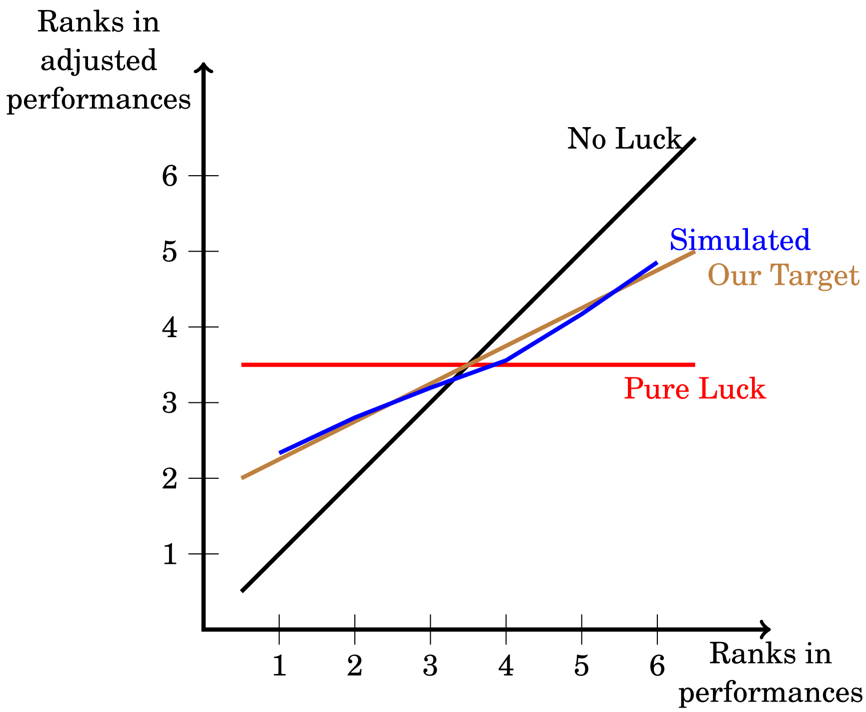


Figure 1: Simulated average ranks

To this end, we conducted a pilot test with six subjects (who are different from the experiment participants,) and learned that the average number of the correct answers in the Encryption Task for five minutes is about 25, and the standard deviation of it is about 5.2. We imagine a "representative" group which consists of six subjects whose actual performance measures are 32, 29, 27, 24, 21, and 18, respectively. For notational simplicity, denote T_r is the actual performance of subject r. To simulate the relationship between the ranks in actual and adjusted performances, we performed the following simulation:

1. Pick two parameters, $\alpha, \beta \in \mathbb{R}^2_+$.

2. Calculate the simulated ranks in adjusted performance. Initiate iteration index $t = 1$.
   - For each iteration $t \leq 500$,
   - draw a number from a discrete uniform distribution between 1 and 6 for each subject r, denoted by d_r^t,
   - calculate $A_r = T_r \times [\alpha + \beta d_r^t]$ for all $r = 1, 2, \ldots, 6$, and
   - record the relative rank of subject r, denoted by R_r^t.
   - Repeat the above steps with $t \Leftarrow t + 1$.
3. Check if $\frac{\overline{R}_6 - \overline{R}_1}{6-1} \in [0.5 - tol, 0.5 + tol]$, where $\overline{R}_r = \frac{1}{500}\sum_{t=1}^{500} R_r^t$ and $tol > 0$ is a small tolerance level.
   - If $\frac{\overline{R}_6 - \overline{R}_1}{6-1} \notin [0.5 - tol, 0.5 + tol]$, repeat from Step 1 with new parameters α and β.

After numerous simulations, we find $(\alpha, \beta) = (2, 1)$ reasonably achieves our goals—see blue line in Figure 1. This is how we ended up with [dice + 2] as a scaling parameter.