HARVARD Kennedy School
JOHN F. KENNEDY SCHOOL OF GOVERNMENT

Informational Inequity Aversion and Performance

Faculty Research Working Paper Series

Iris Bohnet
Harvard Kennedy School

Farzad Saidi
Stockholm School of Economics & CEPR

February 2019

RWP19-008

Informational Inequity Aversion and Performance*

Iris Bohnet†
Harvard University

Farzad Saidi‡
Stockholm School of Economics & CEPR

February 1, 2019

Abstract

In labor markets, some individuals have, or believe to have, less data on the determinants of success than others, e.g., due to differential access to technology or role models. We provide experimental evidence on when and how informational differences translate into performance differences. In a laboratory tournament setting, we varied the degree to which individuals were informed about the effort-reward relationship, and whether their competitor received the same or a different amount of information. We find performance is adversely affected only by worse relative, but not absolute, informedness. This suggests that inequity aversion applies not only to outcomes but also to information that helps achieve them, and stresses the importance of inequality in initial information conditions for performance-dependent outcomes.

JEL classification: D81, D82, M50

Keywords: tournament, information, inequity aversion, performance, effort task

*We wish to thank David Cesarini, Pedro Dal Bó, Armin Falk, Guillaume Fréchette, Xavier Gabaix, Muriel Niederle, Andy Schotter and Richard Zeckhauser, as well as seminar audiences at Universitat Autònoma de Barcelona (MOVE Workshop), Harvard Kennedy School, NYU Stern, and the 2011 Stanford Institute for Theoretical Economics Summer Workshop for helpful comments. Financial support from the Women and Public Policy Program and the Women's Leadership Board at Harvard Kennedy School is gratefully acknowledged.

†Harvard University, Kennedy School of Government, 79 John F. Kennedy Street, Cambridge, MA 02138. E-mail: iris_bohnet@harvard.edu

‡Stockholm School of Economics, Swedish House of Finance, Drottninggatan 98, SE-111 60 Stockholm, Sweden. E-mail: farzad.saidi@hhs.se

1 Introduction

In labor markets and other competitive contexts, the degree of inequality in outcomes is crucially determined by access to information. Having more information on the determinants of success often serves to reduce uncertainty attached to different economic choices. However, such career-relevant information is not distributed equally, even in relatively small groups. Many factors, including income and technology, play a role. In addition, the number of similar others available to learn from, as determined by country-of-origin or same-language social networks and other demographic characteristics (Bertrand, Luttmer, and Mullainathan (2000); Edin, Fredriksson, and Åslund (2003); Munshi (2003); Jackson and Schneider (2011); Figlio, Hamersma, and Roth (2015)), can serve as a channel of information transmission. Similarly, role models and mentors typically have the same demographic characteristics as their mentees (e.g., Holmes and O'Connell (2007), Ibarra (1992), Ibarra (1993), and Ragins (1999) for a review), making it harder for minorities to learn from comparable others.

Informational differences lead to some people being better informed on how desirable outcomes can be achieved than others. For instance, better-informed individuals might have a clearer notion of how effort translates into rewards. A worse understanding of the effort-reward relationship could rationally depress effort exertion by less-informed individuals. This would, in turn, imply that *absolute informedness* matters for performance. However, it remains unexplored whether this impact on individual performance is complemented, or even superseded, by the effect of *relative informedness*.

In this paper, we present controlled experimental evidence that the distribution of information within groups matters more for individual effort exertion than the absolute level of information provision. This is consistent with a large literature in psychology and economics which suggests that individuals care not only about themselves but also about their position relative to others (Loewenstein, Thompson, and Bazerman (1989); Fehr and Schmidt (1999); Ockenfels and Bolton (2000)). Various psychological mechanisms could be at play. In the realm of stereotype threat (Steele and Aronson (1995); Steele (1998); Steele, James, and Barnett (2002)), similar effects have been found. Namely, self-stereotyping can be particularly pronounced when the stereotype-disadvantaged group has to compete with the stereotype-advantaged group.

In addition, research on procedural fairness suggests that individuals do not only care about outcomes, but also about how outcomes come to be (see Kahneman, Knetsch, and Thaler (1986); Lind and Tyler (1988); and Frey, Benz, and Stutzer (2004) for an overview). That is, individuals may have procedural-fairness concerns, and dislike systems that lead to informational differences, independent of whether they are the beneficiaries of having more information or not.

The design of our laboratory experiment is inspired by labor market reality where some individuals have less precise information on the determinants of economic success than others. Economic outcomes are typically imperfectly correlated with effort – be it due to the subjective assessment of one's performance by superiors or, in the extreme case, luck. Having more information about the non-effort-related determinant of success reduces the uncertainty associated with individuals' prospects. This is all the more important under nonlinear incentive schemes that can have a winner-take-all nature, such as promotion tournaments.

In our experiment, we varied informational precision by changing the number of observations individuals received on the potential values of the random component affecting their performance evaluation, with one group receiving a large sample of data and the worse-informed group receiving a small sample. At the same time, we instructed all subjects that their random values would be drawn from the same distribution. In order to ascertain whether the sample size of the data affected individuals' perception about the effort-reward relationship, we measured their beliefs about the mean and the variance of the random component.

Most importantly, we estimate the impact of such informational differences on effort in a real-effort task where pairs of subjects competed with each other for a prize in a tournament (much like in the classical setting of Lazear and Rosen (1981)). In the baseline condition, both subjects in a competing pair were equally well informed, i.e., both received either a large sample or a small sample of data, allowing them to infer the potential importance of the random component. In the treatment condition, subjects in a competing pair were differentially informed, with one of the subjects receiving a large sample and the other receiving a small sample of data.

If informational disadvantage – i.e., absolute informedness – alone carries the day, then it should hold that the less well individuals are informed on the impact of the stochastic component on economic success, the less effort they will exert, independent of how much information their counterpart has. In contrast, if subjects exhibit *informational inequity aversion*, the informationally disadvantaged will reduce their effort more in the treatment condition where their counterpart is better informed. If subjects dislike informational differences due to procedural-fairness concerns, we might even find that both, the worse and the better informed, cut back on their effort in the treatment condition.

Experimental participants receiving less precise information on the effort-reward relationship performed significantly worse than better-informed participants only in the treatment condition where the competing counterpart was better, rather than equally well, informed. Conversely, there was no performance difference between subjects receiving a small sample of data and those receiving a large sample of data when each of them competed with equally-well-informed counterparts. Thus, members of the worse-informed group were only affected by said informational difference when they competed with better-informed individuals, suggesting that the difference did not affect how well they understood the effort-reward relationship but rather how inequity averse they felt.

Subjects in the better-informed group were not affected by informational differences – they exerted about the same amount of effort in the treatment and in the control condition, hence a general dislike of processes leading to unequal informal allotments did not drive our results. Furthermore, we show that the performance difference in the treatment condition cannot be explained by differences in beliefs about the mean or the variance of the random component.

What mechanism best explains our finding that relative informedness, but not absolute informedness, matters for the performance of seemingly worse-informed subjects? In a rank-order tournament (Lazear and Rosen (1981)), such a performance difference would arise if the contestants faced a de-facto different effort-reward relationship. For example, one type of contestants could be handicapped by a larger variance of the luck component. As long as there is asymmetric information (across these different types of contestants) regarding this difference in the effort-reward relationship, the handicapped group would exert less effort, even under risk neutrality.

This would be in line with the so-called discouragement effect among heterogenous players, resulting in lower individual and aggregate effort in contests, as has been documented experimentally by, for instance, Schotter and Weigelt (1992).[1] Individual effort by disadvantaged, or discouraged, players decreases because they find it relatively unprofitable to exert effort in an attempt to beat their stronger opponents. In addition, advantaged players may react to this by saving on costly effort exertion, compared to a tournament in which they face more symmetric competition. This potential convergence in performance when the heterogeneity of the players is public information is also in line with the model of Lazear and Rosen (1981).

Our experiment shares some similarities with a setup that can give rise to the discouragement effect. In particular, the degree of informedness of all players was public information. However, all subjects were aware that they received their random bonus from the same distribution as their opponents. In this regard, informational inequity aversion is distinct from the discouragement effect documented in the literature as, first, only the disadvantaged subjects reacted by performing worse under asymmetric competition and, second, their receiving a smaller sample of potential bonus values did not matter for the actual effort-reward relationship, which was kept constant within competing pairs. We also show this had no differential effect on subjects' beliefs when they faced asymmetric, rather than symmetric, competition.

Besides introducing informational inequity aversion, our work contributes to earlier experimental studies examining the role of uncertainty and information on individual performance

[1] See Dechenaux, Kovenock, and Sheremeta (2015) for a review of this literature.

in competitive environments. Insofar as we varied information that experimental subjects received on the random component in a tournament, our paper is related to Delfgaauw, Dur, Non, and Verbeke (2015). Furthermore, Bull, Schotter, and Weigelt (1987) and Freeman and Gelber (2010) varied the amount of information on the past performance of competitors that their experimental subjects had available. For a hypothetical effort task, Bull, Schotter, and Weigelt (1987) found that subjects who were informed of their counterparts' decisions after each round exerted less effort than those who did not receive any information. In contrast, Freeman and Gelber (2010) who used a real-effort task (mazes) found that providing more information on the historical performance of competitors led to higher effort on average.

In both cases, the uncertainty about the effort-reward relationship is influenced by the amount of information subjects have available on their competitors' past performance, making best-responding a difficult problem as past performance may not necessarily map directly into future performance, and responses are interdependent.

Connecting with this literature, in our experiment, we varied absolute informedness exogenously and independently of subject performance, allowing us to determine the impact of perceived informedness regarding the importance of luck on effort. By additionally varying the degree of relative informedness, we find that individuals are not just inequity averse in terms of outcomes, but also in terms of information that helps them achieve those outcomes.

The remainder of the paper is organized as follows. In Section 2, we present our experimental design. Section 3 reports the results, and Section 4 concludes.

2 Experimental Design and Procedures

The goal of our experiment is to test whether individuals who may or may not be well informed about the extent to which luck determines outcomes in a tournament perform differently, depending on whether or not they compete with somebody who receives the same amount of information. This enables us to disentangle the effects of relative as opposed to absolute informedness.

Our experiment involved a tournament where pairs of subjects competed with each other for a prize in a word-find task. Subjects were confronted with a matrix containing letters: most letters appeared in random order, but some formed words. Provided with a list of 20 words, subjects could find these words in the matrix by marking sequences of letters horizontally, vertically, or diagonally. An example of such a letter matrix is given in Appendix A. The subjects' task was to mark as many words as possible in three minutes. For every correct word marked, subjects received 10 points. In addition to their task score, subjects received a random bonus, i.e., their luck component.

Bonuses were drawn from a hat containing 18 different bonus values which, in turn, were drawn from a uniform distribution. In every round, *all* subjects were informed of the highest and the lowest possible bonus value, but not that they were uniformly distributed, plus they were given a randomly drawn sample of actual bonus values in the hat (without replacement). The sample was either large (for L-players), namely 12 out of 18 possible bonus values, or small (for S-players), namely 3 out of 18 bonus values.

The final score was equal to the sum of the task score and the value of the bonus. The subject with the higher final score in a pair won the tournament and received a winner prize (10 tokens), whereas the subject with the lower final score received a loser prize (2 tokens). One token was worth $1.

To examine the role of one's competitor's information condition, we created two treatments: the baseline where subjects had perfectly identical information conditions, and a second treatment where they did not. In the *same-sample tournament*, both subjects in a pair received information on either the small or the large sample of bonus values in the hat, and we informed all subjects of this. That is, S-players faced the competition of other S-players, whereas L-players competed with L-players. In the *different-sample tournament*, one person per pair received information on the large sample and the other person in the pair received information on the small sample, and this was common knowledge. That is, S-players competed with L-players.

After an initial practice round, the task was repeated four times (with a different letter matrix and word list in every round). Subjects remained in the same pair for the duration of the experiment. In rounds 1 and 2, subjects were confronted with a wide range of potential bonus values from 0 to 100. In rounds 3 and 4, we decreased the range of bonus values by limiting them to be between 30 and 70.

At the end of each round, subjects were informed of their task score, their final score, their counterpart's final score, and the tokens they won. They did not receive information on their counterpart's task score and were, thus, unable to determine with certainty whether they won/lost because of their counterpart's performance or the randomly drawn bonus. As an example, Appendix B provides the main instructions for S-players in the different-sample tournament.

After subjects had been informed of the highest and the lowest possible bonus values, and had seen their random draw of sample values, we asked them to estimate the mean bonus value and report a 90% confidence interval for their estimate. We did so to elicit their beliefs about the bonus component.

Furthermore, after the completion of the main experiment, we had subjects participate in a risky-choice task to measure their risk preferences, which could also explain differential effort choices.[2] We employed the task introduced by Holt and Laury (2002) with identical incentives. The resulting measures for risk aversion that we use in the empirical analysis are binary variables for the categories from 1 (very risk loving) to 10 (very risk averse).

Finally, the study concluded with a short questionnaire collecting demographic information: in the empirical analysis, we will use binary variables for the categories capturing the subjects' economic background, namely from 1 (very poor) to 6 (very rich), matching the six options given in the questionnaire. We also create a dummy variable for the subject's gender.

Subjects were paid for their performance in all four rounds and, in addition, received their earnings from the risky-choice task. Average earnings, including a $10 show-up fee,

[2] Given the uncertain nature of the tournament, more risk-averse players potentially exert less effort and, thus, score lower.

were about $36 for a study that lasted one hour.

We ran the experiment in the Harvard Decision Science Laboratory. We held nine sessions with up to 24 subjects in each of them, and we yield valid data for 158 subjects.[3]

3 Results

In this section, we empirically examine whether subjects who were provided with a smaller sample of potential bonus values (S-players) performed worse than their counterparts with more information (L-players), and whether this effect depends on whether the subjects were differentially treated (same-sample vs. different-sample tournament). Furthermore, we will explore to what extent any such effect could be explained by the subjects' beliefs about the luck component. Before moving to a regression analysis, we first report descriptive statistics regarding the scores on the word-find task.

Raw Performance Statistics by Treatment

On average, subjects found 10.45 words (with a standard deviation of 3.88) out of a total of 20 words available in a given letter matrix. Women and men differed slightly in their performance, with women marking 10.70 words correctly and men finding 10.06 words on average. However, this difference was entirely driven by performance in the first round, and women and men did not differ (neither economically nor statistically) in their performance in the remaining three rounds. Figure 1 presents the distribution of the number of words people found in the pooled sample. Typical outcomes in the middle 10 to 90% of the distribution ranged from 6 to 16 words per matrix. Four participants – i.e., 2.5% of our subjects in the sample – found the maximum of 20 words in one round (but never more often than that).

We first review raw differences in the mean number of words found by L- and S-players. Table 1 reports the data pooled across both treatments (first panel) and separately for same-

[3] We dropped all scores of a subject after incidents involving IT or other problems during the experiment. In addition, to safeguard consistency, we require subjects to have valid data for all aspects of the experiment.

sample and different-sample tournaments (second and third panels). Within each panel, in the first row we present performance levels aggregated over all four rounds, in the second for the wide-range rounds (rounds 1 and 2), in the third for the narrow-range rounds (rounds 3 and 4), and in the last row for the rounds where people had already gained one round's experience within a given range condition (rounds 2 and 4). Across all four rounds, L-players found 0.84 words more than S-players on average ($p = 0.007$), which corresponds to more than one-fifth of a standard deviation. Interestingly, the difference grows to, and stays roughly at, one word after the first round, as can be seen in Table 2, which presents mean scores by round.[4]

In Table 2, individual scores improved over time, but no clear learning pattern is observable. In particular, scores decreased between rounds 2 and 3 for L- and S-players under both the same-sample and the different-sample tournament, refuting simple learning but suggesting the existence of adjustment costs to the new bonus range in round 3. We do not assign particular importance to this, other than noting that learning alone cannot explain the dynamics we observe. However, performance generally did increase from round 1 to 2 and from round 3 to 4, giving rise to the possibility of learning within each range condition.

Figure 2 presents our main finding. It displays the differences in average task scores between L- and S-players across same-sample and different-sample tournaments. The first two bars correspond to the average scores indicated in the first row of the second and third panel of Table 1, respectively. We find that L-players did not significantly outperform S-players in the same-sample treatment (10.844 vs. 10.408, $P = 0.330$). That is, the variation in absolute informedness by previewing different numbers of potential bonus values did not have any impact on performance under the same-sample treatment.

This is, however, not the case in the different-sample treatment where L-players performed just as well as under the same-sample treatment (10.841), but S-players performed worse (9.576). The difference in scores between the two types of players is significant only in the different-sample treatment ($P = 0.003$). In addition, the last two bars show that the

[4] Note that standard errors become relatively large for the tests by round given the small number of observations.

performance differential in different-sample, but not in same-sample, tournaments holds also when considering only the highest score per individual. This attests to the idea that relative informedness played a role in inducing such performance difference.

In the following, we will buttress our findings in a multivariate-regression framework, and will further scrutinize whether S-players' worse performance under the different-sample treatment could be explained through a rational response to changes in beliefs, rather than through informational inequity aversion.

Regression Analysis

In Table 3, we provide the summary statistics for all variables employed in the regressions, of which we have already discussed the subjects' scores. In addition, we test whether the performance difference between L- and S-players was potentially driven by their beliefs, i.e., whether our experimental treatment had any effect on the subjects' beliefs about the luck component.

For this purpose, we make use of our survey measures on the subjects' estimated mean bonus and their reported confidence intervals around their estimates of the mean bonus. As the range of potential bonus values differed between the first two and the last two rounds, we use a measure of both the estimated mean and the reported confidence interval that is normalized to be between 0 and 1 – namely the estimated mean and the reported confidence interval over the actual range (100 in the first two and 40 in the last two rounds) – to make them comparable across all rounds. This yields *Perceived mean bonus* with a mean of 0.49 and a standard deviation of 0.13, and *Perceived range* with a mean of 0.46 and a standard deviation of 0.25. We use $1 -$ *Perceived range* in our regressions so as to capture the notion of subjects' *precision* of their estimated mean bonus. This can be understood as a measure of the degree of uncertainty subjects perceived, beyond the information on the luck component that was known to all subjects.

In the first column of Table 4, we regress the subjects' per-round scores on an indicator for whether they were L- or S-players (*Large sample*), and – as already seen in the first row of Table 1 – find that L-players, on average, found 0.84 words more than S-players. In the second column, this insight holds up to including round-specific dummy variables, a gender dummy, as well as controls for our survey-based measures for the subjects' economic background and degree of risk aversion. We include separate indicators for rounds 2 and 4, and the last two (narrow-range) rounds in an attempt to disentangle any effect of reducing the range of potential bonus values from within-range experience. For the latter, one would hypothesize that subjects perform better in rounds 2 and 4 than in rounds 1 and 3. This was indeed the case: our subjects, on average, found 1.2 words more in rounds 2 and 4. We also find some suggestive evidence that narrow-range rounds were associated with higher performance, although the effect is robustly significant only for the fourth round (for which one needs to add up the coefficients on the two above-mentioned indicator variables).

In the third column, we present our main finding of this paper, and explore whether there is any differential impact of the different-sample vs. same-sample treatment on the performance difference between L- and S-players. The corresponding difference-in-differences estimate is the coefficient on the interaction between *Large sample* and *Diff. sample*, which is significant at the 3% level. L-players did not outperform S-players in the same-sample tournament, as captured by the insignificant coefficient on *Large sample*. However, S-players performed significantly worse – namely, they found 1.3 fewer words – in the different-sample than in the same-sample tournament, as captured by the coefficient on *Diff. sample*. This effect corresponds to one-third of a standard deviation in task scores.

On the other hand, the performance of L-players was roughly the same across both the same-sample and the different-sample treatment, as the sum of the coefficients on *Diff. sample* and *Large sample* $\times$ *Diff. sample* is not significantly different from zero ($P = 0.472$).

The subjects learned about their scores and how much they earned after each round, and they were paid for their performance in all four rounds. This raises the question as to whether depending on their accumulation of wealth over these rounds, they may have exerted different amounts of effort. In the fourth column, we control for this possibility by including the number of times a subject won prior to round n (*Times won before*), i.e., at most $n-1$ times. Note that this variable also implicitly captures the past realizations of the luck component. The coefficients on *Large sample*, *Diff. sample*, and their interaction are virtually unaltered. On the other hand, the coefficients on *Rounds 2 & 4* and *Rounds 3 & 4* are somewhat reduced, as they relate to the subjects' experience during the run-time of the tournament, as does *Times won before*.

In the fifth column, we control for subjects' self-reported beliefs regarding the mean and the range of the bonus component, *Perceived mean bonus* and *Perceived range*, but fail to find any effect. Most importantly, the coefficients on *Large sample*, *Diff. sample*, and their interaction are again unaffected. In addition, in the last column, we show that the effects on performance are robust to our exogenously imposed variation in the range of potential bonus values between the first two (wide-range) and last two (narrow-range) rounds.

These results suggest that L-players did not outperform S-players due to the actual amount of additional information because otherwise, the performance differential should have been apparent in the same-sample tournament. Instead, S-players performed significantly worse and, subsequently, found fewer words when they faced a seemingly better informed counterpart. In addition, any decrease in S-players' effort is unlikely to be a response to the suspicion that the game was in any way rigged against them, as all subjects were informed that they received their random bonus from the *same* hat/distribution as their counterparts (cf. instructions for "Stage 2" in Appendix B).

Finally, we scrutinize whether S-players' worse performance could be explained through a rational response to changes in beliefs, rather than through informational inequity aversion. If the performance difference between L- and S-players pertaining to the different-sample, rather than the same-sample, treatment was driven by beliefs about the bonus component, then we should be able to estimate a significant impact of the interaction between *Large sample* and *Diff. sample* on the subjects' beliefs.

In Table 5, we re-run the regressions from (all but the fifth column of) Table 4, and use as dependent variable the subjects' estimated mean bonus. While L-players perceived the mean of their bonus to be significantly higher, this effect was invariant to the same-sample and different-sample treatments. The magnitude of this effect is very large, and amounts to over one-half of a standard deviation (cf. third and fourth columns). However, as can be seen in the last column, the effect persisted only throughout the first two rounds. Most importantly, we do not find any significant impact of the interaction between *Large sample* and *Diff. sample* which could explain our previous findings for the subjects' task performance.

We next use as dependent variable the confidence interval around the subjects' estimate of the mean bonus. One can interpret this measure as a proxy for the perceived variance of the bonus component or – more conservatively – for the perceived uncertainty about the functional form of the distribution, beyond the information on the bonus component that was known to all subjects. As mentioned before, we use $1 - \textit{Perceived range}$ as a measure of subjects' precision of their estimated mean bonus. The results are in Table 6. Throughout all columns, receiving a larger sample of information implies attaching a tighter confidence interval around the estimated mean value of the bonus. The effect of receiving a larger sample on $1 - \textit{Perceived range}$ is equal to slightly less than one-half of a standard deviation (cf. third and fourth columns).[5]

However, as in Table 5, we fail to find any effect for our difference-in-differences estimate, namely the coefficient on *Large sample* $\times$ *Diff. sample*. Therefore, while we do find that receiving a larger sample mattered for beliefs, this did not induce a significant performance gap between L- and S-players in the same-sample tournament. Most importantly, the difference between L- and S-players' self-reported beliefs did not vary between the same-sample and the different-sample tournaments and, thus, cannot explain our previous findings for the subjects' task performance.

[5] The negative effect of the narrow-range rounds 3 and 4 on $1 - \textit{Perceived range}$ might be confusing. Subjects, unsurprisingly, reported significantly smaller *absolute* confidence intervals in rounds 3 and 4 as opposed to rounds 1 and 2. However, the reported *relative* confidence intervals reverse, as we divide confidence intervals by the actual round-specific ranges of the bonus component to make the measures comparable. We cannot use absolute confidence intervals as our dependent variable, although the respective coefficient in unreported regressions suggests a significantly smaller reported absolute confidence interval in rounds 3 and 4, as that correlation is partly mechanical. Still, the correlation between the reported relative and absolute confidence intervals is 0.68. We therefore use the relative confidence interval, *Perceived range*, in all regressions to avoid mechanical correlations. Also, our results are virtually unchanged if one drops the indicator for rounds 3 and 4 from the right-hand side of the regression specifications.

4 Conclusion

When and how do informational differences translate into performance differences? Does the distribution of information matter more for individual effort exertion than the absolute level of information provision? To shed light on these issues, we presented the results of a real-effort experiment in which we varied the amount of information that subjects had on the role of luck for tournament outcomes. In addition, we matched subjects with equally or differentially informed counterparts. This enabled us to disentangle the effects of *relative* as opposed to *absolute* informedness.

We found that differences in absolute informedness were not associated with performance differences. Instead, informational differences had an effect on performance only when subjects were differentially treated, and the seemingly less well informed subjects performed significantly worse.

This suggests that even a subtle perturbation in individuals' perception of how well informed they are compared to others, even when their outcomes are equally affected by the same source of uncertainty, can depress the performance of seemingly worse-informed individuals. This mechanism is potentially generalizable to settings in which peers compete for a prize or a promotion, and face noisy performance evaluation. Informational differences may then arise regarding different parameters of this performance evaluation, some of which may carry greater weight for actual outcomes than others.

The resulting pattern of behavior also corresponds more generally with other theories and field observations: the activation of self-stereotyping (Steele and Aronson (1995); Steele (1998); Steele, James, and Barnett (2002)) when the stereotype-disadvantaged group has to compete with the stereotype-advantaged group, perceived informational injustice leading to a pessimistic outlook (Bénabou and Tirole (2006)), and women shying away from competition only in mixed-sex, rather than same-sex, groups (Gneezy, Niederle, and Rustichini (2003); Niederle and Vesterlund (2007); Booth and Nolen (2012)).

Our findings can serve as a first basis for arguing that informational inequity aversion could explain how initial relative information conditions, often determined by social networks and other group structures, influence future labor-market outcomes. Testing the potential generality of this insight gives rise to fruitful avenues for future research that explores the role of informational inequity aversion in different settings and for different groups.

References

Bénabou, R., and J. Tirole (2006): "Belief in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, 121(2), 699–746.

Bertrand, M., E. F. P. Luttmer, and S. Mullainathan (2000): "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, 115(3), 1019–1055.

Booth, A. L., and P. Nolen (2012): "Gender Differences in Risk Behaviour: Does Nurture Matter?," *Economic Journal*, 122(558), F56–F78.

Bull, C., A. Schotter, and K. Weigelt (1987): "Tournaments and Piece Rates: An Experimental Study," *Journal of Political Economy*, 95(1), 1–33.

Dechenaux, E., D. Kovenock, and R. M. Sheremeta (2015): "A Survey of Experimental Research on Contests, All-pay Auctions and Tournaments," *Experimental Economics*, 18(4), 609–669.

Delfgaauw, J., R. Dur, A. Non, and W. Verbeke (2015): "The Effects of Prize Spread and Noise in Elimination Tournaments: A Natural Field Experiment," *Journal of Labor Economics*, 33(3), 521–569.

Edin, P.-A., P. Fredriksson, and O. Åslund (2003): "Ethnic Enclaves and the Economic Success of Immigrants – Evidence from a Natural Experiment," *Quarterly Journal of Economics*, 118(1), 329–357.

Fehr, E., and K. M. Schmidt (1999): "A Theory Of Fairness, Competition, And Cooperation," *Quarterly Journal of Economics*, 114(3), 817–868.

Figlio, D. N., S. Hamersma, and J. Roth (2015): "Information Shocks and the Take-Up of Social Programs," *Journal of Policy Analysis and Management*, 34(4), 781–804.

Freeman, R. B., and A. M. Gelber (2010): "Prize Structure and Information in Tournaments: Experimental Evidence," *American Economic Journal: Applied Economics*, 2(1), 149–164.

Frey, B. S., M. Benz, and A. Stutzer (2004): "Introducing Procedural Utility: Not Only What, but Also How Matters," *Journal of Institutional and Theoretical Economics*, 160(3), 377–401.

Gneezy, U., M. Niederle, and A. Rustichini (2003): "Performance In Competitive Environments: Gender Differences," *Quarterly Journal of Economics*, 118(3), 1049–1074.

Holmes, M., and S. O'Connell (2007): "Leaks in the Pipeline: Why Do Women Remain Curiously Absent from the Ranks of Academia?," *Nature*, 446(7133), 346.

Holt, C. A., and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), 1644–1655.

Ibarra, H. (1992): "Homophily and Differential Returns: Sex Differences in Network Structure and Access in an Advertising Firm," *Administrative Science Quarterly*, 37(3), 422–447.

——— (1993): "Personal Networks of Women and Minorities in Management: A Conceptual Framework," *Academy of Management Review*, 18(1), 56–87.

Jackson, C. K., and H. S. Schneider (2011): "Do Social Connections Reduce Moral Hazard? Evidence from the New York City Taxi Industry," *American Economic Journal: Applied Economics*, 3(3), 244–267.

Kahneman, D., J. Knetsch, and R. Thaler (1986): "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, 76(4), 728–741.

Lazear, E. P., and S. Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89(5), 841–864.

Lind, E. A., and T. R. Tyler (1988): *The Social Psychology of Procedural Justice.* Plenum Press, New York, NY.

Loewenstein, G. F., L. Thompson, and M. H. Bazerman (1989): "Social Utility and Decision Making in Interpersonal Contexts," *Journal of Personality and Social Psychology*, 57(3), 426–441.

Munshi, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics*, 118(2), 549–599.

Niederle, M., and L. Vesterlund (2007): "Do Women Shy Away from Competition? Do Men Compete Too Much?," *Quarterly Journal of Economics*, 122(3), 1067–1101.

Ockenfels, A., and G. E. Bolton (2000): "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 90(1), 166–193.

Ragins, B. R. (1999): "Gender and Mentoring Relationships: A Review and Research Agenda for the Next Decade," in *Handbook of Gender & Work*, ed. by G. N. Powell, pp. 347–370. SAGE Publications, Inc.

Schotter, A., and K. Weigelt (1992): "Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results," *Quarterly Journal of Economics*, 107(2), 511–539.

Steele, C. M. (1998): "Stereotyping and its Threat are Real," *American Psychologist*, 53(6), 680–681.

Steele, C. M., and J. Aronson (1995): "Stereotype Threat and the Intellectual Test Performance of African-Americans," *Journal of Personality and Social Psychology*, 69(5), 797–811.

Steele, J., J. James, and R. Barnett (2002): "Learning in a Man's World: Examining the Perceptions of Undergraduate Women in Male-Dominated Academic Areas," *Psychology of Women Quarterly*, 26(1), 46–50.

5 Figures

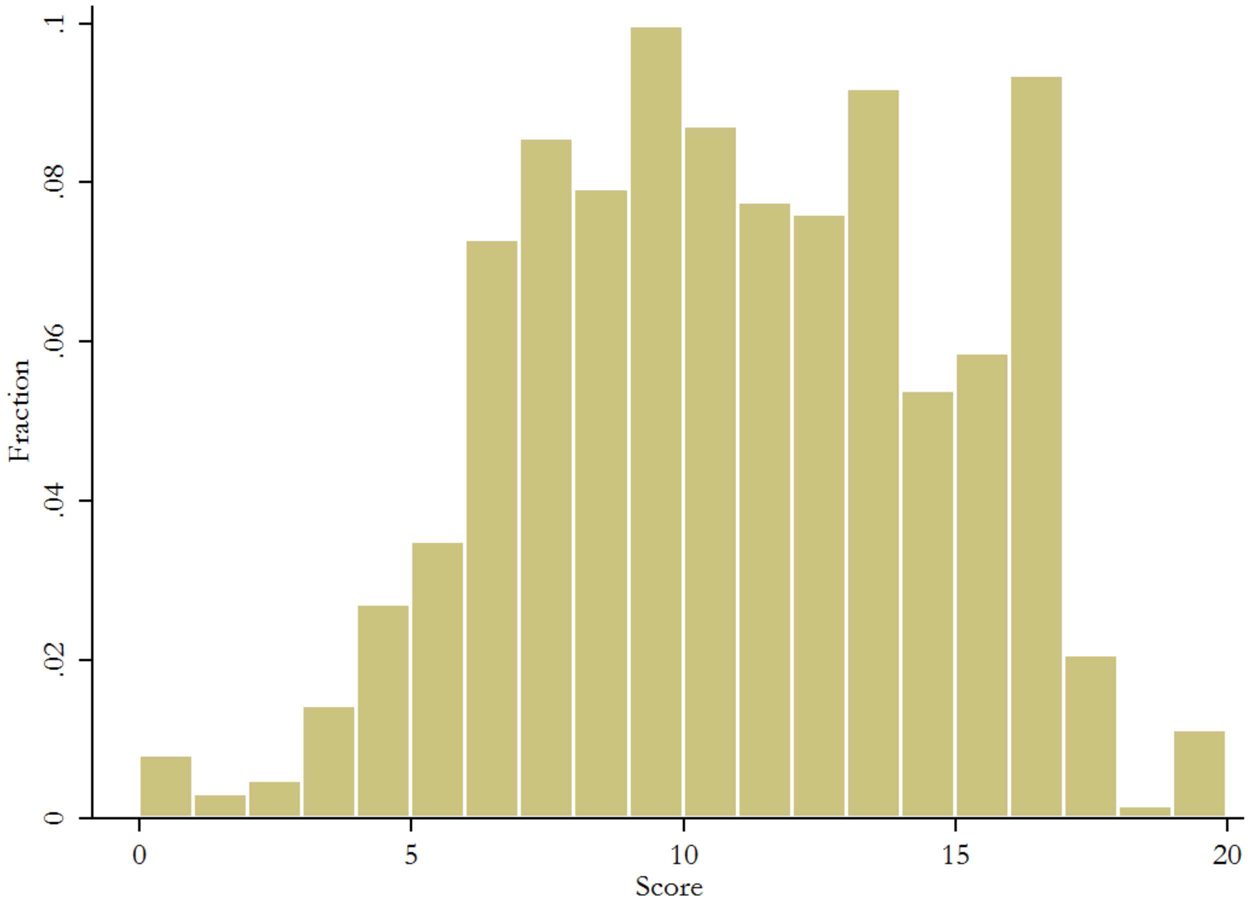

Figure 1: **Histogram of Scores.** This figure presents the distribution of the number of words people found in the pooled sample.

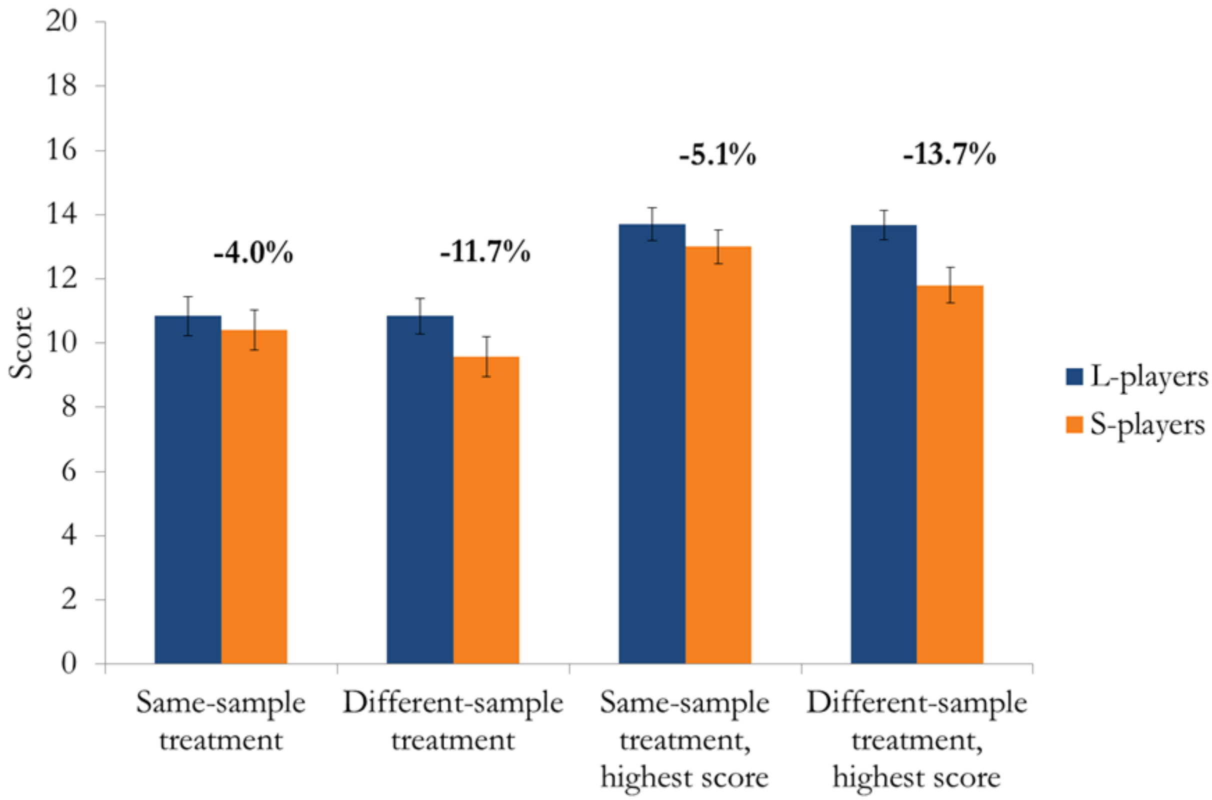

Figure 2: **Differences in Mean Scores between L- and S-players across Treatments.** The first two bars build on all available scores for L- and S-players, whereas the last two bars use only the highest achieved score (across all rounds) for each individual. Error bars indicate 95% confidence intervals.

6 Tables

Table 1: **Differences in Mean Scores**

Treatment	Rounds	Large sample ($N = 84$)	Small sample ($N = 74$)	Difference (p-value)
Both	All	10.842	10.003	0.007
		(3.844)	(3.885)	
Both	1 & 2	10.411	9.716	0.089
		(3.542)	(3.683)	
Both	3 & 4	11.274	10.291	0.033
		(4.090)	(4.070)	
Both	2 & 4	11.542	10.500	0.019
		(3.929)	(3.904)	
Treatment	Rounds	Large sample ($N = 40$)	Small sample ($N = 38$)	Difference
Same sample	All	10.844	10.408	0.330
		(3.937)	(3.956)	
Same sample	1 & 2	10.438	10.145	0.626
		(3.635)	(3.843)	
Same sample	3 & 4	11.250	10.671	0.384
		(4.202)	(4.073)	
Same sample	2 & 4	11.475	10.934	0.394
		(3.978)	(3.920)	
Treatment	Rounds	Large sample ($N = 44$)	Small sample ($N = 36$)	Difference
Different sample	All	10.841	9.576	0.003
		(3.769)	(3.776)	
Different sample	1 & 2	10.386	9.264	0.044
		(3.475)	(3.476)	
Different sample	3 & 4	11.295	9.889	0.030
		(4.009)	(4.054)	
Different sample	2 & 4	11.602	10.042	0.013
		(3.906)	(3.862)	

Notes: N denotes the number of subjects. Standard deviations are in parentheses. The last column indicates the P-values from two-sided difference-in-means tests.

Table 2: **Differences in Mean Scores by Rounds**

Treatment	Round	Large sample ($N = 84$)	Small sample ($N = 74$)	Difference (p-value)
Both	1	9.926	9.554	0.511
		(3.456)	(3.691)	
Both	2	10.893	9.878	0.082
		(3.580)	(3.694)	
Both	3	10.357	9.459	0.149
		(3.817)	(3.959)	
Both	4	12.190	11.122	0.105
		(4.170)	(4.034)	
Treatment	Round	Large sample ($N = 40$)	Small sample ($N = 38$)	Difference
Same sample	1	9.900	10.132	0.790
		(3.699)	(3.967)	
Same sample	2	10.975	10.158	0.326
		(3.534)	(3.767)	
Same sample	3	10.525	9.632	0.322
		(3.955)	(3.962)	
Same sample	4	11.975	11.711	0.781
		(3.965)	(4.3647)	
Treatment	Round	Large sample ($N = 44$)	Small sample ($N = 36$)	Difference
Different sample	1	9.955	8.944	0.176
		(3.263)	(3.320)	
Different sample	2	10.818	9.583	0.137
		(3.662)	(3.644)	
Different sample	3	10.205	9.278	0.288
		(3.727)	(4.004)	
Different sample	4	12.386	10.500	0.041
		(4.024)	(4.067)	

Notes: N denotes the number of subjects. Standard deviations are in parentheses. The last column indicates the P-values from two-sided difference-in-means tests.

Table 3: **Summary Statistics**

Pooled	Mean	Std. dev.	Min	Max	N
Task score	10.449	3.883	0	20	632
Perceived mean bonus	0.485	0.133	0	1	632
Perceived range	0.463	0.252	0	1	632
Individual level	Mean	Std. dev.	Min	Max	N
Large sample	0.532	0.501	0	1	158
Different sample	0.506	0.502	0	1	158
Female	0.614	0.488	0	1	158
Economic background	3.715	0.972	1	6	158
Risk aversion	6.608	1.588	3	10	158

Notes: N corresponds to the number of observations in the first panel, and the number of subjects, each of which played four rounds, in the second panel. *Task score* is the subject's score, between 0 and 20, on the word-find task in a given round. *Perceived mean bonus* denotes the subject's reported estimate of the mean bonus. *Perceived range* is the subject's reported confidence interval of her estimate of the mean bonus. The latter two measures are normalized to be between 0 and 1, i.e., the estimated mean and the reported confidence interval over the actual range (100 in the first two and 40 in the last two rounds). We use 1 − *Perceived range* in our regressions so as to capture the notion of subjects' *precision* of their estimated mean bonus. *Large sample* indicates that the subject is an L-player, rather than an S-player. The different-sample tournament, under which the two players received samples of different size, is labeled as *Diff. sample*. Self-reported economic background is scaled from 1 (very poor) to 6 (very rich), and risk aversion is measured on a scale from 1 (very risk loving) to 10 (very risk averse).

Table 4: **Determinants of Task Performance**

	Task score $\in [0, 20]$					
Large sample	0.839**	0.915**	0.119	0.101	-0.001	-0.024
	(0.392)	(0.412)	(0.556)	(0.554)	(0.574)	(0.717)
Diff. sample			-1.264**	-1.263**	-1.289**	-1.314*
			(0.529)	(0.521)	(0.534)	(0.681)
Large sample × Diff. sample			1.658**	1.613**	1.699**	1.659*
			(0.744)	(0.738)	(0.750)	(0.985)
Rounds 2 & 4		1.209***	1.209***	1.016***	1.263***	1.209***
		(0.216)	(0.216)	(0.260)	(0.232)	(0.217)
Rounds 3 & 4		0.728*	0.728*	0.336	0.794**	0.526
		(0.384)	(0.381)	(0.455)	(0.400)	(0.751)
Female		0.602	0.524	0.572	0.537	0.524
		(0.402)	(0.394)	(0.392)	(0.398)	(0.395)
Times won before				0.374		
				(0.233)		
Perceived mean bonus					0.620	
					(1.259)	
Perceived range					-0.611	
					(0.709)	
Large sample × Rounds 3 & 4						0.286
						(1.109)
Diff. sample × Rounds 3 & 4						0.099
						(1.065)
Large × Diff. sample						-0.002
× Rounds 3 & 4						(1.508)
Constant	10.003***					
	(0.282)					
Economic-background FE	N	Y	Y	Y	Y	Y
Risk-aversion FE	N	Y	Y	Y	Y	Y
N	632	632	632	632	632	632

Notes: The dependent variable is the subject's score, between 0 and 20, on the word-find task in a given round. *Large sample* indicates that the subject is an L-player, rather than an S-player. The different-sample tournament, under which the two players received samples of different size, is labeled as *Diff. sample*. *Rounds 2 & 4* and *Rounds 3 & 4* are indicator variables for the respective rounds. *Female* indicates the subject's gender, *Times won before* is the number of times (always 0 for the first round) the subject has won before the round in question, *Perceived mean bonus* is the subject's reported estimate of the mean bonus, normalized to be between 0 and 1, and *Perceived range* is the subject's reported confidence interval of her estimate of the mean bonus, normalized to be between 0 and 1. Self-reported economic background is scaled from 1 (very poor) to 6 (very rich), and included in the regressions as separate dummy variables. Risk aversion is measured on a scale from 1 (very risk loving) to 10 (very risk averse), and also included in the regressions as separate dummy variables. Robust standard errors (clustered at the level of teams by set of rounds (first two vs. last two)) are in parentheses.

Table 5: **Determinants of Perceived Mean of Bonus Component**

	Perceived mean bonus $\in [0,1]$				
Large sample	0.058***	0.062***	0.078***	0.078***	0.133***
	(0.011)	(0.011)	(0.015)	(0.015)	(0.016)
Diff. sample			0.017	0.017	-0.002
			(0.018)	(0.018)	(0.019)
Large sample × Diff. sample			-0.033	-0.033	-0.023
			(0.022)	(0.022)	(0.023)
Rounds 2 & 4		-0.068***	-0.068***	-0.069***	-0.068***
		(0.009)	(0.009)	(0.010)	(0.009)
Rounds 3 & 4		0.019*	0.019*	0.017	0.064***
		(0.010)	(0.010)	(0.014)	(0.022)
Female		0.010	0.011	0.011	0.011
		(0.009)	(0.009)	(0.009)	(0.009)
Times won before				0.003	
				(0.008)	
Large sample × Rounds 3 & 4					-0.109***
					(0.028)
Diff. sample × Rounds 3 & 4					0.039
					(0.031)
Large × Diff. sample × Rounds 3 & 4					-0.021
					(0.037)
Constant	0.454***				
	(0.009)				
Economic-background FE	N	Y	Y	Y	Y
Risk-aversion FE	N	Y	Y	Y	Y
N	632	632	632	632	632

Notes: The dependent variable is the subject's reported estimate of the mean bonus, normalized to be between 0 and 1. *Large sample* indicates that the subject is an L-player, rather than an S-player. The different-sample tournament, under which the two players received samples of different size, is labeled as *Diff. sample*. *Rounds 2 & 4* and *Rounds 3 & 4* are indicator variables for the respective rounds. *Female* indicates the subject's gender, and *Times won before* is the number of times (always 0 for the first round) the subject has won before the round in question. Self-reported economic background is scaled from 1 (very poor) to 6 (very rich), and included in the regressions as separate dummy variables. Risk aversion is measured on a scale from 1 (very risk loving) to 10 (very risk averse), and also included in the regressions as separate dummy variables. Robust standard errors (clustered at the level of teams by set of rounds (first two vs. last two)) are in parentheses.

Table 6: **Determinants of Perceived Range of Bonus Component**

	1 − Perceived range $\in [0,1]$				
Large sample	0.102***	0.101***	0.117***	0.118***	0.131***
	(0.027)	(0.027)	(0.038)	(0.038)	(0.048)
Diff. sample			0.022	0.022	0.044
			(0.034)	(0.034)	(0.050)
Large sample × Diff. sample			-0.034	-0.032	-0.096
			(0.055)	(0.055)	(0.073)
Rounds 2 & 4		-0.019**	-0.019**	-0.013	-0.019**
		(0.009)	(0.009)	(0.012)	(0.009)
Rounds 3 & 4		-0.128***	-0.128***	-0.116***	-0.126**
		(0.024)	(0.024)	(0.028)	(0.051)
Female		-0.034	-0.032	-0.034	-0.032
		(0.025)	(0.026)	(0.026)	(0.026)
Times won before				-0.012	
				(0.016)	
Large sample × Rounds 3 & 4					-0.027
					(0.067)
Diff. sample × Rounds 3 & 4					-0.044
					(0.069)
Large × Diff. sample × Rounds 3 & 4					0.124
					(0.098)
Constant	0.483***				
	(0.020)				
Economic-background FE	N	Y	Y	Y	Y
Risk-aversion FE	N	Y	Y	Y	Y
N	632	632	632	632	632

Notes: The dependent variable is 1 minus *Perceived range*, which is the subject's reported confidence interval of her estimate of the mean bonus, normalized to be between 0 and 1. *Large sample* indicates that the subject is an L-player, rather than an S-player. The different-sample tournament, under which the two players received samples of different size, is labeled as *Diff. sample*. *Rounds 2 & 4* and *Rounds 3 & 4* are indicator variables for the respective rounds. *Female* indicates the subject's gender, and *Times won before* is the number of times (always 0 for the first round) the subject has won before the round in question. Self-reported economic background is scaled from 1 (very poor) to 6 (very rich), and included in the regressions as separate dummy variables. Risk aversion is measured on a scale from 1 (very risk loving) to 10 (very risk averse), and also included in the regressions as separate dummy variables. Robust standard errors (clustered at the level of teams by set of rounds (first two vs. last two)) are in parentheses.

Supplementary Appendix

A Letter Matrix (Example: Nations of the World)

B	T	U	W	T	T	B	P	M	S	K	L	L	L	T	W	Q	N	B	V	ALGERIA
O	O	M	E	X	J	A	E	B	K	I	D	A	M	A	R	B	H	W	Y	BELGIUM
F	F	H	N	G	N	G	Y	O	M	U	H	J	I	U	T	K	U	B	W	CANADA
W	C	P	I	A	Y	L	S	T	A	K	C	S	A	R	I	V	Y	B	P	EGYPT
S	Z	E	M	U	T	P	E	V	I	R	T	K	Q	P	E	G	I	M	X	FINLAND
R	A	A	A	L	S	B	T	U	U	A	Z	Q	D	A	A	G	L	A	Q	GREECE
D	N	A	L	A	E	Z	W	E	N	I	X	W	D	V	O	N	L	E	F	HONDURAS
S	B	V	V	T	W	U	A	I	R	N	X	N	J	U	W	R	I	A	B	INDONESIA
N	G	Y	K	H	Y	D	F	T	B	E	A	R	C	B	V	V	X	G	H	JAPAN
S	F	T	V	A	R	Y	B	R	L	W	Y	A	O	I	I	N	T	M	Y	KOREA
A	G	D	D	I	M	D	G	U	R	A	U	S	E	D	E	C	J	A	U	LATVIA
R	U	W	P	L	B	G	Z	X	G	Q	M	T	F	M	A	J	I	X	W	MALTA
U	D	B	N	A	C	O	T	S	Z	U	N	P	E	N	U	S	D	F	H	NEW ZEALAND
D	H	N	X	N	O	I	D	B	H	A	R	Y	A	P	E	K	I	I	U	PANAMA
N	O	C	P	D	O	M	X	V	M	G	P	D	G	N	O	N	V	R	K	RWANDA
O	Q	L	L	J	M	H	I	B	R	A	A	H	O	R	L	Q	T	T	R	SINGAPORE
H	F	B	V	G	R	E	E	C	E	M	V	D	E	A	J	Q	B	L	P	THAILAND
A	B	I	T	L	K	P	H	C	W	B	N	A	N	P	J	P	G	U	O	UKRAINE
C	X	Q	U	G	I	Z	F	V	J	I	C	D	O	Z	K	N	T	I	V	VIETNAM
B	W	Y	L	D	I	S	I	N	G	A	P	O	R	E	V	I	D	K	M	YEMEN

B Experimental Instructions

You are participating in a study in which you will earn some money. The amount will depend on how well you do in a task plus a bonus (described below). At the end of the study, your earnings (1 token = $1) will be added to a show-up fee, and you will be paid in cash.

Main task We will show you matrices containing letters. Some letters appear in random order and some form words. Words can be found by combining letters next to each other horizontally (moving from left to right or from right to left), vertically (moving from top to bottom or from bottom to top), or diagonally (moving from left to right or from right to left). A list of all words contained in a given matrix is displayed next to each matrix. You will be shown a matrix for 3 minutes. You are asked to find as many words from the list as possible. Your point score for the task is calculated as follows:

- For every correct word marked in the matrix, 10 points are added to your score.
- Words that are not marked receive no points.

You are randomly matched with another person present. You and your counterpart see the same letter matrix, and you are both asked to find as many words as possible.

Your *final point score* depends on your point score from the task plus a bonus corresponding to a number between X and Y (you will be informed of the values of X and Y in each round). The number will be randomly drawn from a hat containing 18 balls.

In order for you to get a better sense of the likely value of the bonus, we will first draw 3 balls from the hat at random and inform you of their values. We will then put the balls back into the hat.

In order for your counterpart to get a better sense of the value of the bonus, we will draw 12 balls from the same hat at random and inform your counterpart of their values.

Your *final point score* equals: TASK POINT SCORE + BONUS NUMBER

For example (choosing numbers outside of the range of possibilities in this study), if you identified 1,000 correct words and your counterpart identified 900 correct words, the value of your bonus (randomly drawn number) was 2,000 and the value of your counterpart's bonus (randomly drawn number) was 1,500, then your final point score would be: 10,000 + 2,000 = 12,000. Your counterpart's final point score would be: 9,000 + 1,500 = 10,500.

Calculation of payout The person getting the higher final point score in your pair will receive 10 tokens. The person with the lower final point score will receive 2 tokens.

How the study is conducted It is conducted in five stages.

Stage 1 You will be informed of the lowest and the highest possible bonus (X and Y) contained in a given hat and the value of the balls randomly drawn from the hat. We will ask you to give us your best guess of the average value of all balls in the hat.

Stage 2 We will draw a ball at random for you and then put the ball back into the hat. This ball is your bonus, i.e., the value of the ball counts towards your final score. From the same hat, we will also draw a bonus at random for your counterpart and then return the ball into the hat. You will be informed of your bonus after completing the word-find task.

Stage 3 You will complete the word-find task.

Stage 4 We will then calculate your scores and inform you of your final score (i.e., your score for the task plus the value of your bonus) and your counterpart's final score.

Stage 5 The person getting the higher final score in your pair will receive 10 tokens, the other person will receive 2 tokens.

The study is conducted anonymously and without communication between you and your counterpart. Participants will be identified only by code numbers.

The exercise is *repeated four times*.

- You will remain in the same pair for all four rounds.
- We will always draw 3 balls for you and 12 balls for your counterpart.
- The composition of the hat, that is the average value of the bonus (all balls), will be different in each round. At the beginning of each round, you will be informed of the highest and the lowest possible value of the bonus (X and Y).
- The prize for the winner, 10 tokens, and the prize for the other person, 2 tokens, will remain the same in each round. At the end of each round, we will determine the winner and inform you whether you won 10 or 2 tokens.

Specific instructions for how to mark the words Once we start, you will see a letter matrix on your screen. You can highlight the words you find by marking them with your mouse. Your task is to mark as many correct words as possible. We will practice this in a trial round.

If you have any questions, please press the help button now. Once we have addressed all questions, we will start with the trial round. You will have another chance to ask questions after you have familiarized yourself with the methodology in the trial round.