When "hope springs eternal:" The role of chance in risk taking[*]

Natalia Karelaia[♣#] & Robin M. Hogarth[♦]

[♣] INSEAD, Fontainebleau, France

[♦]ICREA & Universitat Pompeu Fabra, Barcelona, Spain

natalia.karelaia@insead.edu

robin.hogarth@upf.edu

June 11, 2009

[*] The authors appreciate the insightful comments on earlier versions of this paper by Daniel Read and Don Moore as well as comments by seminar participants at the London Business School, the University of Trento, and Universitat Pompeu Fabra. They also thank Roger Rodrigo and Augusto Rupérez-Micola for help in conducting the experiments. This research was financed partially by grants from the Swiss National Science Foundation and the INSEAD Alumni Fund (Karelaia) and the Spanish Ministerio de Educación y Ciencia (Hogarth).

[#] Corresponding author: Natalia Karelaia, INSEAD, Boulevard de Constance, 77305 Fontainebleau, France, phone: + 33 1 6072 4511, fax: + 33 1 60 74 55 00.

Abstract

In most naturally occurring situations, success depends on both skill and chance. We contrast experimental market entry decisions where payoffs depend on skill as opposed to combinations of skill and chance. Our data show differential attitudes toward chance by those whose self-assessed skills are low and high. Making chance more important induces greater optimism for the former who start taking more risk, while the latter maintain a belief that high levels of skill are sufficient to overcome the vagaries of chance. Finally, although we observed "excess entry" (i.e., too many participants entered markets), this could not be attributed to overconfidence.

Keywords: Skill, chance, overconfidence, optimism, competition, risk taking, gender differences.

JEL classification: C91, D03, D81.

It is common for people to compete for access to limited resources. Consider, for example, applications by researchers for grants, attempts to obtain positions in large organizations or public office, and decisions by potential entrepreneurs to enter new markets. These situations typically share four characteristics: (1) a limited number of resources can be given to only k of N agents where $k < N$; (2) each agent has to decide independently whether or not to enter the competition; (3) agents who enter the competition are ranked according to a criterion and the resources are allocated to those with the k highest ranks. The unsuccessful entrants incur additional costs; (4) prior to deciding, each agent receives a signal that is probabilistically related to his or her score on the ranking criterion.

Perhaps the best known example of this paradigm in the economic literature is the work by Camerer and Lovallo (1999) who used it to model entrepreneurial entry. Their focus was on possible effects of overconfidence which they tested by comparing responses (i.e., entry decisions) between conditions where entrants were ranked at random as opposed to their knowledge measured by a test. Camerer and Lovallo noted higher entry rates when ranks were established on the basis of tests as opposed to random orderings, a result they interpreted as indicating overconfidence.[1] Using a similar experimental design, Moore and Cain (2007) varied test difficulty and found higher entry rates for easy as opposed to hard tests.

The experimental market entry paradigm has proven to be a rich stimulus for research. And yet, we contend here that the work reported to date has been limited by failing to illuminate the complex nature of the interaction between skill and chance in economic decision making. In particular, the paradigm to date can be characterized by two extreme positions. In one, behavior has been investigated under purely random conditions where no skill is involved (as in Camerer and Lovallo's random ranking condition). That is, skill plays no role in the outcomes obtained by agents. In contrast, in the other it is relative skill that

[1] Parenthetically, whether Camerer and Lovallo's (1999) results really demonstrate overconfidence can be questioned (Benoît & Dubra, 2009; Hogarth & Karelaia, 2009). However, this is not the goal of the present paper.

determines outcomes and the only uncertainties faced by agents concern the validity of the signal they receive about their skill and the actions of other agents.

The outcomes of naturally occurring markets, however, are more accurately characterized by mixtures of chance and skill. A potential entrepreneur, for example, should not only be concerned about the level of her skills relative to competitors but also the other uncertainties inherent in economic conditions. What will be the demand for the new product? Will there be a general economic downturn? Will there be unanticipated surprises, e.g., a presently unknown competitor introduces a technologically superior product? And so on. Clearly, the potential effects of such random factors are important and should be modeled in the experimental paradigm.

Our goal therefore in this paper is to expand the market entry paradigm by investigating what happens when both chance and skill play important roles in determining outcomes. Specifically, what happens when, in addition to both the probabilistic nature of the signal-criterion relation and uncertainty concerning the actions of others, agents learn that the rankings have been perturbed by an explicit random factor? A priori, one might imagine that increasing uncertainty would discourage agents from entering risky markets. However, could it be that, when skill is involved, people prefer noisier evaluation procedures? In addition, how do attitudes toward noise interact with different levels of skill?

More generally, our research follows a long research tradition that has investigated risky choice in situations where outcomes depend on both skill and chance. Starting with the innovative work of Cohen and his colleagues who investigated bus drivers' attempts to make tight maneuvers (Cohen, Dearnaley, & Hansel, 1956), soccer players' shots at goal (Cohen & Dearnaley, 1962), and success at hitting dart boards (Cohen & Hansel, 1959), this work has documented the difficulty people have in assessing the relative contributions of chance and skill in determining outcomes. The general finding is that for events that are difficult (low probability of success), the introduction of a skill component leads to overestimating the probability of success, a phenomenon that has also been labeled the "illusion of control" (Langer, 1975; Hogarth, 1987).

When it comes to interpreting outcomes of tasks involving both skill and chance, there is much evidence that people tend to attribute good outcomes to skill and bad outcomes to chance (Miller & Ross, 1975). This self-serving tendency can in turn lead to overconfidence in skill (Langer & Roth 1975; Gervais & Odean 2001). Moreover, although arguments can be made that, for some tasks, such illusions are beneficial in that they encourage proactive behavior (Taylor & Brown, 1988), they are clearly dangerous when payoffs are large (Makridakis, Hogarth, & Gaba, 2009).

At one level, the introduction of the chance component can be conceptualized as diminishing the validity of the signal that agents receive about their level on the criterion. Absent the chance component, this was already imperfect; with chance added, it becomes even more imperfect. Therefore, following the results of Camerer and Lovallo (1999) and Moore and Cain (2007), it is reasonable to hypothesize that agents will be less likely to enter the game as signal validity decreases, i.e., with the noise due to chance.[2]

On the other hand, consider the differential attitudes that might be exhibited by agents whose signals suggest that they have high and low criterion scores, respectively. Recall, first, that what matters is whether an agent is ranked among the first k of those who enter the market. Thus, an agent whose true ability merits a place in the first k can only be hurt by chance (i.e., demoted below the k^{th} position) but not helped. However, an agent whose true ability does not merit a place in the first k can only be helped by chance (i.e., promoted into the top k positions) but not hurt. This reasoning therefore leads to an alternative hypothesis. With the introduction of chance, agents who believe that they have relatively high scores on

[2] In Moore and Cain's (2007) study, the signal validity (correlation between agents' true and estimated scores on the criterion) was greater in the easy as opposed to the hard task (D. Moore, personal communication, July 2007).

the criterion will enter competition at a lower rate than in the absence of chance. On the contrary, agents who believe they have low scores will enter at a higher rate when the random factor is increased. It is, after all, the only way they can succeed. We test these hypotheses in a within-participant experiment described below.

As noted, agents need to estimate their probability of being ranked among the *k* best entrants and, yet, they only receive an imperfect signal that informs them about their absolute score on the criterion and not their relative ranking as such. An interesting issue therefore also centers on how accurate agents are in estimating both their criterion scores and relative rankings as well as what might affect this. For example, there is an extensive literature that demonstrates that people over- (under)estimate their abilities in hard (easy) tasks and simultaneously under- (over)place themselves (i.e., relative to others) (Moore & Healy, 2008). In our work, we did not manipulate the difficulty of the task that determines the signal, and thus have no hypothesis concerning the effects of confidence. On the other hand, we examine our data to observe whether any "excess entry" is accompanied by overconfidence.

The role of gender has also often been highlighted as affecting differential degrees of confidence and attitudes toward risk. It is generally claimed, for example, that men are more overconfident than women (Barber & Odean, 2001) and that women are both more risk averse than men (Byrnes, Miller, & Schafer, 1999) and less willing to place themselves in competitive situations (Niederle & Vesterlund, 2007). We therefore examine our data for gender effects bearing these considerations in mind.

We tested our main hypotheses of the differential effects of chance on those with low and high scores on the ranking criterion in an experiment with two conditions and using a within-subject design to control for individual differences. In one condition, participants were ranked by their skill and payoffs were fully determined by relative skill. In the other, skill rankings were adjusted by randomly drawn individual chance components such that the relation between relative skill and payoffs was imperfect. Moreover, participants were made aware of this fact.

Our results showed that low-ability participants (i.e., those who performed relatively poorly on the test) entered competition more when a combination of chance and skill, rather than only skill, determined their payoffs. The evidence that participants with high performance were more willing to enter competition when payoffs depend uniquely on skill is weaker. The high-ability participants, it seems, still hoped to achieve good outcomes whatever the role of chance. Interestingly, our data also show that participants tended to attribute success to their skill, while chance was more likely to be blamed for failures.

Overall, there was evidence of excess entry in that mean group profit was negative in most rounds as more participants entered the markets than should have. At the same time, we also found accurate assessments of numbers of entrants, no effect of initial confidence, and an understanding that past experience with chance is not predictive of future outcomes. Since the outcomes of market entry decisions in naturally occurring environments inevitably involve elements of both skill and luck, our results contribute to explanations of why low-skill participants still enter markets and, for the most part, fail.

Experiment

Participants and sessions. Participants were recruited through invitations sent to the members of the database system of the Leex laboratory of Universitat Pompeu Fabra and the experiment was conducted on computers in the laboratory using the z-Tree software (Fischbacher 1999). No participant took part in more than one session. Upon arrival participants were randomly assigned to seats. They were identified by code numbers only. Each participant had an individual printed copy of the instructions. Instructions also appeared on the screens. In addition, one experimenter read the instructions aloud. All questions were answered in private. Sessions lasted about one hour. There were six separate sessions, each with fifteen participants.

Procedure. Participants were informed that they were being given a credit of 15€ at the beginning of the experiment, and that at the end, their net earnings would be added to (or deducted from) the 15€. Moreover, their net earnings would depend on both their choices and the choices of other participants taking part in the same session. Participants were further told that the study was conducted anonymously and that their individual choices would not be known to other participants. Table 1 summarizes the experimental procedure for one of the experimental conditions explained in detail below.

Insert Table 1 about here

There were two parts to the experiment. The first involved ten problems of multiplying two 2-digit numbers. Participants could use a pen and scratch paper. They were paid 0.50€ for each correct answer and informed that the more questions they got correct, and the faster they were, the easier it would be for them to earn more money in the second part of the experiment. The ten problems appeared one by one on the screen. Participants had 30 seconds to solve each problem. If no answer was provided within 30 seconds, the question was counted as "incorrect" and the next problem appeared automatically. After solving ten problems, participants were asked how many problems they thought they had solved correctly and how many of the fourteen other participants in the room they believed did better than them (i.e., answered more questions correctly and/or faster). At the end of the experiment, they were paid 0.50€ for each correct estimate. Feedback on the multiplication test was not given until the end of the experiment.

The second part of the experiment was a market entry game involving an experimental market with a capacity of five entrants and a total payoff to be divided among successful entrants of 25€. Each participant had to decide privately whether or not to enter. If they decided to stay out, the additional payoff was zero. If they decided to enter, the payoffs were determined by the number of entrants and the results of the multiplication test. If the number of entrants was five or less, the 25€ was divided equally among the entrants. If more than five participants decided to enter the competition, all entrants were ranked according to the number of problems solved correctly in the multiplication task (time taken to complete the test was used to break ties). The five best entrants earned 5€ each and the others lost 10€ each.

Before the first entry decision was made, participants were informed that there would be a total of twelve similar rounds. At the end of the experiment, four of the twelve rounds were chosen randomly by throwing a die. Final payoffs for the market entry game were determined by the mean of individual payoffs in these four rounds. Participants were informed in advance about how the final payoffs would be determined.

Before the first round, four quiz questions were administered to make sure that everyone understood the rules of the game. Participants were given three attempts to answer each question. Eighteen participants failed to provide a correct answer to one or two questions. We analyzed the data both with and without these participants and found that excluding these 18 participants did not change results significantly. We therefore report below the analyses of all data.

At the beginning of each round, participants were asked to forecast how many entrants would enter (including themselves) in the round. Accurate forecasts were rewarded with 0.50€. Participants then decided individually whether or not to enter the competition. Post-decision feedback for each round included the number of entrants and the participant's individual payoff. After the first six and the last six rounds participants were asked to indicate on an 11-point scale how risky these rounds were in their opinion (0 = "not risky at all" and 10 = "extremely risky").

For the last six rounds, the procedure for determining the performance ranking was changed by including an explicit chance component. Participants were told about the changes in the rules of the game immediately before the block of the modified rounds. In particular, at the beginning of each round, individual random *chance parameters* were drawn from a uniform distribution. There were seven possible “levels of chance”: an improvement of the position in the original ranking by three places, two places, and one place; a worsening of the position in the ranking by three places, two places, and one place; and no change of the original position. The process of generating chance parameters was described to participants using the analogy of an urn containing seven balls of different colors where colors determined chance. In each modified round, after receiving performance feedback, participants were additionally asked to indicate on 11-point scales how lucky they thought they were in the round (0 = “not at all” and 10 = “extremely lucky”) and how fair they thought their payoff was in the round (0 = “not at all” and 10 = “extremely fair”). Finally, after the block of six modified rounds, participants additionally indicated on 11-point scales how fair they thought their result was overall in these rounds (0 = “not at all” and 10 = “extremely fair”) and what role, in their opinion, chance played in determining their payoffs in these six rounds (0 = “chance has not played any role”; and 10 = “chance has been decisive”).

After all twelve rounds, participants received feedback on the number of correct answers and their total time spent on the multiplication test, indicated their age and sex, and answered 29 questions of the internal-external(IE)-scale questionnaire (Rotter, 1966) that purports to measure individual locus of control. Locus of control refers to the extent to which individuals believe that they can control events. Individuals with a high internal locus of control (a low score on the Rotter scale) believe that their actions mainly determine their future outcomes whereas those at the other extreme believe that they have little control over what happens to them.

Finally, a 12-faced die was then thrown to determine the four rounds to be used to calculate total payoffs and the information on individual total payoffs appeared on the participants' screens. Participants were paid privately one-by-one at the end of the experimental session.

Design. The experimental design included two within-subject variables and two between-subject variables. The within-subject variables were round number and whether a randomly generated *chance* component was (*chance rounds*) or was not used (*baseline rounds*) to determine performance ranking in a given round. The between-subject variables were the level of chance and order of rounds. In four sessions (of eight), chance component was enlarged. In these sessions, the position in the original ranking could increase or decrease by nine, six, or three places (vs. three, two, one place in the other chance sessions). We refer to these sessions as the *big chance* condition and to the other sessions as the *small chance* condition. As to the order of rounds, baseline rounds preceded chance rounds in four sessions, and chance rounds preceded baseline rounds in the other four sessions. In summary, there were four between-subject conditions: (1) baseline – small chance, (2) small chance – baseline, (3) baseline – big chance, and (4) big chance – baseline.

Main hypotheses. The main question we pose in this study is how adding noise to agents' evaluations affects their willingness to enter competitive situations. We hypothesize that there is a difference between the responses of low-skill and high-skill participants.

> *Hypothesis 1:* Participants who score low on the test will enter the competition more in the chance than in the baseline condition.
>
> *Hypothesis 2:* Participants who score high on the test will enter the competition less in the chance than in the baseline condition.

Insert Figure 1 about here

To illuminate how the random perturbation of rankings in the chance conditions would affect individual ranks, we ran simulations for each value of initial rank (from 1 to 15). The results are shown in the upper and lower parts of Figure 1 for small and big chance, respectively. The simulations were done over 10,000 trials and show, for each rank, the distribution of ranks that would obtain. (Ties were resolved randomly). Thus, with small chance, someone ranked 1st would rarely fall below the 5th rank whereas someone ranked 5th would have a good chance of losing by falling below that position. Similarly, participants with ranks worse than the 10th position would rarely succeed (i.e., be better than 6th) while those in the middle ranks (e.g., 6th through 9th) start to have chances of success that would have been denied in the absence of chance. With large chance, the simulations show the same kinds of trends except that the distributions are naturally much flatter.

Equilibrium predictions. Assuming risk neutrality and no private information about the probability of success on entry, there are multiple pure-strategy Nash equilibria with seven players (47%) entering a market that has a capacity of five. Provided that players cannot coordinate, there is a mixed-strategy equilibrium in which each risk-neutral player (without private information about the probability of success on entry) enters with probability of 53%. This corresponds to 7.9 players (out of 15). Detailed calculations are presented in the Appendix.

If all players know their relative performance on the test, then only the top five players (33% of all potential entrants) enter. However, if players have imperfect information about their test performance, the equilibrium number of entrants will be higher when players overestimate their probabilities of success and lower when the probability of success is underestimated. This idea is developed further in the Appendix.

We do not make any specific assumption about participants' risk preferences but control for it within-subject since all participants took part in both baseline and chance conditions (cf., Camerer & Lovallo, 1999; Moore & Cain, 2007). Our primary measure of interest is within-subject differences in behavior between the baseline and chance conditions.

Results

Of the 120 participants, 57 were male. Participants were between 18 and 36 years of age, 21 on average. Total earnings per participant were between 5.50€ and 25.90€ with a mean of 18.23€ and median of 18.63€. The mean number of correctly solved problems in the multiplication test was 8.1 (out of 10), the median was 9.0. There were on average nine entrants per round. Mean group payoff in the market entry game across all sessions was -17€. Figure 2 details the distribution of group payoffs.

The effect of chance. Our main question is whether adding the chance component to the evaluation procedure affected participants' decisions to enter the competition and whether the effect was different for participants who scored low as opposed to high on the test. The data support the following main results:

Result 1: Low-skill participants enter competition more when the outcome-determining procedure is noisier.

Result 2: High-skill participants enter competition somewhat less when the outcome-determining procedure is noisier.

Result 3: High-skill participants are less sensitive to changes in the outcome-determining procedure than low-skill participants.

Tables 2 and 3 report the results of several population-averaged logit analyses of the probability of entry. The dependent variable was coded as 1 if the participant decided to enter competition in a given round and as 0 otherwise. As there is an observation for each individual in each round, each individual appears in the sample 12 times (six baseline rounds and six chance rounds). Because of the multiple observations on individuals, standard errors were calculated accounting for the correlation across repeated observations.

Insert Tables 2 and 3 about here

Table 2 reports the coefficient estimates of several logit specifications which capture the effect on entry decisions when the evaluation criterion is perturbed by chance. Table 3 reports the estimated change in the probability of entry and the standard errors on this change in probability separately for high and low-skill participants. Model 1 is a specification with a constant; indicator variables for the chance conditions, big chance, gender; rank (indicating participants' relative performance in the test); and the interaction between rank and chance. It additionally controls for the IE score, forecasted number of entrants, possible order effects (an indicator variable) and round (a series of indicator variables). The interaction term between rank and chance is significant in this specification.[3] Figure 3, left panel, presents the predicted probability of entry for all participants as a function of rank, separately for the baseline and chance conditions. First, the figure shows that, as might be expected, the probability of entry is lower for lower ranked participants. Second, it shows that for participants who ranked low on the test, the probability of entry is higher in the chance as opposed to the baseline condition. For example, for participants who ranked 12th, the probability of entry was 0.40 in the baseline but 0.53 in the chance condition. Participants ranked higher than the median (8) do not appear to be sensitive to the presence of noise.

Insert Figure 3 about here

The estimated changes in the probability of entry presented separately for high and low-skill participants in Table 3 provide further evidence on how the impact of chance

[3] An analogous model without the interaction term yields a lower model fit (χ^2=81.98). The difference in fit is statistically significant ($p<0.001$).

depends on skill. We split participants by their median performance classifying them as low-skill if they ranked 8 or worse and as high-skill otherwise. The results of these specifications imply that entry rates of low-skill participants are 13 percentage points higher in the chance as compared to the baseline condition. The effect is significantly different from zero ($p<0.001$). For high-skill participants, entry rates are 3 percentage points lower in the chance as compared to baseline condition. The effect, however, is not statistically significant, suggesting that high-skill participants are equally willing to enter competition in the chance and baseline conditions. The data are thus consistent with hypothesis 1 in that low-skill individuals entered competition more in the chance than baseline condition. On the other hand, evidence in favor of hypothesis 2 is not strong in that the difference in entry behavior of high-skill participants, although pointing in the hypothesized direction (i.e., less entry in the chance than in baseline condition), was not statistically significant.

Parenthetically, descriptive data are consistent with the results of the logit models. On average, participants with the worse rankings (8^{th} or worse) entered the market in 2.5 baseline rounds and 3.3 chance rounds (both out of six). Participants with the better rankings (7^{th} or better) entered the market in, on average, 4.6 baseline rounds and 4.5 chance rounds. In short, chance induced the poorer but not the better performers to enter the competition more. The effects of chance, however, were generally not positive. Mean losses of the less skilled participants (8^{th} or worse) were approximately the same in the baseline and chance rounds: 3.82€ and 3.43€. The difference is not statistically significant ($z = -1.35$, $n = 63$, Wilcoxon signed-rank test). The higher ranked participants (7^{th} or better) gained less in the chance than baseline rounds: 0.83€ vs. 2.41€ (the difference is statistically significant, $p<0.001$, $z = 3.35$, $n = 57$, Wilcoxon signed-rank test).

The second specification in Table 2 (Model 1a) includes in addition the entry decision made in the previous round and the payoff earned in that round. (It thus excludes the first round for which no lagged data are available). Controlling for the lagged decision and payoff does not eliminate the interaction effect of chance and rank. (The plot of predicted probabilities as a function of rank looks similar to that presented in Figure 3, left panel). Both lagged variables are significant in this model implying that the probability of entry in a given round is higher when entry occurred in the previous round and when the payoff of the previous round was larger.

In addition, the effect of lagged payoff differed in the baseline and chance rounds (the logit specifications are not presented here). Analogous models fit separately on the data from the baseline and chance conditions showed that in the baseline conditions, earning 1€ more in a given round increases the probability of entry by 4 percentage points (se = 0.007, $p<0.001$). In the chance condition, the marginal effect of lagged payoff was not significant ($dy/dx = 0.001$, se = 0.004). This implies that participants correctly understood that past performance in the competition was less predictive of future performance in the chance as opposed to baseline rounds.

Although our experiment was designed to test the effect of chance *within* subjects, we also investigated whether the differential behavior of low- and high-skill individuals in the face of chance was mirrored by differences *between* the big and small chance conditions. Model 2 (Tables 2 and 3) compares entry rates of low- and high-skill participants in the big chance and small chance conditions (between-subject comparison). Neither the indicator variable *big chance* nor its interaction with rank is significant in this model (Table 2). Marginal effects of the indicator variable *big chance* (Table 3) calculated separately for low- and high-skill participants are not significant ($p>0.05$) implying that participants were not sensitive to the levels of chance. However, the signs of the marginal effects are in the hypothesized direction: negative effect on entry rate for high-skill participants and positive for low-skill participants.

Finally, the specifications presented in Tables 2 and 3 show that there was no consistent difference in the predicted probabilities of entry for participants of different gender, participants who had different IE-scores, and participants who forecasted more vs. less entrants in experimental rounds.

Excess entry. The data reveal "excess entry" in that there were between six and fourteen entrants each round (40% - 93% of all participants). (Recall market capacity was five entrants). Figure 4 shows the actual number of entrants in each session and the mean number of entrants predicted by participants at the beginning of each experimental round. As a reference point, we have added horizontal lines corresponding to the mixed-strategy equilibrium prediction of 7.9 entrants (assuming risk-neutrality and no private information about relative performance on the test).

Insert Figure 4 about here

The decreasing slope of entry curves in Figure 4 suggests some learning from performance feedback. However, the number of entrants only approached the mixed-strategy equilibrium of 7.9 players in some rounds. A surprising finding is the relative accuracy with which participants anticipated the number of entrants. There was no consistent tendency to over- or underestimate the number of entrants. The mean difference between the mean predicted and actual number of entrants was - 0.1, and the mean absolute difference was 2.0 persons. There was no significant difference between the baseline and chance rounds in terms of the magnitude of the errors.[4]

Confidence in test performance. Participants were better at estimating their score (i.e., absolute "skill") than their rank (i.e., comparative "skill"). Of the 120 participants, 43 correctly estimated their score and only 11 correctly estimated their rank. The correlation

[4] The mean difference between expected and actual entries was 0.1 persons in the baseline and -0.3 persons in the chance condition. The mean absolute difference between expectations and realizations was 2.1 persons in the baseline and 2.0 persons in the chance condition.

between actual and estimated scores was 0.65 (Pearson, $p<0.001$, $n = 120$); the correlation between actual and estimated rank was 0.38 (Pearson, $p<0.001$, $n = 120$). Moreover, estimated score was a better predictor of actual rank (i.e., the criterion relevant to entry decision) than estimated rank (Pearson coefficient of -0.63, $p<0.001$, $n = 120$). Overall, participants' estimates of their scores and ranks were imperfect.

To quantify overconfidence and underconfidence, we defined confidence about score as estimated score minus actual score and confidence about rank as estimated rank minus actual rank. Positive values indicate overconfidence, negative imply underconfidence. Of the 120 participants, 58 were overconfident about their score and 19 were underconfident. As to confidence about ranks, 70 were overconfident and 39 were underconfident. Figure 5 depicts mean confidence for different performance levels. The upper panel shows that participants who solved more problems correctly on the test were less overconfident, on average, than those who solved less. The lower panel shows that participants who ranked among the best five were on average underconfident about their rank, while participants who ranked worse were on average overconfident.

Insert Figure 5 about here

Note that participants estimated their performance in the test before the market entry game. The decreasing slopes of the entry curve in Figure 4 suggest that these initial estimates were refined during the market entry game as participants received feedback on their performance. Indeed, as reported above, the probability of entry in a given round was higher when the payoff from the previous round was larger. Thus, confidence in skill elicited before the market entry game can only be taken to represent an initial level of confidence rather than a characteristic that was stable during all experimental rounds.

In order to estimate the impact of initial confidence on entry decisions, we included confidence about score and confidence about rank in Model 3 (Table 2). Our interest was to understand whether, given the same level of performance in the tests, greater confidence implied higher probability of entry. We therefore also control for actual score and rank. The effects of neither confidence about score nor confidence about rank are significant in this specification, and the marginal effects of these variables are virtually null (not shown in the table). The implication is that the initial level of confidence did not affect entry decisions.[5]

Gender effects in confidence and entry. Men were more confident about their performance than women. Table 4 details the proportions of underconfident, overconfident, and well-calibrated (i.e., estimates equal to actual values) male and female participants as well as mean confidence within each group. While the details on confidence about test score look similar for men and women (most of them slightly overestimating their score, a few underestimating), women who underestimate their rank do so much more than men (mean underestimation of 3.7 vs. 2.2 positions in the ranking).[6] There was no gender difference in actual ranks (means of 8 for women and men, $z = 0.77$, *ns*, $n = 120$, Wilcoxon rank-sum test). A linear regression model of estimated rank with an indicator *female* as independent variable and actual rank as control yielded a coefficient of 2.1 for *female* (robust se = 0.49, $p<0.001$,

[5] Initial confidence did not affect entry decision in the first round either. A logit model of entry that was fit on the data from the first rounds and included estimated and actual score and rank as independent variables, yielded no significant coefficients. Analogous models fit separately on the data from the baseline and chance rounds showed that participants with larger estimated scores were slightly more likely to enter the competition in the first baseline rounds than participants with smaller estimated scores (10 percentage points per each addition unit of estimated score, se = 0.06, $p = 0.08$, $n = 60$). Estimated score and rank were not significant in explaining the probability of entry in the first chance rounds ($n = 60$).

[6] Overall, the gender difference in confidence about scores was not statistically significant ($z = 0.94$, $n = 120$, Wilcoxon rank-sum test), while the difference in confidence regarding placement in the ranking was significant ($p<0.01$, $z = 2.95$, $n =120$, Wilcoxon rank-sum test).

n = 120, $R^2 = 0.26$) suggesting that women were indeed less confident about their rank than men.[7]

As to the decision to enter competition, the specifications presented in Table 2 and the marginal effects presented in Table 3 fail to detect any consistently significant gender difference in entry behavior. The probability of entry was not lower for women than for men. If anything, women were somewhat more likely to enter competition, as suggested by the positive effect of the indicator variable *female* in these specifications.

To investigate possible gender effects in how low- and high-skill participants behaved when chance was added to the evaluation procedure, we analyzed the probability of entry separately for men and women using specifications identical to Model 1 presented in Table 2. The predicted probability of entry from these gender-specific logit models is plotted in the middle (women) and right (men) panels of Figure 3. (See also Table A2 in the Appendix that details the results from the logit models). The gender-specific predicted probabilities mirror the general result in that the probability of entry for low-skill female and male participants (higher values of rank) is larger in the chance than baseline condition. As to high-skill participants, women tend to enter less with chance than without (i.e., the interaction of rank and chance) while there is no apparent difference between the chance and baseline conditions among high-skill male participants.

Since further splits of data (by skill, within gender) reduce sample size, we prefer caution in drawing inferences about the behavior of high-skill participants and thus conclude that, if anything, the actions of high-skill participants were only modestly affected by chance (as marginal effects in Table 3 indicate, first column).

[7] The gender gap in relative confidence echoes the findings of Niederle and Vesterlund (2007). In their experiment, men and women performed equally well on a test, but women were more likely to underrate their relative performance: 75% of men thought they were best in their group (of four participants), while only 43% of women shared this belief. In Niederle and Vesterlund’s experiment, higher overconfidence among men partially explained why men were more willing than women to select into a competitive environment.

Attributions, risk, and fairness. Participants' answers to the post-round questions of how risky the rounds were indicate beliefs that baseline rounds were less risky than chance rounds. The median evaluation of risk (0-10 scale) was 5 for baseline rounds and 7 for chance rounds. The difference is statistically significant ($p<0.001$, $z = -4.87$, $n = 120$, Wilcoxon signed-rank test). This result can be considered a manipulation check that participants understood that payoffs were more unpredictable in the chance as opposed to baseline condition.

Participants also correctly indicated that chance played a bigger role in big as opposed to small chance sessions (where the possible impact of chance was three times smaller) – medians of 8 vs. 5 on a scale where 0 = "chance has not played any role" and 10 = "chance has been decisive" ($p<0.001$, $z = -3.46$, $n = 120$, Wilcoxon rank-sum test, between-subject comparison).

Individuals who earned less money in the chance rounds were more likely to attribute the result to chance than those who earned more, as evidenced by a negative correlation between the mean profit that each participant earned in the chance rounds and beliefs about the role of chance (Spearman's rank correlation of -0.25, $p<0.01$, $n = 120$).

Finally, larger individual payoffs were considered as being "fairer." Specifically, the Spearman's rank correlation between mean individual payoffs across the chance rounds and reported fairness (0 = "not fair at all" and 10 = "extremely fair") was 0.66 ($p<0.001$, $n = 120$).

General Discussion

Our experimental innovation augmented the validity of the market entry paradigm by introducing an explicit element of chance in how participants are ranked. An analog in naturally occurring situations could be, say, uncertainty concerning the membership of a committee deciding on grant proposals or job applications. Consider also new product launches. An entrepreneur might well have a good sense of her domain-specific competence relative to competitors, but also be aware that unknowable consumer tastes could perturb the market ranking of potential entrants. Book publishing provides a further example. An author can have a good track record and rank highly among peers; but this does not guarantee that a new novel will be accepted by the public or critics.

In summary, our results provide evidence of excess entry that was exacerbated when participants knew that chance would affect how they were ranked. Interestingly, the effects of chance were moderated by how well participants scored in the evaluation procedure (i.e., the math test). Consistent with our first hypothesis, participants with lower scores were more likely to enter the competition in the presence of chance (Result 1) but, contrary to our second hypothesis, the actions of higher scoring participants were only modestly affected by chance (Results 2 and 3).

Additional findings were no differences in entry decisions due to levels of chance (i.e., big chance versus small chance), and whereas men were more overconfident than women, men did not enter the competition more than women. Contrary to the claims of Camerer and Lovallo (1999), we found no relation between overconfidence and excess entry. Moreover, in post-round questionnaire responses participants recognized that chance rounds were riskier than their baseline counterparts and that big chance conditions were riskier than little chance conditions. On the other hand, there was evidence of a self-serving bias in that participants who earned less money in the chance rounds were more likely to attribute their outcomes to chance than those who earned more.

We can view our main results from different perspectives. The first is simply that the participants were confused by the introduction of the chance component and reacted by responding more to what they perceived as the game implicit in the task, i.e., matching more noise in the task by more noise in their responses. As noted previously, people have a tendency to seek risk when games involve explicit components of both skill and chance.

The results might also be viewed from the viewpoint of Atkinson's (1957) work on "need for achievement." The key idea here is that, when choosing between skill-related tasks of different difficulty levels, people motivated to achieve success as opposed to avoid failure tend to select tasks that are intermediate in difficulty. In other words, there is some preference for uncertainty in outcomes that is consistent with preferring outcomes based on skill and chance as opposed to skill alone. Whereas we suspect that our experimental participants are more likely to be motivated by achieving success than avoiding failure, we did not explicitly measure "need for achievement" and thus must leave testing this hypothesis to future work.

A problem with these two explanations is that they do not account explicitly for the differential increase in risk taking observed by participants who scored low as opposed to high on the test. Instead, we suspect that participants with low scores on the criterion realized that the introduction of an explicit chance element could only improve their probability of success and were thus willing to gamble more on entering the competition. However, this sensitivity did not lead to different behavior in the big as opposed to small chance condition. In other words, our data showed sensitivity to the introduction of chance but not to its level.[8] At the same time, it is puzzling why those who scored well on the evaluation criterion didn't reduce their risk taking activity in the presence of chance.

If we accept that our participants' behavior is not fully rational, there are further ways of explaining the behavior of the high scoring participants. One is a belief in the "law of small numbers" (Tversky & Kahneman 1971) which implies that the high scoring participants didn't anticipate much variation in outcomes from their expected scores. This hypothesis could be tested in future research by changing the nature of the chance component. For

[8] Note, however, that whereas the baseline-luck comparison was made within subjects, our design only permitted a between-subjects comparison between small and big luck. It will be useful, therefore, to extend our design to allow for within-subject comparisons in future work.

example, if instead of all participants' rankings being perturbed, what would happen if only the rankings of a limited subset of participants (chosen at random) were affected? More generally, there are many different ways in which chance could be introduced. For example, *k*, the number of successful entrants, could vary randomly across trials.

Second, more entry in the presence of both skill and chance is consistent with concerns about sustaining positive self-image when tested by skill-sensitive tasks (Larrick, 1999; Koszegi, 2006). In the case of failure, chance can be blamed for doing worse than expected such that positive self-image is maintained. In the case of success, self-image is enhanced. Therefore, more entry should be observed when skill together with chance determines outcomes. In fact – and as noted above – the participants who earned less money in our market entry game when both chance and skill determined their performance were more likely to attribute the result to chance. This result is also consistent with the literature on self-attribution bias (Miller & Ross, 1975).

Overall, our participants took too much risk and, as a result, their payoffs were lower than could have been achieved by participating less in the competition. In addition, although there was evidence of some learning through repeated experience, this did not eliminate excess entry. At the same time, our participants seemed to understand correctly that past performance is less predictive of future performance when chance plays a bigger role.

In our market entry game, confidence in own performance, measured on both absolute and relative levels, did not affect entry decisions. What we observed resembled imperfect self-selection into the game based on skill (i.e., test performance). (For a theoretical model of entry illustrating such self-selection, see Hogarth & Karelaia, 2009). In addition, men were more confident than women in their performance (see also Lenney, 1977; Barber & Odean 2001). And yet, men did not compete with other participants overall relatively more often than women contrary to findings of Niederle and Vesterlund (2007) that showed men entering a competitive tournament more than women for any performance level. Why might these findings differ? First, although Niederle and Vesterlund also used a math test to measure performance and confidence, their participants took decisions in tournaments that differed from our experimental task. Second, whereas decisions in our experiment were made in semi-private cubicles, Niederle and Vesterlund's participants were involved in a more face-to-face situation that could have had an impact. It is also possible that in our (mostly undergraduate) sample, women were, on average, used to performing academically as well as men and thus did not take less risk in a task related to performance.

Our investigation poses an intriguing puzzle. Our respondents were remarkably accurate at predicting the number of participants entering each round (across all conditions). At the same time, they only overplaced their ranking on average by one position. Why then did they enter markets where the most likely gain was 5€ but the cost of being wrong was 10€? In both the baseline and chance condition, there was excess entry in that mean profits were negative each round. There seemed to have been a pervasive overoptimism or "myopic self-focus" (Moore et al., 2007) that was only marginally affected by learning. We suspect that the fact that payoffs reflected relative skill contributed to some measure of illusion of control (Langer, 1975). Further research should therefore investigate conditions of illusion of control in market entry competitions. What levels of feelings of skill are necessary for participants to feel they have "control"?

In summary, we augmented the realism of the market entry paradigm by including an explicit chance component in determining payoffs and found that people take more risk when both skill and chance, as opposed to skill alone, determine outcomes of their actions. Our data support the explanation that for people who assess their own skill as low, greater uncertainty induces more risk taking. Although not entirely "rational," the reason, we suggest, is not unreasonable. People with low skill know they cannot succeed if outcomes only depend on skill. Chance is their only path to success even though, on average, by taking action most will fail. On the other hand, people who assess their skill as high still hope that this will ensure good outcomes whatever the role of chance. In Alexander Pope's famous words "Hope springs eternal" but it also seems that it needs (the) chance to do so.

References

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*, 359-372.

Barber, B., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, *116 (1)*, 261-292.

Benoît, J.-P., & Dubra, J. (2009). *Overconfidence?* Unpublished working paper, London Business School and Universidad de Montevideo.

Byrnes, J. P., Miller, D. C., & Schafer, W.D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125(3)*, 367-383.

Camerer, C. F., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, *89 (1)*, 306-318.

Cohen, J., & Dearnaley, E. (1962). Skill and judgment of footballers in attempting to score goals: A study of psychological probability. *British Journal of Psychology*, *53*, 71-86.

Cohen, J., Dearnaley, E. J, & Hansel, C. E. M. (1956). Risk and hazard: Influence of training on the performance of bus drivers. *Journal of the Operational Research Society*, *7*, 67–82.

Cohen, J., & Hansel, M. (1959). Preferences for different combinations of chance and skill in gambling. *Nature*, *183*, 841-843.

Fischbacher, U. (1999). *Ztree: A Toolbox for Readymade Economic Experiments*. Working Paper, Institute For Empirical Research In Economics, University of Zurich.

Gervais, S., & Odean, T. (2001). Learning to be overconfident. *Review of Financial Studies*, *14 (1)*, 1-27.

Hogarth, R. M. (1987). *Judgement and Choice: The Psychology of Decision*, (2nd Ed.), Chichester, UK: John Wiley and Sons.

Hogarth, R. M., & Karelaia, N. (2009). *Entrepreneurial success and failure: Confidence and fallible judgment*. Working Paper, Universitat Pompeu Fabra, Barcelona.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science, 39 (1),* 17-31.

Koszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association, 4 (4),* 673-707.

Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology, 77 (2),* 221-232.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32 (2),* 311-328.

Langer, E. J., & Roth, J. (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. *Journal of Personality and Social Psychology*, *32 (6),* 951-955.

Larrick, R. P. (1993). Motivational factors in decision theories: The role of self-protection. *Psychological Bulletin*, *113*, 440-450.

Lenney, E. (1977). Women's self-confidence in achievement settings. *Psychological Bulletin, 84*, 1–13.

Makridakis, S., Hogarth, R. M., & Gaba, A. (2009). *Dance with Chance: Making Chance Work for You.* Oxford, UK: Oneworld Publications.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*, 213-225.

Moore, D, A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior and Human Decision Processes*, *103*, 197-213.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115 (2),* 502-517.

Moore, D. A., Oesch, J. M., & Zietsma, C. What competition? Myopic self-focus in market-entry decisions. *Organization Science, 18 (3)*, 440-454.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics, 122 (3),* 1067-1101.

Rapoport, A., Seale, D., Erev, I., & Sundali, J. (1998). Equilibrium play in large group market entry games. *Management Science, 44 (1),* 119-141.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcements. *Psychological Monographs, 80 (1),* 1-28.

Taylor, S. & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103(2),* 193-210.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76(2),* 105-110.

Table 1. Summary of experimental procedure

Task 1

10 multiplication problems (0.50€ for each correct answer)

"How many correct?"

"How many of the 14 others did better than you?"

Task 2

Market entry game

1. Quiz to check understanding

2. First 6 rounds (no chance component)

For each round:

a. Forecast number of entrants (0.50€ for accurate forecasts)

b. Decide to enter or not

c. Feedback: number of entrants and individual payoff

3. After first 6 rounds:

"How risky were the rounds?"

3. Second 6 rounds (chance component)

For each round:

a. Forecast number of entrants (0.50€ for accurate forecasts)

b. Decide to enter or not

c. Feedback: number of entrants and individual payoff

d. "How lucky in round?" & "How fair was payoff?"

5. After second 6 rounds:

"How risky were the rounds?"

"How fair were rounds overall?"

"What role of luck in payoffs?"

Feedback on performance in Task 1 (multiplication test)

Completion of Rotter's (1966) I-E test

Participants remunerated individually

Note: The procedure is shown for the experimental condition for which there was no explicit chance condition in the first six rounds but there was in the second six rounds.

Table 2: Logit models of entry

	Model 1	Model 1a	Model 2 (Chance condition)	Model 3
Constant	0.81 (0.56)	-0.23 (0.58)	0.28 (0.59)	2.85 (1.63)
Chance (0/1)	-0.45 (0.30)	-0.46 (0.29)		-0.48 (0.30)
Rank	-0.22 (0.04) ***	-0.16 (0.04) ***	-0.18 (0.05) ***	-0.28 (0.07) ***
Rank*Chance	0.08 (0.03) **	0.09 (0.03) **		0.09 (0.03) **
Big chance (0/1)	-0.10 (0.26)	-0.01 (0.23)	-0.73 (0.61)	-0.02 (0.23)
Rank*Big chance			0.08 (0.07)	
Order (0/1)	0.31 (0.26)	0.46 (0.23)	0.25 (0.41)	0.58 (0.26) *
Female (0/1)	0.57 (0.27)	0.59 (0.24) *	0.45 (0.30)	0.52 (0.26)
IE score	0.06 (0.03)	0.05 (0.03)	0.05 (0.04)	0.05 (0.03)
Forecasted number of entrants	0.07 (0.03)	0.04 (0.03)	0.10 (0.04) *	0.04 (0.03)
Entry in the previous round		1.08 (0.23) ***		1.07 (0.23) ***
Payoff in the previous round		0.05 (0.02) **		0.05 (0.02) **
Confidence about rank				0.00 (0.04)
Confidence about score				-0.03 (0.11)
Score				-0.27 (0.13) *
Indicator variables for Rounds 2-12	Included	Included	Included	Included
Model χ^2	98.78	178.00	34.45	191.62
Number of observations	1440	1320	720	1320
Number of participants	120	120	120	120

Notes:

Population-averaged models were fitted. Standard errors, reported in parentheses, correct for correlation across repeated observations on individuals. *** p<0.001; ** p<0.01; * p<0.05.

Table 3: Predicted change in the probability of entry: Results from logit models

	Model 1		Model 2 (Chance condition)	
	high-skill	low-skill	high-skill	low-skill
Chance (0/1)	-0.03 (0.04)	0.13 (0.04) ***		
Rank	-0.06 (0.02) ***	0.00 (0.02)	-0.05 (0.02) **	0.00 (0.02)
Big chance (0/1)	-0.04 (0.07)	0.01 (0.08)	-0.02 (0.08)	0.05 (0.09)
Order (0/1)	0.08 (0.07)	0.07 (0.08)	0.13 (0.11)	0.03 (0.14)
Female (0/1)	0.08 (0.07)	0.12 (0.08)	-0.01 (0.08)	0.17 (0.10)
IE score	0.00 (0.01)	0.02 (0.01)	0.01 (0.01)	0.01 (0.01)
Forecasted number of entrants	0.01 (0.01) *	0.02 (0.01)	0.01 (0.01)	0.03 (0.02) *
Indicator variables for Rounds 2-12	Included	Included	Included	Included
Model χ^2	92.82	81.56	24.52	24.89
Mean predicted probability	0.80	0.49	0.77	0.55
Number of observations	684	756	342	378
Number of participants	57	63	57	63

Notes: High-skill individuals are ranked 7 or better. Low-skill individuals are ranked 8 or worse. Population-averaged models were fitted.

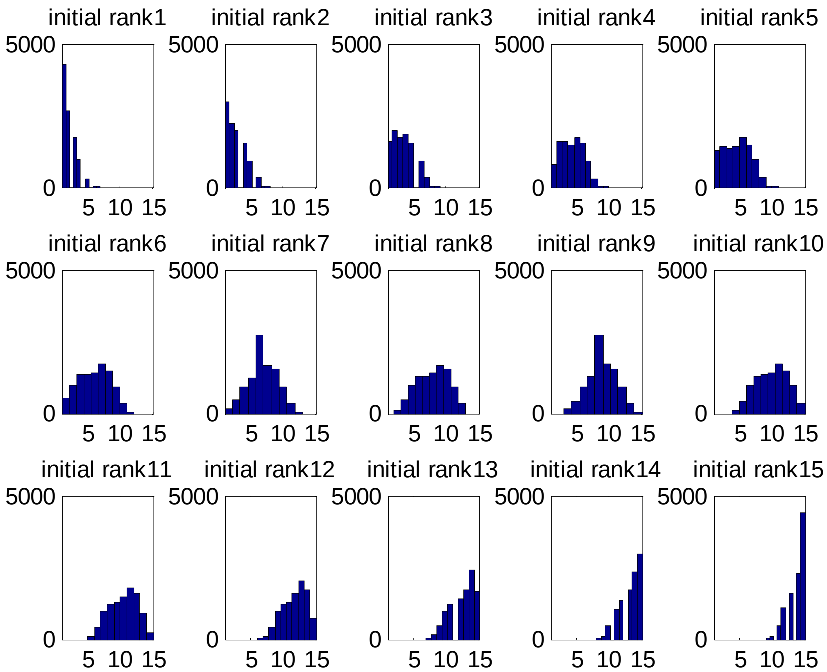
Standard errors, reported in parentheses, correct for correlation across repeated observations on individuals. *** p<0.001; ** p<0.01; * p<0.05.

Table 4: Confidence about performance in the test

		Men		Women		Total	
		n	Mean confidence	n	Mean confidence	n	Mean confidence
Score							
	overconfident	30	1.8	28	1.7	58	1.8
	underconfident	9	-1.0	10	-1.5	19	-1.3
	well-calibrated	18		25		43	
	Total	57	0.8	63	0.5	120	0.7
Rank							
	overconfident	40	4.8	30	4.1	70	4.5
	underconfident	12	-2.2	27	-3.7	39	-3.3
	well-calibrated	5		6		11	
	Total	57	2.9	63	0.3	120	1.6

Figure 1: Effects of chance on relative rankings

Small Chance



Big chance

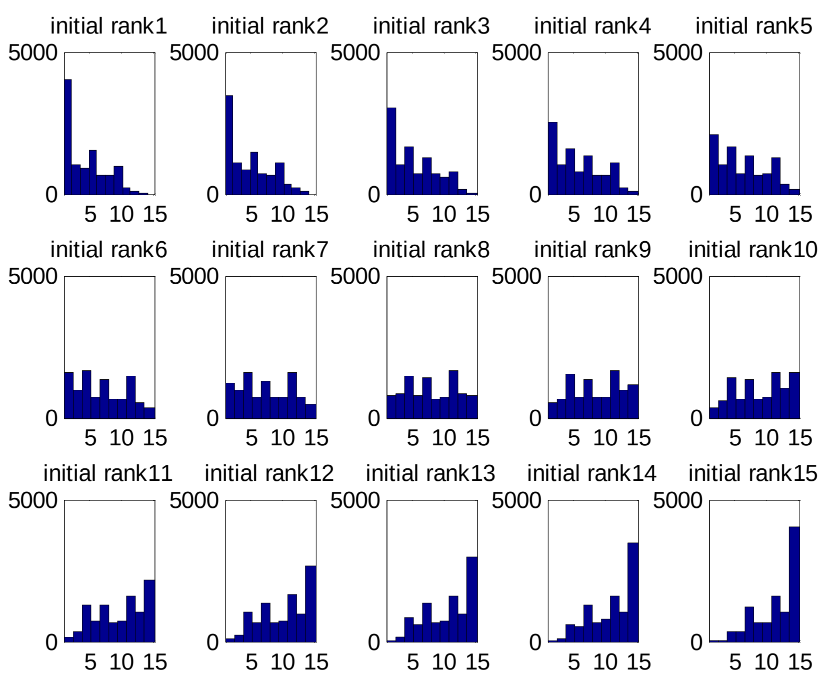
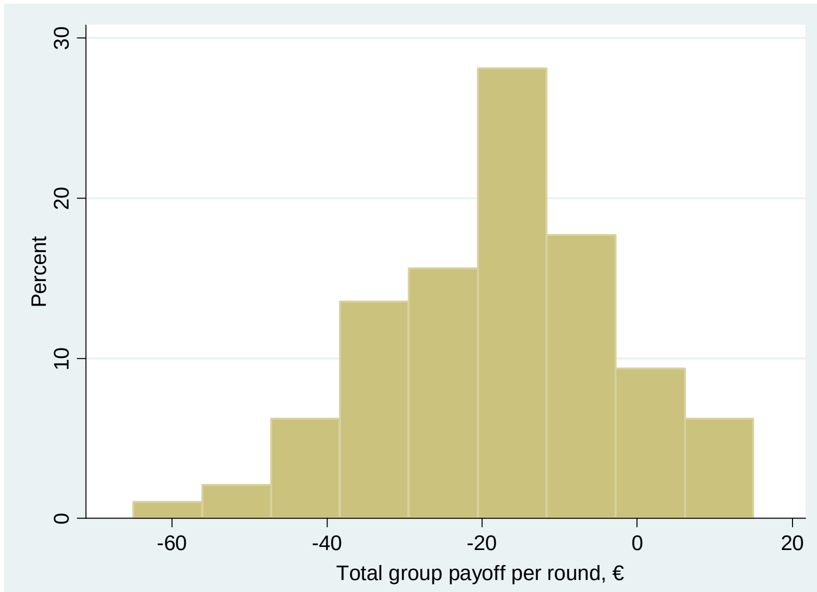
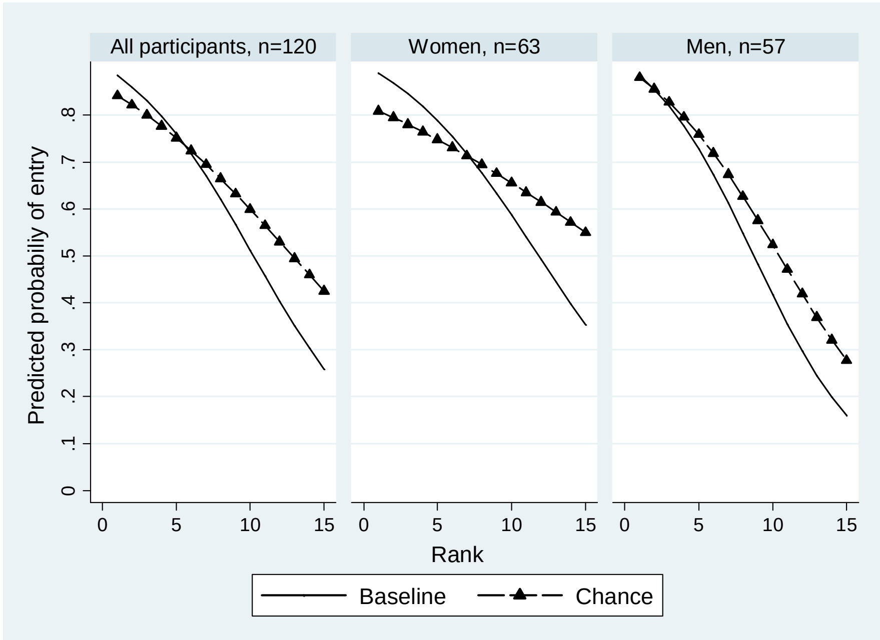


Figure 2: Histogram of group payoffs.



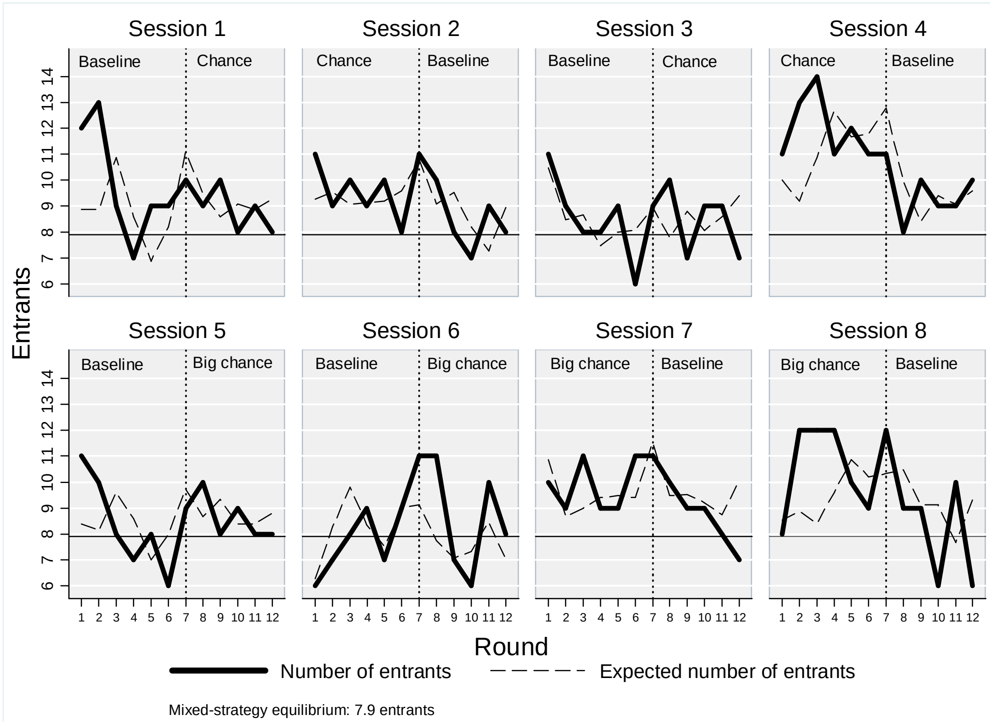
n = 96 rounds

Figure 3: Predicted probability of entry: Results from logit models



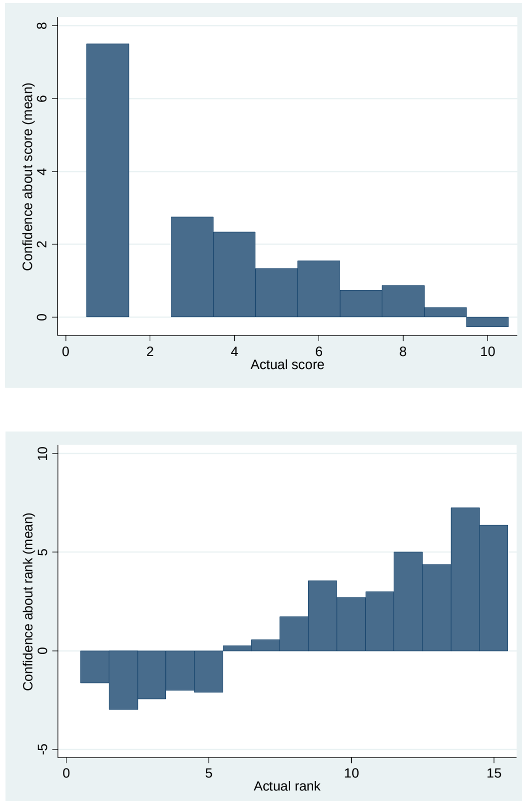
Note: The coefficients of the models are presented in Tables 2 (Model 1) and A2.

<u>Figure 4: Entry by session and round</u>.



n = 15 participants in each session.

Figure 5: Confidence about performance in the test



Note: Confidence about score (upper panel) = estimated score – actual score
Confidence about rank (lower panel) = estimated rank – actual rank

Appendix A: Equilibrium entry predictions.

Assuming risk neutrality and no private information about the probability of success on entry, there are multiple pure-strategy Nash equilibria with seven players (47%) entering a market that has a capacity of five. At equilibrium, participants do not expect to receive a larger payoff by changing their strategy, i.e., by entering if the decision was to stay out and staying out if the decision was to enter. In our game, when there are fewer than seven entrants, a participant who stayed out could have received a positive payoff by entering, and with more than seven entrants, a participant who entered could have avoided an expected loss by staying out. In particular, the expected payoff of each of seven entrants is $\frac{5}{7}\frac{25€}{5}+\frac{2}{7}(-10€)\approx 0.71€$. If there are eight entrants, the individual expected payoff is $\frac{5}{8}\frac{25€}{5}+\frac{3}{8}(-10€)\approx -0.63€$. Expected payoffs for all numbers of entrants are detailed in Table A1 (lower panel, first column).

Provided that players cannot coordinate, there is a mixed-strategy equilibrium in which each risk-neutral player (without private information about the probability of success on entry) enters with a probability *p*. The value of *p* is found by equating the expected payoff of entry and the payoff of staying out (see also Rapoport et al., 1998): $\sum_{E=1}^{5}\binom{15}{E}p^{E}(1-p)^{15-E}\frac{25€}{E}+\sum_{E=6}^{15}\binom{15}{E}p^{E}(1-p)^{15-E}\left[\frac{5}{E}\frac{25€}{5}+\left(1-\frac{5}{E}\right)(-10€)\right]=0$. In this game, each player enters with probability *p* of 53%. That is, 7.9 players (out of 15) will enter on average each round (Table A1, lower panel, first column).

If all players know their relative performance on the test, then clearly only the top five players (33% of all potential entrants) will enter. However, if players have imperfect information about their test performance, it is instructive to speculate how they might take account of competitors when assessing relative performance. Considerable evidence suggests that people tend focus on themselves and neglect others, thereby adopting a so-called “inside view” (Kahneman & Lovallo, 1993; Camerer & Lovallo, 1999; Kruger 1999; Moore, Oesch, & Zietsma, 2007). Thus, imperfect information about test performance could imply biased subjective estimates of probabilities of success. For example, assume that this bias is captured by a parameter α (-1> α >1) that adjusts the probability of success on entry. Then, a player’s (biased) expected payoff of entry when there are E entrants (E > 5) is $(1+\alpha)\frac{5}{E}\frac{25€}{5}+\left(1-(1+\alpha)\frac{5}{E}\right)(-10€)$. Table A1 provides equilibrium results for $\alpha \neq 0$. For example, if $\alpha = 0.2$, pure-strategy Nash equilibria occur when nine players (60%) enter the market, and if $\alpha = -0.2$, six players (40%) enter. In terms of mixed strategies, if $\alpha = 0.2$, the equilibrium probability of entry is 62% (9.3 entrants), and if $\alpha = -0.2$, it is 46% (6.9 entrants).

Table A1: Equilibrium analysis.

PURE-STRATEGY EQUILIBRIUM

	Expected payoff, €				
			alpha*		
Number of entrants	0	-0.4	-0.2	0.2	0.4
1	25.00	25.00	25.00	25.00	25.00
2	12.50	12.50	12.50	12.50	12.50
3	8.33	8.33	8.33	8.33	8.33
4	6.25	6.25	6.25	6.25	6.25
5	5.00	**5.00**	5.00	5.00	5.00
6	2.50	-2.50	**0.00**	5.00	7.50
7	**0.71**	-3.57	-1.43	2.86	5.00
8	-0.63	-4.38	-2.50	1.25	3.13
9	-1.67	-5.00	-3.33	**0.00**	1.67
10	-2.50	-5.50	-4.00	-1.00	**0.50**
11	-3.18	-5.91	-4.55	-1.82	-0.45
12	-3.75	-6.25	-5.00	-2.50	-1.25
13	-4.23	-6.54	-5.38	-3.08	-1.92
14	-4.64	-6.79	-5.71	-3.57	-2.50
15	-5.00	-7.00	-6.00	-4.00	-3.00

MIXED-STRATEGY EQUILIBRIUM

	Expected payoff, €				
			alpha		
Number of entrants	0	-0.4	-0.2	0.2	0.4
1	0.01	0.09	0.03	0.00	0.00
2	0.02	0.22	0.10	0.00	0.00
3	0.07	0.45	0.23	0.01	0.00
4	0.17	0.71	0.44	0.03	0.00
5	0.33	0.88	0.66	0.08	0.01
6	0.31	-0.51	0.00	0.22	0.06
7	0.13	-0.66	-0.29	0.27	0.12
8	-0.13	-0.57	-0.43	0.20	0.20
9	-0.30	-0.36	-0.38	0.00	0.21
10	-0.30	-0.17	-0.23	-0.20	0.10
11	-0.19	-0.06	-0.10	-0.28	-0.10
12	-0.09	-0.01	-0.03	-0.21	-0.24
13	-0.03	0.00	-0.01	-0.10	-0.22
14	0.00	0.00	0.00	-0.03	-0.11
15	0.00	0.00	0.00	0.00	-0.02
sum	0.00	0.00	0.00	0.00	0.00
Probability of entry	0.53	0.41	0.46	0.62	0.72
Number of entrants	**7.9**	**6.2**	**6.9**	**9.3**	**10.8**
Group profit**, €	-4.4	13.1	6.2	-18.5	-33.0

Notes:

* alpha is an adjustment coefficient of the subjective probability of success on entry.

**=25€ if 5 or fewer enter; 25€ -10€*(E-5) if E>5 enter.

Table A2: Logistic models of entry, by gender

	Model 1	
	Men	Women
Constant	1.61 (0.96)	1.31 (0.78)
Chance (0/1)	-0.10 (0.37)	-0.74 (0.42)
Rank	-0.26 (0.05) ***	-0.19 (0.05) ***
Rank*Chance	0.05 (0.04)	0.10 (0.04) **
Big chance (0/1)	0.24 (0.33)	-0.63 (0.39)
Order (0/1)	0.09 (0.34)	0.61 (0.38)
IE score	0.04 (0.04)	0.04 (0.05)
Forecasted number of entrants	0.10 (0.04) *	0.02 (0.03)
Indicator variables for Rounds 2-12	Included	Included
Model χ^2	75.1	39.5
Number of observations	684	756
Number of participants	57	63

Notes:

Population-averaged models were fitted. Standard errors, reported in parentheses, correct for correlation across repeated observations on individuals. *** p<0.001; ** p<0.01; * p<0.05.