IFN Working Paper No. 1149, 2017

Grind or Gamble? An Experimental Analysis of Effort and Spread Seeking in Contests

Ola Andersson, Håkan J. Holm and Erik Wengström

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Grind or Gamble?

An Experimental Analysis of Effort and Spread Seeking in Contests

November 2016

Ola Andersson, Håkan J. Holm and Erik Wengström*

Abstract: We conduct a contest experiment where participants can invest in increasing both the mean and the spread of an uncertain performance variable. Subjects are treated with different prize schemes and in accordance with theory we observe substantial investments in spread. We find that both types of investments can be controlled with a three level prize scheme. However, the control is imperfect and behavior is characterized by inertia. The winner-take-all prize scheme has many disadvantages including high spread and heterogeneous behavior. The scheme where only one loser is punished appears superior; it generates high mean, low spread and is most popular.

JEL Codes: C7; D8; D02; D03

Keywords: Contest; Risk; Spread; Incentives; Institutional Choice, Experiment

*Andersson: Research Institute of Industrial Economics (IFN), Lund University, ola.andersson@ifn.se; Holm: Lund University, Department of Economics: hj.holm@nek.lu.se Wengström: Lund University, Department of Economics and University of Copenhagen, Department of Economics, erik.wengstrom@nek.lu.se. We thank the Swedish Competition Authority for financial support and Claes Ek for helping us run the experiment. We are grateful to participants at the 8th CNEE Workshop, the 11th Nordic Conference on Behavioral and Experimental Economics in Oslo, the Contest: Theory and Evidence Conference in Norwich, 2016, and the seminar at Department of Banking and Finance at University of Innsbruck for helpful comments.

1. Introduction

A core question in economics is to provide people with appropriate incentives. It is therefore hardly surprising that many different incentive systems used in practice have been analyzed both theoretically and empirically. One such scheme entails using fixed payments based on rank, so called rank-order tournaments or contests for short. Such payments are practical in many cases, especially when the performance variable is difficult to translate to a cardinal scale and where ordinal assessments are easy. Lazaer and Rosen (1981) showed that payments based on rank leads to efficient choices under risk-neutrality, and also that participants may under certain circumstances prefer to be paid according to rank. The latter property is important whenever the participants in the contest (e.g., workers and managers) have a say in the decision of what incentive system to implement.

Most theoretical studies of contests have either focused solely on effort (see e.g., Nalebuff and Stiglitz 1983, Rosen, 1986 and Moldovanu and Sela 2001; see Konrad 2009 for a review) or solely on risk taking (see e.g., Dekel and Scotchmer, 1999, Tsetlin, Gaba, and Winkler, 2004) affecting the mean and the spread, respectively, of the rank determining performance variable. However, it has been convincingly argued by Hvide (2002) that in many important contest situations the contestants have a possibility to affect both the mean and the spread of the performance variable. For instance, CEOs can enter stable mature markets or unstable emerging markets and fund managers can choose a safe or risky portfolio. If this is the case and contestants at no cost can increase the variance of their performances, Hvide (2002) shows that contestants, facing contests that award the top ranking candidate, will end up in an equilibrium characterized by low mean and high spread.[1] Since it is reasonable to assume that the principal, i.e. the “contest organizer”, is positively affected by mean increases in the performance and negatively affected by the spread of the performance, this is a particularly bad equilibrium for the contest organizer.

In some of the literature, spread increases are somewhat loosely referred to as increases in risk without being precise of its meaning, in particular concerning who is exposed to the risk. To avoid confusion we interpret investments in mean as (productive) effort since it will increase the expected performance to the benefit of the contest organizer. Furthermore, investments in spread are interpreted as unproductive since it is costly and can be assumed to increase the risk the contest organizer faces without affecting the expected performance.[2] The distinction between spread of the rank-determining performance variable and the spread of the contestant’s payoffs is worth noting. Investments in mean increases of performance will normally affect the spread of the contestant’s payoff (and hence his risk) without necessarily affecting the variance of the contest organizer.

The observation by Hvide (2002) that the contestants seek spread at the cost of effort may be problematic for society as a whole when there are externalities (as in the banking sector), but as noted before, it will be especially costly for the contest organizer. Gilpatric (2009) shows how such spread seeking can be tamed. In a model with three or more contestants and where it is assumed to be costly to increase the spread, he shows that three payoff levels (a prize to

[1] In a somewhat similar model Kräkel and Sliwka (2004) analyzes a two-contestant tournament in which contestants differ in abilities and first choose a high or low-spread strategy and then choose effort (which is assumed to affect the mean). In this setting, they show that diverse equilibria are possible and depend on the magnitude of the ability difference, the shape of the cost function, and the prize spread.
[2] Note, investments in spread can also be seen as a form of effort, but since these investments are assumed to be costless in some of the literature (see e.g., Hvide, 2002) we reserve the term for investments affecting the mean.

the contestant ranked first, an intermediate prize, and a “loser prize” to the contestant ranked last) are sufficient to induce any combination of effort and spread under certain assumptions. The intuition is straightforward: Increasing spread (symmetrically) raises the probability of ending up first and last. Using standard prize schemes, the latter is not punished which distorts spread-choices upwards. By introducing a ‘’looser prize’’ for ending up last such incentives can be tamed.

One important difference in assumptions between Gilpatric and Hvide is that the former assumes strictly positive and convex cost of increasing the spread, whereas the latter assumes that increases in spread are free. We think that both assumptions are possible to defend in different empirical settings. Choosing a stock with a large spread instead of one with a low spread appears costless. However, if one assumes that projects possess a normal level of initial spread, then search theory would suggest that it is costly to find projects with the same mean but that have an unusually large spread. One example is the degree of originality in the design of new products. It is time consuming and costly come up with creative deviations from a standard design and it is not clear that consumers eventually consider these deviations to be improvements. On the other hand, a creative design that deviates from the standard product in many attributes that consumers appreciate will have a large market. Hence, increased originality of a new product is costly and increases the probability for both fiascos and best-sellers, which in turn generates a larger spread of returns.

While the predictions from theory by Gilpatric (2009) are promising in the sense that spread and productive effort choices can be carefully tailored, we know from various experiments that actual behavior does not always follow crisp equilibrium predictions. In the case with contest behavior one can think of various disturbing behavioral factors such as risk preferences, positional concerns beyond what is motivated by rank based payments and cognitive factors. We therefore conduct a contest experiment with a varying prize structure to investigate if the theoretical predictions get support. We also address the important but empirically open question regarding which prize structure the contestants prefer. In addition, our new design allows us to explore if different prize schemes generate a different level of behavioral heterogeneity when spread choices also are directly involved. Furthermore, the two-variable design makes it possible to study how the simultaneous effort and spread choices are connected.

We find clear evidence of investments in spread, which suggest that the concern raised by Hvide (2002) is motivated even in a setting where increasing spread is costly. Furthermore, both effort and spread seeking can be controlled to a certain extent with a three-level prize scheme as suggested by Gilpatric (2009). However, the observed behavior is characterized by inertia and the theoretically predicted treatment differences are starker than the observed ones. We also present results on how the prize schemes perform in this new environment of both investments in mean increases and spread. The prize scheme where only one loser is punished appears superior to the other schemes since it is associated with relatively high effort, low spread seeking and a low behavioral heterogeneity within the competing groups. Somewhat surprisingly, it is also the most popular scheme among the subjects. On the other side, winner-take-all appears to be the worst performing scheme with high spread, highly heterogeneous behavior and lowest popularity scores at the same time as it does not generate significantly higher efforts.

We also explore the results on how effort and spread choices are correlated both “within” the subject over time and “across” subjects. We find some evidence that subjects trade-off effort against spread, but this is happens only in one treatment (with two winners and one loser), which suggest that this effect is contingent on the prize structure. However, when we compare the correlation across subjects, we find robust evidence for a positive correlation. Hence, those who make the greatest productive efforts are also the ones most likely to make large investments in destructive spread. This ought to be an important lesson for any contest organizer who suspects that the contestants can affect the spread of their performance.

2. Related empirical studies

To our knowledge our study is the first experiment that manipulates the prize structure in a contest where contestants can choose both the mean and the spread of the performance variable. There is one study by Nieken (2010) where subjects in pairs choose a low or high variance distribution first and then effort. This study contains no treatment manipulation and can be considered a laboratory test of Hvide’s (2002) predictions, which partly get support in that when contestants have chosen high variance they exert less productive effort. At the same time, about 50 percent of the contestants do not end up in choosing the high variance distribution even after 27 rounds as they should in theory. This suggests that there is some heterogeneity in how contestants choose spread, and possibly also that all contestants do not fully understand the strategic upside of the high spread strategy. We take this further by investigating how spread choices can be controlled by treatment manipulations of the prize structure. This ought to be highly relevant from a risk-management perspective. We also elicit individual attributes, like personality attributes and cognitive measures to understand the heterogeneity better.

Investments by a contestant typically affect the probability distribution of the ranks the contestant ends up with, but will often also affect the contestant's expected payoff directly. Both effects are likely to affect the payoff distribution of the contestant and (depending on definitions) thus the risk he faces. Eriksen and Kvaløy (2016) focus on this risk when they study betting behavior in a lottery contest where the prize is partly contingent on the size of the bets (i.e., the investment) and where it is rational to bet zero. They find sizeable irrational risk taking in this lottery contest and that risk taking increases when feedback about the winner's strategy (in the previous round) is given and when the number of contestants increases. Eriksen and Kvaløy (2016) explain that competition "per se" triggers risk seeking even if it is irrational by referring to psychological mechanisms such as the "the contingency of reinforcement" by Skinner (1969) and the "availability heuristic" by Kahneman and Tversky (1973). Although, our contest and spread-seeking concept are different from the contest and risk-taking concept in Eriksen and Kvaløy (2016) it is not obvious why not the same mechanism should be triggered in our contest.[3] Hence, if spread seeking is mainly triggered by irrational psychological mechanisms as suggested by Eriksen and Kvalöy (2016) we think it is legitimate to ask if these mechanisms can be tamed (at least to some extent) with appropriately chosen prize schemes.

[3] To be fair there are also other differences in our contests. For instance, in Eriksen and Kvaløy (2016) any level of risk taking is excessive. In contrast, we allow for contests with both too high and too low spread seeking compared to equilibrium predictions.

In line with the theoretical studies, most contest experiments use a framework where contestants only can invest in effort (see the review by Dechenaux et al., 2015).[4] The first study to examine contests was clearly inspired by the theoretical findings by Lazear and Rosen (1981) and conducted by Bull et al. (1987). They found that tournament and piece-rate pay schemes generated the same mean effort, though the contest pay scheme induced a higher variance in effort. Observations like this one has motivated researchers to more carefully study the impact the prize structure has on effort (see e.g., Orrison et al., 2004; Harbring and Irlenbusch, 2008, Müller and Schotter, 2010, Sheremeta 2011, Shupp et al. 2013, Dutcher et al. 2015 and Andersson et al. 2016b). The results from these studies are somewhat mixed, but it is clear that manipulations of the prize structure matters and that the standard winner-takes-it-all (WTA) prize structure may discourage some contestants to exert effort. Our study can be seen as complementary to these since the choice of prize structure get somewhat more involved but also more realistic in cases when subjects also can affect the spread of their performance variable. One interesting empirical question that can be addressed in our study is if contestants who chose low productive effort also chose low unproductive spread and if this is associated with personal characteristics. It is also the case that prize schemes like WTA will in theory tempt contestants to increase the spread of their performance variable whereas other schemes will not. To learn more about this we need both between treatment manipulations of prize schemes and within-subject analyses, which this paper can contribute with.

[4] We have found one study where the contestants only choose risk and not effort. In this study (Nieken and Sliwka 2010) the correlation between the contestants performance variable is experimentally manipulated and found to matter for the contestants spread choices.

3. Hypotheses

The purpose of the experiment is to investigate if effort and spread choices can be controlled in contests when subjects choose both effort and risk. To accomplish this we investigate four different prize schemes with three contestants each. Each contestant $i \in \{1,2,3\}$ can make costly choices to increase the mean, $\mu_i \in [0,30]$, (i.e., effort) and spread, $\sigma_i \in [0,30]$, of their performance variable $Y_i = \mu_i + \varepsilon_i$ where $\varepsilon_i : U(-(50+\sigma_i), 50+\sigma_i)$ and iid.[5] The cost of choosing μ and σ are given by $C_\mu(x) = C_\sigma(x) = C(x) = x^2/20$.[6]

In all prize schemes the contestants compete for the same total prize sum (of 360 Danish crowns), but the schemes differ in what theory predicts about effort and spread seeking.[7] The first scheme is the winner-take-all scheme (WTA), which gives all money to one single winner. The second scheme is called single-loser (SL) since all but the loser divide the prize sum equally. The third and fourth schemes are called something-for-all since all contestants receive something but where the winner get the most and the contestant in the 2nd place get the average payoff and the contestant ending up last receives the least. The third and fourth prize schemes have either small or large differences in prizes, respectively and are denoted SFAS and SFAL. Each prize scheme represents a specific treatment and their exact prizes in Danish crowns are given in Table 1.

[5] I.e. when both μ and σ are set to zero the performance variable is uniformly distributed according to U(-50, 50).
[6] See Appendix B for details of the model and the predictions underpinning our hypothesis.
[7] The exchange rate between US dollars and Danish crowns was about 0.17 at the time of the experiment, which meant that the sum of prizes was equivalent to 61 US dollars.

Table 1. The treatments in terms of the different prize schemes and the theoretical predictions associated with each in terms of effort and spread.

Treatment	**WTA** Winner takes it all	**SL** Single loser	**SFAS** Something for small differences	**SFAL** Something for all large differences
Prizes	W1 =360 W2 = 0 W3 = 0	W1 = 180 W2 = 180 W3 = 0	W1 =150 W2 =120 W3 = 90	W1 = 210 W1 = 120 W1 = 30
Predicted effort (μ)	30	18	6	18
Predicted spread (σ)	10	0	0	0

The prize schemes are associated with various predictions if risk neutrality is assumed. These predictions are given in Table 1 and their derivations can be found in the Appendix. Our first hypotheses about effort are based on these predictions. In theory WTA is the scheme that would generate the highest effort; we therefore arrive at the following hypothesis:

Hypothesis 1: a) $\mu_{WTA} > \mu_{SL}$ *b)* $\mu_{WTA} > \mu_{SFAL}$ *c)* $\mu_{WTA} > \mu_{SFAS}$

It is also the case that the relative small differences in the prizes in SFAS generate weak incentives to exert costly efforts. We therefore also have the following hypothesis:

Hypothesis 2:a) $\mu_{SL} > \mu_{SFAS}$ *b)* $\mu_{SFAL} > \mu_{SFAS}$

Our theoretical predictions also vary with respect to the expected risk. All but the WTA scheme punishes the loser sufficiently much to make costly increases in spread unprofitable in expected terms. Hence, our third hypothesis is as follows:

Hypothesis 3: a) $\sigma_{WTA} > \sigma_{SL}$ *b)* $\sigma_{WTA} > \sigma_{SFAL}$ *c)* $\sigma_{WTA} > \sigma_{SFAS}$

We also elicit the popularity of the various schemes after the contestants have played all games. This can be regarded as our fifth prize scheme and is denoted CC (contest choice). In this treatment each contestant in a group ranked the four treatments after which one treatment was randomly drawn. Higher ranked treatment increased the probability that a treatment was drawn (see instructions in the Appendix). According to standard economic theory, risk neutral contestants who are interested in their own payoff should prefer the prize scheme associated with the lowest expected cost since the expected prize payoff is constant in the different treatments. If this view is accepted then it should be clear from Table 1 that SFAS is the least costly alternative, WTA the most costly and SL and SFAL equally costly. Our fourth hypothesis is therefore:

Hypothesis 4: a) SFAS is more popular than WTA, SL and SFAL. b) WTA is less popular than SL, SFAL.

Clearly, the hypotheses above guide our expectations about behavior in the various treatments. At the same time, we know that behavioral mechanisms and preferences can matter and cause deviations from the theoretical predictions. For instance, the popularity of the prize schemes may depend on social preferences. For instance, if inequity preferences according to Fehr and Schmidt (1999) are assumed, then one can expect that WTA is less popular than SL, since the payoff differences are smaller and envy is assumed be stronger than guilt.[8] In the similar vein since SFAS is more equal than SFAL, the former would be preferred under inequity aversion.

In addition to the predictions above we investigate determinants that may affect the contestants' choices. This part is more explorative then the previous one. Because, of the findings of Nieken (2010) and that the game is relatively complicated we hypothesize that it takes time for the contestants to realize how to balance effort and spread. Thus, we will control for the contestants' experience. We also think that cognitive factors may affect their choices. It has been shown that people with low cognitive ability make noisier choices which may affect how their behavior is interpreted (see Andersson et al., 2016a). We therefore elicit the subjects' cognitive reflection ability. It is also possible that individual personality characteristics affect the contestants' choices. For instance, it is possible that risk aversion affect the subjects' choices in the contest and preferences for a given prize scheme. It is also possible that achievement oriented personalities are more willing to exert effort than others. To investigate these issues the subjects participated in a risk elicitation task and answered a personality questionnaire.

4. Experimental design

The experiment was implemented using z-Tree (Fischbacher 2007) and consisted of 12 sessions conducted in the Laboratory for Experimental Economics (LEE) at University of Copenhagen during the fall semester of 2014. In total 237 subjects participated. Subjects were

[8] This follows from that in WTA (SL) two contestants (one contestant) end(s) up envying a substantial (smaller) inequality and one (two) winning contestant(s) feeling guilty about a large (smaller) inequality. Thus, SL gives less inequity and fewer contestants in the strong negative emotion of envying.

divided into 39 groups of 3, which were kept constant throughout the whole session (i.e., we used a partners' matching). Each group first played five trial rounds to learn how the contest worked. In order not to ''prime'' subjects with any particular prize scheme no payoff structure was implemented and the subjects in a group was simply informed about choices and rank outcomes. Subsequently they played five rounds of each of the four games, which differed with respect to prize scheme.[9] After these rounds the contestants played the CC treatment, where they ranked their three most preferred prize schemes. One subject's ranking in the group got randomly selected to be played out and a second random draw determined which of the three ranked prize schemes to be played out.[10] The scheme that got selected was subsequently implemented and played for five rounds. One round of the 25 was randomly chosen for payments. The sessions differed in respect of the order the groups played the four games. The four treatments are given in Table 2.

[9] Note, since the tournaments were played a finitely number of periods, which was clearly stated to the subjects the one-shot equilibrium predictions are not affected by the repetition, which follows from the standard backward induction argument.

[10] See Appendix A for exact rules and instructions on how this was played out. Most importantly subjects had incentives to be truth telling.

Table 2. Treatment design and organization of sessions.

	Session type			
	1	**2**	**3**	**4**
1st subsession	WTA	SL	SFAS	SFAL
2nd subsession	SFAS	SFAL	SL	WTA
3rd subsession	SFAL	SFAS	WTA	SL
4th subsession	SL	WTA	SFAL	SFAS
5th subsession	CC	CC	CC	CC
Number of sessions	3	3	3	3
Number of groups	20	21	20	18
Number of subjects	60	63	60	54

Each subject got extensive information about the game they should play. A key challenge was to explain to the subject that they could affect both the mean (interpreted as effort) and the spread of the distribution. We accomplished this by an illustrative example moving and changing an old canon. The subjects got the following information:

"Consider a contest against two other co-participants where you fire an old cannon and your place in the contest (1st, 2nd or 3rd) depends on how far your cannon ball gets before it hits the ground compared to the other co-participants. However, firing a cannon is associated with some uncertainty (due to factors that are difficult to control like the quality of gunpowder and the weather). You start with a cannon associated with a certain distribution (in the target zone). You may affect this distribution in two ways:

1. You can move the cannon forward and thereby increase the probability that the cannon ball reaches longer.

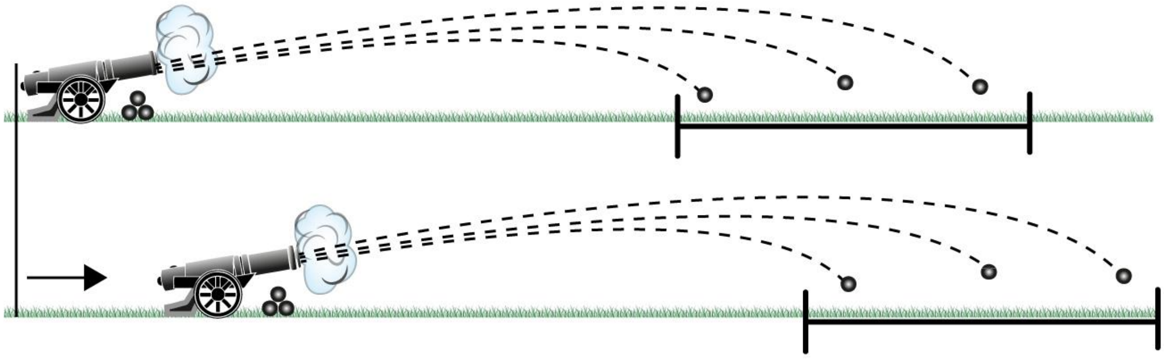


Figure 1: Illustration of increase in mean (effort).

2. You can increase the length of the cannon pipe. This makes it possible for the cannon ball to reach longer, but at the cost that the cannon gets more unstable and therefore also increases the probability for shorter shots. Hence, increasing the length of the cannon pipe increases both the probability for very long shots and very short shots.

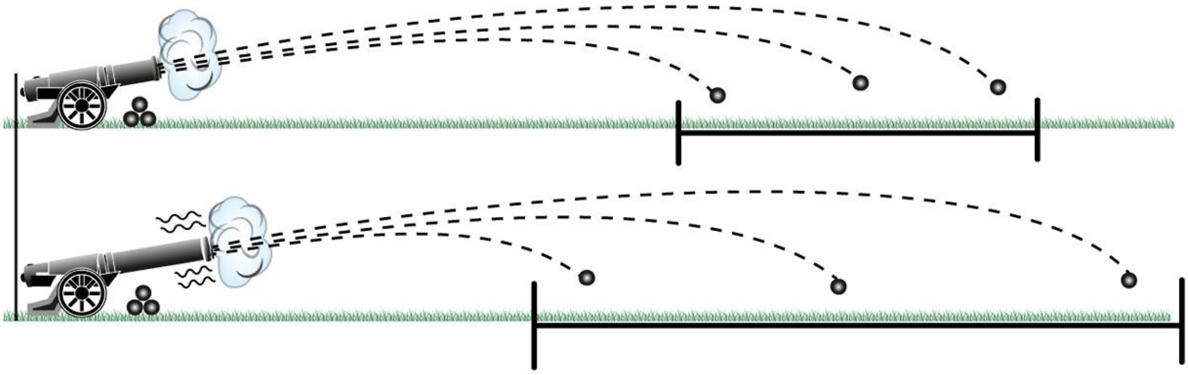


Figure 2: Illustration of increase in spread (spread).

The changes to the cannon—moving it forward or increasing the length of its pipe—is associated with a certain cost. Furthermore, in the contest your reward will only depend on your place in the contest and thus only indirectly through how far your shot is. These issues will be described in detail below."

After this information we also more exactly described the initial distribution, which was uniform, and how they could affect it by mean and spread increases and at what cost.[11] The

[11] See the appendix for instructions.

subjects also got to answer control questions to check that they understood the game. The subjects then played the trial rounds and the 25 real rounds. After this we had one incentivized risk-elicitation task and a dictator game, the reason for this was to control for risk preferences and social preferences that might impact behavior.[12] Once all incentivized tasks were done, the subjects learned their earnings and answered background questions about gender, age, studies etc. They were also asked a number of cognitive reflection questions (CRT), which we extended by two new questions since we were worried that some students already knew the ones used by Frederick (2005). Finally, we asked 17 personal value questions using the Personal Value Questionnaire (PVQ) (Schwartz 1992). These were chosen to elicit value orientation toward universalism, benevolence, achievement and power, since we believed that these were characteristics that may influence play and also the prize scheme preferences. The whole sequence of tasks and information in a session is given in Table 3 below.

[12] It should be noted that due to a mistake some sessions the risk elicitation was not incentivized. Since the risk elicitation was after the tournament and the dictator game it should not affect behavior.

Table 3. Sequence tasks and information in the sessions

Instructions
Control Questions
Trial Rounds (5 Rounds)
Main Treatments (25 Rounds)
Dictator Game
Risk Task
Profit Display
Background questions
Cognitive reflection
PVQ

5. Results

In this section we will present the experimental results. We start with some descriptive statistics. In total 237 subjects took part in the experiment. Average age was 26 years and 46 percent female. The experiment took on average about 2 hours and the average earnings was 242 Danish crowns, which is more than this group of subjects would earn if they would spend the same time on a typical job in Denmark.

5.1. Average behavior

We will first present results on average behavior and discuss them with respect to the hypotheses. The robustness of the results will then be checked using regressions with control variables. In Table 4 we present the average contest behavior in the various treatments over all rounds.

Table 4. Predictions and observed average behavior in the various contests.

Treatment	**WTA**	**SL**	**SFAS**	**SFAL**
Predicted Effort (μ)	30	18	6	18
Predicted Spread (σ)	10	0	0	0
Average Effort (μ)	17.5	17.1	13.8	17.3
Average Spread (σ)	5.5	3.8	3.9	4.8

Our hypotheses are however not expressed in terms of point predictions but in terms of differences between different prize schemes. To make a first formal test of our hypotheses we employ the Wilcoxon matched pairs test using the average group effort choice as the unit of observation (i.e. the effort choices of all subjects across all five periods in a given contest are collapsed into one value).[13] Hypothesis 1a and1b are rejected since the effort in WTA is not significantly larger than the ones for SL (two-sided *p*-value = 0.546) and SFAL (two-sided *p*-value = 0.942). Hypothesis 1c is confirmed since the effort is significantly larger in WTA than in SFAS (two-sided *p*-value < 0.001). Hypothesis 2a and 2b are confirmed since effort is significantly smaller in SFAS compared to SFAL (two-sided *p*-value < 0.001) and SFAS compared to SL (two-sided *p*-value < 0.001). An interpretation is that the subjects consider the incentives induced by the prize structure in WTA, SL, SFAL sufficiently similar to give

[13] The conclusions are similar if we instead restrict attention to performances in the first subsession and compare behavior between subjects using the Mann-Whitney U-test. The only difference is that we cannot confirm hypothesis 2b since mean increases are not significantly different between SFAS and SL in the first period. See Online Appendix X for details.

rise to the same effort but that there is a threshold were the prize structure gets too "flat" to make effort worthwhile and SFAS is below this threshold.

Hypothesis 3 predicts that WTA will be associated with higher spread than all the other prize schemes. This is confirmed by the data, where the spread in WTA is significantly higher than the one in SL (two-sided p-value < 0.001), SFAS (two-sided p-value < 0.001) and SFAL (two-sided p-value = 0.024).[14]

At this point it is worthwhile to make some general remarks about the results on average effort and spread. From Table 4 it is clear that the average observed effort and spread choices differ from the theoretical point predictions. This is not especially strange since we made demanding assumptions when deriving the predictions (e.g., full rationality, risk neutrality, selfish preferences). Even if the point prediction on effort in SFAL and SL are very close to average observed behavior, it is difficult to argue that this is because our model closely captures behavior. In particular, there are large observed individual deviations from the point predictions in both directions in WTA and SFAS. Hence, while the effort predictions are reasonably correct for SFAL and SL they overestimate WTA and underestimate SFAS. In the case of spread increases we find that our model overestimates them for WTA and underestimates them for the others, the latter is quite natural since the point predictions are zero, which means that any noise would lead to underestimations. The overall impression is that the observed behavior is characterized by inertia and that it is less responsive to treatment differences than what the model predicts.

[14] Except the comparison between WTA and SFAL, these conclusions are confirmed if we instead look at between subject differences in the first round of the experiment using the MWU-test. See Online Appendix for details.

Hypothesis 4 concerns the popularity of the schemes and the first prediction (4a) is that SFAS is the most popular scheme. This turns out to be wrong. From Table 5 it should be clear that SL is most popular independent if we rank according to number of times it is “ranked 1” or ranked “top 3”. Hypothesis 4b suggests that WTA is less popular than SL and SFAL. It is true that WTA is the least popular scheme but the differences between WTA and SL and SFAF are not significant (the difference between SL and WTA is borderline significant using a binomial test of equal proportions, two-sided *p*-value = 0.120).

Table 5. Popularity of the various prize schemes according to how often they were ranked 1st, 2nd and 3rd.

	% Rank 1	**% Rank 2**	**% Rank 3**	**# Total top 3**
WTA	21.5	21.94	29.54	173
SL	29.1	24.89	24.47	186
SFAS	23.2	28.27	22.78	176
SFAL	26.2	24.89	23.21	176

Thus far we have only looked at average behavior without respect to behavioral changes over time. However, when scrutinizing the data it becomes clear that time matters and especially so when behavior initially deviates from equilibrium predictions. When there is a notable change over time this change appears to be in the direction of the theoretical equilibrium prediction. For instance, in Figure 3 the only notable change in effort over time is in SFAS, where contestants on average initially deviate substantially from the equilibrium prediciton of 6. However, during the 5 rounds the average effort decreases markedly in line with the equilibrium prediction. If we look at spread choices we can see that these drop somewhat for all treatments except for WTA, which implies that the spread increases go in the direction of the equilibrium predictions for these prize schemes. Hence, a tentative conclusion from this is that learning takes place during these rounds and that experienced subjects are closer to the equilibrium predictions than unexperienced. In the regressions below we will therefore control for experience effects.

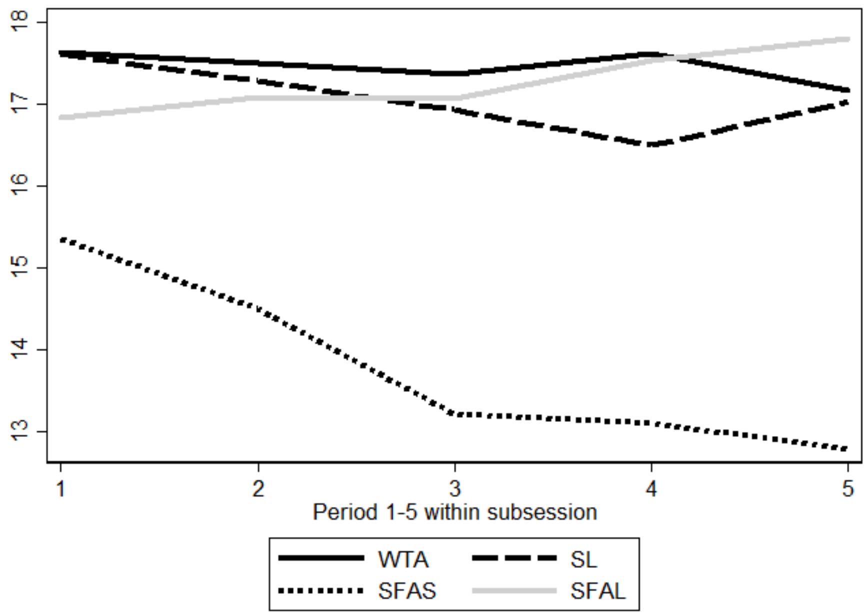


Figure 3: Average effort choice over time (measured by the average means chosen in the various rounds within the treatment subsessions).

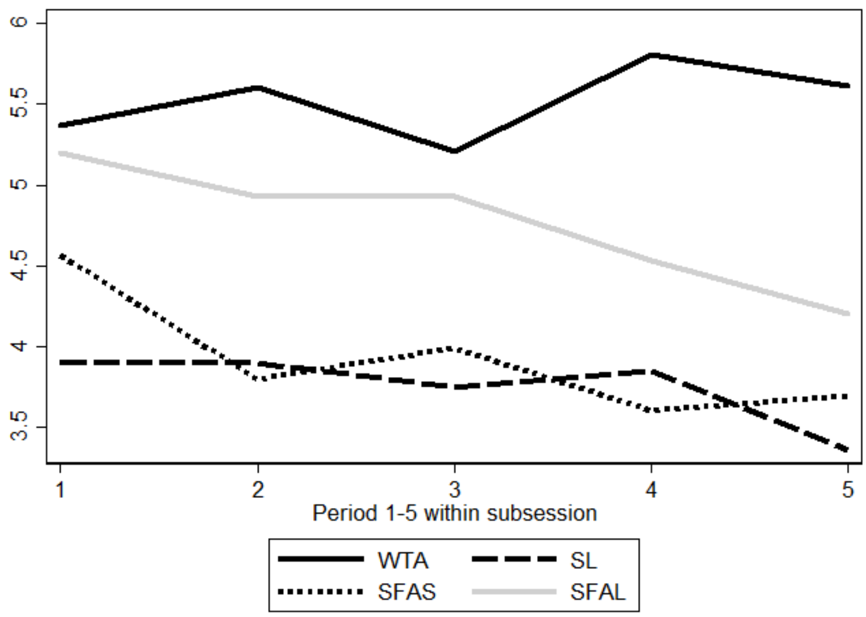


Figure 4: Average spread choice over time (measured by the average spreads chosen in the various rounds within the treatment subsessions).

5.2 Regression analyses

In this section we will inspect if our previous test results of our hypotheses are robust when we control for various factors. We will then analyze individual determinants of behavior. We start by summarizing our control variables in Table 6. The first two are self-explanatory. CRT gives the number of correct answers to the CRT questions, Risk choice the switch point (later switch indicates higher degree of risk aversion), Watch sport measures the number of hours spent watching sport each week and Play sport the number of hours practicing some sport. Danish is an indicator function for being a Danish citizen. Social media measures number of minutes per week spent on social media platforms (e.g Facebook, Instagram). Dictator offer reports the offer made to an anonymous recipient in the dictator game. The last four variables are the four most prominent personality measures from the PVQ (Schwartz 1992).

Table 6.Summay of regression variables.

	Mean	SD	Min	max
Gender	0.456	0.498	0	1
Age	25.658	4.957	19	68
Risk choice (1-8)	3.945	2.067	1	8
CRT (0-5)	3.228	1.580	0	5
Watch sport (hours/week)	2.363	4.581	0	30
Play sport (hours/week)	9.460	8.394	0	50
Danish	0.278	0.448	0	1
Social media (min/week)	98.793	111.508	0	1000
Dictator offer (0-500 DKK)	117.743	129.690	0	500
Universalism	2.326	0.818	1	6
Benevolence	2.426	0.738	1	5
Achivement	2.884	1.070	1	6
Power	3.782	1.000	1	6

5.2.1 Effort

Table 7 reports results from a series of linear random effects regressions with Effort as dependent variable. We let each individual constitute a panel and cluster standard errors on the group level. WTA is used as the baseline treatment dummy. To capture an order effect of the different treatments the Subsession variables indicate in which subsession the decision was made (with the first subsession being the baseline). The Period variable captures the round within the subsession. Data from the last subsession (CC treatment) is excluded from the regressions.

In line with our previous findings we can only confirm Hypothesis 1c as only the coefficient on SFAS is negative and significant. A chi-square test on the most general specification (model 4) reveals that coefficients on SFAS and SL are significantly different (*p*-value < 0.001), confirming Hypothesis 2a, and the coefficients on SFAS and SFAL are different (p-value < 0.001) confirming Hypothesis 2b. In addition, it can be noted that female subjects exert significantly less effort. In terms of personality measures, Dictator offer and Benevolence (which to some extent measures pro-socialness) adds positively to effort choices whereas Achievement leads to lower effort choice. Finally, it can be noted that experience as measured by Subsession is not significant. Thus effort does not change much between the sessions, which suggests that learning does not substantially impact the effort decisions in this respect. However, as Figure 3 indicates the experience effect within subsessions appears treatment contingent. This is also confirmed statistically by a significant effect if we interact subsession with SFAS.

Table 7: Regression, Effort, Subsession 1-4

	(1)	(2)	(3)	(4)
SL	-0.389	-0.380	-0.380	-0.380
	[0.589]	[0.580]	[0.580]	[0.580]
SFAS	-3.673***	-3.661***	-3.661***	-3.661***
	[0.580]	[0.562]	[0.562]	[0.563]
SFAL	-0.191	-0.152	-0.152	-0.152
	[0.502]	[0.511]	[0.512]	[0.512]
Subsession 2		-0.128	-0.128	-0.128
		[0.508]	[0.508]	[0.508]
Subsession 3		0.00849	0.00849	0.00849
		[0.607]	[0.607]	[0.607]
Subsession 4		0.972	0.972	0.972
		[0.600]	[0.600]	[0.600]
Period (1-5)		-0.173**	-0.173**	-0.173**
		[0.0774]	[0.0774]	[0.0775]
Gender		-2.182**	-1.945**	-2.068**
		[0.930]	[0.964]	[1.004]
Age		0.0587	0.0786	0.0744
		[0.0805]	[0.0801]	[0.0748]
Riskchoice			0.232	0.310
			[0.239]	[0.248]
CRT			0.304	0.238
			[0.253]	[0.297]
Dictator offer			0.00551*	0.00527*
			[0.00322]	[0.00314]
Additional Controls				X
Constant	17.46***	17.24***	14.08***	17.56***
	[0.721]	[2.530]	[3.290]	[3.657]
Observations	4,740	4,740	4,740	4,740
Number of id	237	237	237	237
N_clust	79	79	79	79

Notes: Linear random effects regressions. Robust standard errors clustered at the group level in brackets. WTA is used as the baseline treatment and SL, SFAS, SFAL are treatment dummies. The Subsession 2-4 variables indicate in which subsession the decision was made (with the first subsession being the baseline). The Period variable captures the round within the subsession. Riskchoice indicates the switch point in the risk elicitation task with a higher number indicating a higher degree of risk aversion. CRT gives the number of correct answers to the CRT questions. Dictator offer reports the offer made to an anonymous recipient in the dictator game. Additional Controls contains the variables Watchsport, Playsport, Danish, Socialmedia, Universalism, Benevolence, Achivement, Power.

*** p<0.01, ** p<0.05, * p<0.1

5.2.2 Spread choices

Table 8 reports results from a series of linear random effects regressions with Spread as dependent variable, using each individual as a panel and clustering standard errors at the group level.

All treatment coefficients are negative and statistically different from zero indicating that spread increases were larger in WTA compared to all other treatments. This finding confirms Hypotheses 3a-c. It is also interesting to note that there are significant experience effects on spread seeking (as can be seen from the coefficients on the Subsession and Period variables). The more experienced, the less inclined to invest in spread. One interpretation of this is that in three of the four contests the expected returns on spread investments are negative and that subjects learn this the more sessions they play. We also think that that the negative correlation between cognitive reflection ability (CRT) and spread seeking also speaks in this direction. People who are more reflective realize the negative return on spread investments in all prize schemes except for WTA. If we include interactions between CRT and contest types in the regressions, we observe that the negative relation between CRT and spread seeking is found in all contests except WTA. That is, subjects with higher CRT scores are less likely to seek spread in the contests where spread seeking is not sustained in equilibrium.

Table 8: Regression, Spread Increase, Subsessions 1-4

	(1)	(2)	(3)	(4)
SL	-1.775***	-1.789***	-1.789***	-1.789***
	[0.428]	[0.394]	[0.394]	[0.394]
SFAS	-1.595***	-1.573***	-1.573***	-1.573***
	[0.431]	[0.411]	[0.411]	[0.411]
SFAL	-0.764**	-0.794**	-0.794**	-0.794**
	[0.355]	[0.348]	[0.348]	[0.348]
Subsession 2		-0.824**	-0.824**	-0.824**
		[0.354]	[0.354]	[0.354]
Subsession 3		-1.209***	-1.209***	-1.209***
		[0.365]	[0.365]	[0.366]
Subsession 4		-1.885***	-1.885***	-1.885***
		[0.384]	[0.384]	[0.385]
Period (1-5)		-0.120*	-0.120*	-0.120*
		[0.0644]	[0.0644]	[0.0645]
Gender		-0.473	-0.830	-0.733
		[0.750]	[0.683]	[0.629]
Age		0.0974	0.0615	0.0819
		[0.0837]	[0.0743]	[0.0800]
Riskchoice			0.0330	0.0861
			[0.170]	[0.175]
CRT			-0.630***	-0.515**
			[0.213]	[0.247]
Dictator Offer			0.00159	0.00132
			[0.00215]	[0.00228]
Additional Controls				X
Constant	5.522***	4.584**	7.382***	4.896*
	[0.581]	[2.142]	[2.403]	[2.973]
Observations	4,740	4,740	4,740	4,740
Number of id	237	237	237	237
N_clust	79	79	79	79

Notes: Linear random effects regressions. Robust standard errors clustered at the group level in brackets. WTA is used as the baseline treatment and SL, SFAS, SFAL are treatment dummies. The Subsession 2-4 variables indicate in which subsession the decision was made (with the first subsession being the baseline). The Period variable captures the round within the subsession. Riskchoice indicates the switch point in the risk elicitation task with a higher number indicating a higher degree of risk aversion. CRT gives the number of correct answers to the CRT questions. Dictator offer reports the offer made to an anonymous recipient in the dictator game. Additional Controls contains the variables Watchsport, Playsport, Danish, Socialmedia, Universalism, Benevolence, Achivement, Power.

*** p<0.01, ** p<0.05, * p<0.1

5.3 Popularity

It is also of some interest to study determinants behind the popularity of different prize schemes. In Appendix C we estimate a multinomial logit regression model to study this.

However, the only interesting convincing significant effect is that subjects who donate more in the dictator game are less likely to rank WTA at the top. This makes sense if dictator donations reflect distributional preferences since WTA generates the most unequal payoff distribution. Except for this effect there are no strong convincing determinants of the subjects' rankings of the prize schemes.

5.4. Other effects

Since this is the first experiment where subjects can choose both effort and spread of the performance variable under different treatment we think it is worthwhile to dig somewhat deeper in our data and explore some additional effects. To start, we will analyze how the different prize schemes affect the heterogeneity of the subjects' effort and spread choices. Secondly, we will analyze the correlation between effort and spread choices.

5.4.1. Variance of effort and spread

An important property of a prize scheme is the degree of heterogeneity of individual behavior it induces. One early observation in the experimental contest literature is that rank-based contests induce a high variance of efforts and thus performance (Bull et al. 1987).[15] Because, the contestants in the current study can affect the distribution of the performance variable both

[15] This observation has been followed by research to understand the sources of this variance and how it can be mitigated. For instance, Vandergrift and Brown (2003) show that high variance strategies may be triggered by a combination individual capabilities and difficulty of the task and Eriksson et al. (2009) demonstrate that variance in effort can be mitigated by allowing contestant to self-select into tournaments. See Sheremeta (2013) for a survey on heterogeneity of individual behavior in contest experiments.

by productive effort and spread seeking we can provide novel data on how these choices are made with different prize schemes. Table 9 displays the standard deviation of investments in productive effort and spread across prize schemes. The figures are obtained by first calculating the standard deviation of productive effort (spread) for each group and prize scheme and then take the average standard deviation over all groups. We denote this measure the “within-group” variance. Given that the contest organizer has a payoff which is positively correlated with the individuals’ performances then the expected variance of this payoff will be increasing in variance of the within group productive effort and spread. For obvious reasons the expected utility will be decreasing in this expected variance for any risk avert contest organizer, which means that these variances can be seen as additional negative side-effects of the prize schemes.

WTA has the highest standard deviation for investments in Effort and SL the lowest within groups. Making pairwise comparisons between prize schemes using the Wilcoxon matched pairs test, we observe that within groups, Efforts in WTA have a higher standard deviation than the other prize schemes (two-sided p-values ≤ 0.001). SL is also associated with lower spread than SFAS (two-sided p-value = 0.023), but there are no differences between SFAL and SL, or between SFAL and SFAS (two-sided p-values > 0.2).

For Spread, WTA also displays the highest standard deviation and SL the lowest. However, the difference between SL and WTA is only weakly significant (two-sided p-value == 0.061) and the other differences are not statistically significant at the 10 % level.

Table 9. Standard deviations of Efforts and Spread

	WTA	SL	SFAS	SFAL
Within Groups				
Mean SD of Effort	8.313	6.138	6.982	6.548
Mean SD of Spread increase	5.054	4.222	4.474	4.656
Between Groups				
SD of Mean Efforts	6.410	7.350	5.700	6.511
SD of Mean Spread increase	5.159	3.958	3.540	4.843

If we use the group means as our unit of observation, we get an indication how groups differ from each other, which is of relevance for a contest organizer whose payoff is contingent on the average performance of the group and where there are more than one group. [16] Although, it is less straightforward one can presume that the owner is negatively affected by a high between group variance. At least, the predictability of the groups' average performance will be negatively affected by a high between group variance. SL has the highest standard deviation in Effort, but it is only significantly higher than SFAS (using a two-sided F-test (*p*-value = 0.026). Regarding Spread, the standard deviation across groups is larger for WTA and SFAL than the other prize schemes.[17]

The most notable conclusion from this analysis is that WTA appears to have the highest variance in all measures except for the between group measure in Effort. SL generates lowest variance of Effort and Spread in the within group measure, but the prize scheme generates higher variance in the between group measures.

[16] To see that it is possible to have a low within group variance and high between group variance consider a prize scheme that induces contestants to imitate each other so that they will converge to either high or low Effort (or Risk). This will lead to "path dependent" choices so that contingent on the choices in the first round the group may end up at high or low effort, which means that the within group variance is low but the between group variance will be high. An example of the reverse case would be a scheme that induces different contestants to act differently but constant over time.

[17] Two-sided F-tests: WTA vs. SL: *p*-value = 0.020; WTA vs SFAS: *p*-value = 0.001; SFAL vs SL: *p*-value = 0.077; SL vs SFAS: *p*-value = 0.006; all other pairwise comparisons are insignificant (*p*-values > 0.3).

5.4.2. Correlation between Effort and Spread

Another novelty in our study is the relationship between effort and spread choices. Table 10 displays estimates from panel data fixed effects regression of Effort when Spread is included as explanatory variable. The estimates in the first column are based on the full sample and the other columns are based on data from each treatment separately. There is a significant negative relationship between mean and spread in SL, which suggest that for a given individual there is a substitution between effort and spread as suggested by Hvide (2002) and observed by Nieken (2010). However, our results demonstrate that this effect appears to be contingent on the given prize scheme.

Table 10. Panel data fixed effects regressions of Efforts in all treatments and for each treatment.

	All treatments	WTA	SL	SFAS	SFAL
Spread Increase	-0.0585	-0.0543	-0.220**	0.0449	-0.0848
	[0.0459]	[0.0924]	[0.110]	[0.0868]	[0.0956]
SL	-0.485				
	[0.579]				
SFAS	-3.753***				
	[0.553]				
SFAL	-0.199				
	[0.511]				
Subsession 2	-0.176				
	[0.511]				
Subsession 3	-0.0623				
	[0.600]				
Subsession 4	0.862				
	[0.591]				
Period (1-5)	-0.180**	-0.0773	-0.220	-0.647***	0.218
	[0.0772]	[0.189]	[0.135]	[0.156]	[0.134]
Constant	18.16***	18.00***	18.56***	15.55***	17.02***
	[0.637]	[0.778]	[0.553]	[0.551]	[0.568]
Observations	4,740	1,185	1,185	1,185	1,185
R-squared	0.051	0.002	0.033	0.034	0.009
Number of id	237	237	237	237	237
N_clust	79	79	79	79	79

Notes: Robust standard errors clustered at the group level in brackets. WTA is used as the baseline treatment and SL, SFAS, SFAL are treatment dummies. The Subsession 2-4 variables indicate in which subsession the decision was made (with the first subsession being the baseline). The Period variable captures the round within the subsession.
*** p<0.01, ** p<0.05, * p<0.1

At the same time and perhaps even more interesting is that we find a strong positive correlation between subjects so that subjects who invest in Effort are also more likely to invest in Spread. Using individual level averages of Effort and Spread across all treatments, the Pearson's correlation coefficient is 0.29 (*p*-value = 0.000). This positive correlation is at least marginally significant in all treatments separately ($r_{WTA} = 0.24$; p-value = 0.000, $r_{SL} = 0.11$; p-value = 0.085, $r_{SFAS} = 0.28$; p-value = 0.000, $r_{SFAL} = 0.13$; p-value = 0.049). This has the important implication that those subjects who are likely to be most productive in providing high effort are also the ones that are most inclined to the more destructive activity of deliberately increasing the spread. This suggests that investments in effort and spread stem from a common motive, like a willingness to win.

6. Conclusion

We present results from a novel experimental design where contestants directly can choose both the mean (Effort) and variance (Spread) of the rank determining performance variable in a contest. By manipulating the prize schemes so that they yield different theoretical predictions in terms of effort and spread seeking we study if behavior follows theory. In line with the predictions we find clear evidence of investments in spread. Hence, the concern raised by Hvide (2002) that contestants may engage in activities that will increase the variance of the performance variable, possibly at the cost of effort, is motivated even in a setting where increasing variance is costly. As a consequence, even if investment in spread is totally wasteful at an aggregate level and possibly harmful to the contest organizer, it is a phenomenon that will turn up under certain conditions. This is the bad news. The good news is that both effort and spread choices can indeed be controlled to a certain extent with a three-level prize scheme as suggested by Gilpatric (2009). However, the observed behavior is characterized by inertia and the theoretically predicted treatment differences are starker than the observed ones.

We also present results on how individuals combine spread and effort under different prize schemes and also how these perform in this new environment of both effort and spread choices. At the individual level we find statistically significant evidence of substitution only for one prize scheme, which suggest that this effect is contingent the prize structure. Furthermore, there is relatively strong positive correlation between investments in effort and spread between subjects. Hence, people who on average choose high effort are also the ones choosing high spread. One interpretation of this is that there is one common denominator between these choices like a willingness to compete and win.

When it comes to the relative performances of the prize schemes we find that the scheme where only one loser is punished appears superior to the other schemes since it is associated with relatively high effort, low spread increase and a low behavioral heterogeneity within the competing groups. In addition, it is also the most popular scheme among the subjects. The scheme that probably most people associate with a contest, winner-take-all appears to be the worst performing scheme with high risk, highly heterogeneous behavior and lowest popularity scores at the same time as it does not generate significantly higher efforts. The important lesson from these observations is that contest owners need to be very careful when they design contests in situations where contestants can affect both the mean and the variance of their performance. Our findings suggest that contest theory is an important guide to understand behavior in these situations. However, these insights need to be complemented by carefully designed studies of human behavior.

7. References

Andersson Ola, Holm, Håkan, Tyran, Jean-Robert, and Wengström, Erik. (2016a) Risk aversion relates to cognitive ability: preferences or noise? *Journal of the European Economic Association*, 14, 1129–1154

Andersson, Ola, Huysentruyt, Marieke, Miettinen, Topi, and Stephan, Ute. (2016b). Person–Organization Fit and Incentives: A Causal Test. *forthcoming in Management Science.*

Bull, C., Schotter, A. and Weigelt, K. (1987). Tournaments and piece rates: an experimental study. *Journal of Political Economy*, 95, 1-33.

Croson, Rachel and James Sundali (2005). The gambler's fallacy and the hot hand: empirical data from casinos" *The Journal of Risk and Uncertainty*, 30:3, 195–209.

Dechenaux, Emmanuel, Kovenock, Dan and Sheremeta, Roman. (2015). A survey of experimental research on contests, all-pay auctions and tournaments, *Experimental Economics* 18 (4), 609-669.

Dekel, E., and S. Scotchmer. (1999). On the evolution of attitudes towards risk in winner-take-all games. *Journal of Economic Theory* 87, 125–43.

Dutcher, Glenn, Balfoutas, Loukas, Lindner, Florian, Ryvkin, Dmitry and Sutter Matthias Strive to be first or avoid being last: an experiment on relative performance incentives. *Games and Economic Behavior* 94, 39-56.

Eriksen, Kristoffer, W., and Ola Kvaløy, (2016) No guts, no glory: an experiment on excessive risk-taking, forthcoming in the *Review of Finance*.

Eriksson, Tor, Sabrina Teyssier, and Marie Claire Villeval. (2009) Self-selection and the efficiency of tournaments. *Economic Inquiry*, Wiley, 2009, 47 (3), 530-548.

Fehr, E. and Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114, 817-868.

Fischbacher, Urs. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, 10(2), 171-178.

Frederick, Shane. (2005). "Cognitive reflection and decision making." *Journal of Economic Perspectives* 19 (4): 25–42.

Gilpatric, Scott M. (2009). Risk taking in contests and the role of carrots and sticks. *Economic Inquiry* 47 (2), 266-277.

Harbring, C. and Irlenbusch, B. (2008). How many winners are good to have? On tournaments with sabotage. *Journal of Economic Behavior and Organization* 65, 682-702.

Hvide, Hans K., (2002). Tournament rewards and risk taking. *Journal of Labor Economics* 20 (4), 877-898.

Konrad, K.A. (2009). *Strategy and Dynamics in Contests*. New York, NY: Oxford University Press.

Kräkel, M., and D. Sliwka. (2004) Risk Taking in asymmetric tournaments. *German Economic Review* 5, 103–16.

Lazear, E.P. and Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, 89, 841-864.

Moldovanu, B. and Sela, A. (2001). The optimal allocation of prizes in contests. *American Economic Review* 91, 542-558.

Müller, Wieland and Schotter, Andrew. (2010). Workaholics and dropouts in organizations. *Journal of the European Economic Association* 8, 717-743.

Nalebuff, B., and J. Stiglitz. ((1983), Prizes and incentives: towards a general theory of compensation and competition. *The Bell Journal of Economics* 14, 21–43.

Nieken, P. (2010). On the choice of risk and effort in tournaments-experimental evidence. *Journal of Economics and Management Strategy* 19, 3, 811-840.

Nieken, P. and Sliwka, D. (2010). Risk-taking tournaments – theory and experimental evidence. *Journal of Economic Psychology* 31, 254–268.

Orrison, A., Schotter, A. and Weigelt, K. (2004). Multiperson tournaments: an experimental examination. *Management Science* 50, 268-79.

Rosen, S. (1986) Prizes and incentives in elimination tournaments. *American Economic Review* 26, 701–15.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theory and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25) (pp. 1-65). New York: Academic Press.

Sheremeta, R.M. (2011). Contest design: an experimental investigation. *Economic Inquiry* 49, 573-590.

Sheremeta, Roman. (2013) Overbidding and heterogeneous behavior in contest experiments, Journal of Economic Surveys 27, 499-514.

Shupp, Robert, Sheremeta, Roman, Schmidt, David, and Walker, James. (2013). Resource allocation contests: Experimental evidence, *Journal of Economic Psychology* 39, 257-267.

Skinner, B. F. (1969) Contingencies of reinforcement: a theoretical analysis, Prentice Hall, Englewood Cliffs, New Jersey.

Tsetlin, I., A. Gaba, and R. Winkler. (2004) Strategic choice of variability in multiround contests and contests with handicaps. *Journal of Risk and Uncertainty* 29,143–158.

Tversky, A. and Kahneman, D. (1973) Availability: a heuristic for judging frequency and probability. Cognitive Psychology 5, 207–232.

Vandegrift, Donald and Paul M. Brown. (2003) "Task difficulty, incentive effects, and the selection of high - variance strategies: an experimental examination of tournament behavior." Labour Economics, 10(4) , 2003, 481-97.

Appendix A. Instructions

A1. General Instructions

We here present detailed instructions and screen shots from the experiment. In the experiment we first gave some general information about the contest to be played, which was handed out in paper form to the subjects. The content of those instructions are displayed (in italics here for presentational purposes) below.

General Information

The purpose of this study is to gain insights into economic behavior. You will make choices in different situations that will be explained later.

The experiment consists of three main parts, in which you can earn money, and a questionnaire. All amounts stated in the experiment are in Danish DKK (DKK) and your earnings will be paid privately in cash at the end of the experiment. The amount of money you earn will depend on the choices you and the other participants make. You start out with a show up fee of DKK 90. You will then be given the opportunity to earn more money but you may also lose some of the show up fee.

The possibility to earn real money is important in economic experiments and there are strict rules against deceiving participants. Hence, all information given here about money and other aspects are true and will be carried out according to the information given. Please, note also that there are no "right" or "wrong" choices in the decisions you are going to make. Therefore, make decisions according to what you prefer. Your choices and the choices of other participants will remain anonymous. The choices will be used for research purposes only, and will be kept strictly confidential.

On the next pages of this document you will find instructions for the first part of the experiment. The instructions of the succeeding parts will be shown to you on the computer screen.

Please, read the instructions carefully. If there is anything you do not understand, please raise your hand and one of the Experimenters will come to your booth.

Instruction Part 1

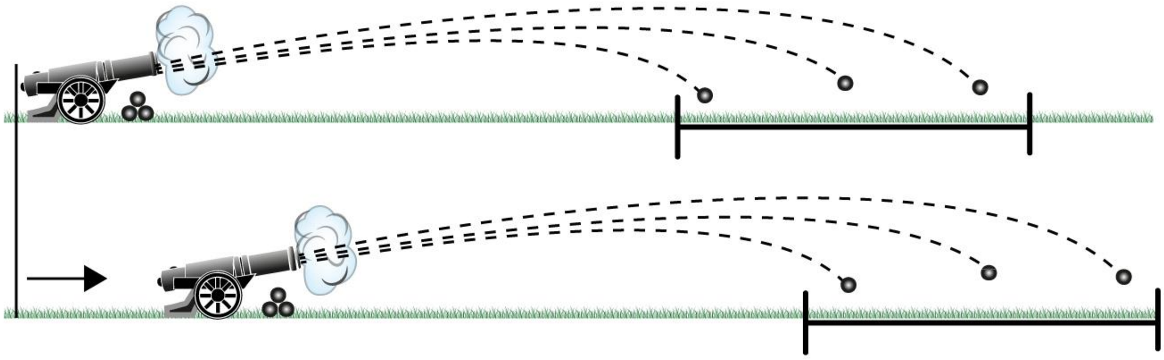
The first part of the experiment consists of different contests.

In each contest you will be matched with two other participants who will make their choices at the same time as you. You have to make your own decision based on what you think the other two will choose. They are in the same situation as you so they will not know your decision when they make their choices. Hence, all participants are in the same situation, have the same information and have the same alternatives to choose from. You will now be introduced to a description of the contests. We start by describing how your position in the contest is determined and after that what payment you will get which depends on your choices and your position in the contest.

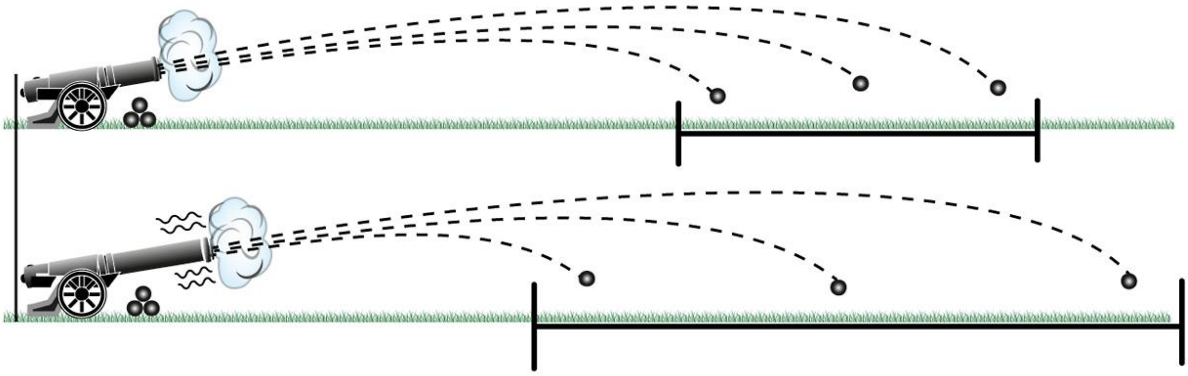
An illustration of the contest: Shooting cannon balls

To get a feeling for the contest we first describe an illustrative example of it before we go through the details. Consider a contest against two other co-participants where you fire an old cannon and your place in the contest (1st, 2nd or 3rd) depends on how far your cannon ball gets before it hits the ground compared to the other co-participants. However, firing a cannon is associated with some uncertainty (due to factors that are difficult to control like the quality of gunpowder and the weather). You start with a cannon associated with a certain distribution (in the target zone). You may affect this distribution in two ways:

1. You can move the cannon forward and thereby increase the probability that the cannon ball reaches longer.



2. You can increase the length of the cannon pipe. This makes it possible for the cannon ball to reach longer, but at the cost that the cannon gets more unstable and therefore also increases the probability for shorter shots. Hence, increasing the length of the cannon pipe increases both the probability for very long shots and very short shots.

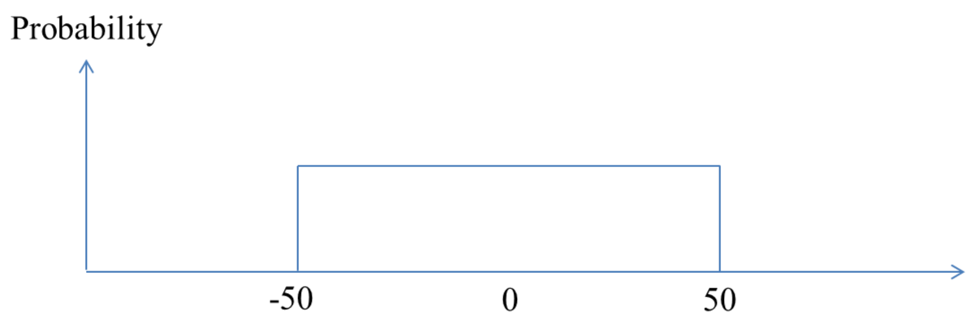


The changes to the cannon—moving it forward or increasing the length of its pipe—is associated with a certain cost. Furthermore, in the contest your reward will only depend on *your place in the contest and thus only indirectly through how far your shot is. These issues will be described in detail below.*

Detailed description of the contest

You will compete with two other participants in getting a high number (corresponding to the length of the cannon shot). Your number will be randomly drawn from a certain distribution that you can affect.

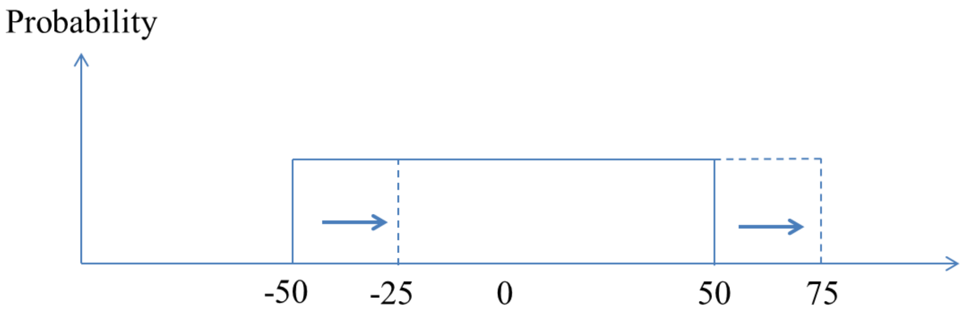
To start with, the distribution which you and your co-participants' numbers will be drawn from has a minimum of -50 and a maximum of 50 and a mean of 0. (This corresponds to the distribution of cannon ball shots when you have not moved the cannon nor increased its pipe length.) The spread of the distribution is measured by the difference between the maximum number and the mean (that is 0). This difference is 50 to start with. To illustrate we have provided a figure of the distribution.



When your number is drawn, all outcomes between the minimum and the maximum are equally likely. For the initial distribution displayed in the figure above, you are equally likely to get any number between -50 and 50.

Now, you can affect the distribution in two ways. You can both increase the mean and the spread of the distribution. However, this will cost you some money that will be deducted from your total earnings (including your show-up fee). This will be described below.

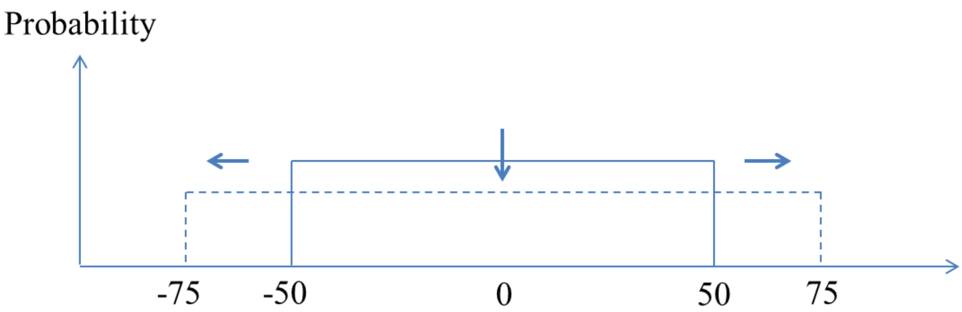
- Increasing the mean: By increasing the mean you will move the distribution to higher numbers by increasing the minimum number as well as the maximum number. This is illustrated below. For any combination of distributions that your co-participants have chosen this will increase the probability of getting a higher number than the other participants. (This corresponds to moving the cannon forwards in the example.) In the figure we have illustrated an increase of the mean by 25.



There is a cost of increasing the mean. The cost of a unit increase of the mean gets more and more costly the higher the mean is. For example, increasing the mean from 0 to 10 costs 5 DKK, but increasing it from 20 to 30 costs 15 DKK. (Mathematically, to increase the mean by x units the cost is given by the square of x divided by 20, that is $C_{\mu} = \frac{x^2}{20}$*.) You can choose any number (between 0 and 30) that you want to invest in increasing the mean. Below we give examples of costs for some mean increases.*

Units increase of mean	*0*	*10*	*20*	*30*
Cost	*0*	*5*	*20*	*45*

- Increasing the spread: By increasing the spread you will increase the probability for very high numbers and very low numbers while you reduce the probability of numbers in the mid-range (see figure below). For any combination of distributions that your co-participants have chosen this will increase the probability of getting the highest and the lowest number of all contestants. (This corresponds to increasing the length of the cannon pipe in the example.) In the figure we have illustrated an increase of the spread by 25. The initial spread (measured as the maximum value minus the mean) was 50 and the spread of the new distribution is 75.

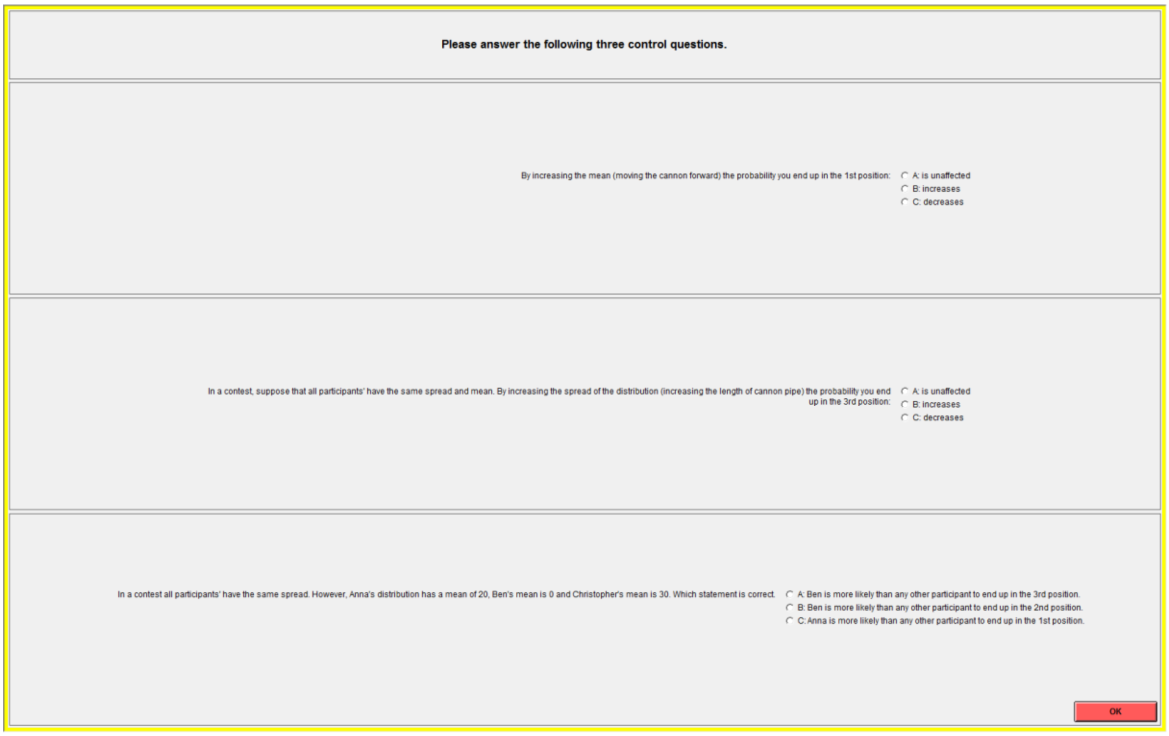


There is a cost of increasing the spread. The cost of a unit increase of the spread gets more and more costly the higher the spread is. For example, choosing to increase the spread by 10 (i.e. increasing the spread from 50 to 60) costs 5 DKK, but going from an increase of 20 to 30 (i.e. increasing the spread from 70 to 80) costs 15 DKK. (Mathematically, to increase the spread by x units the cost is given by the square of x divided by 20, that is $C_\sigma = \frac{x^2}{20}$*.) You can choose any number (between 0 and 30) that you want to invest in increasing the spread. Below we give examples of costs for some spread increases.*

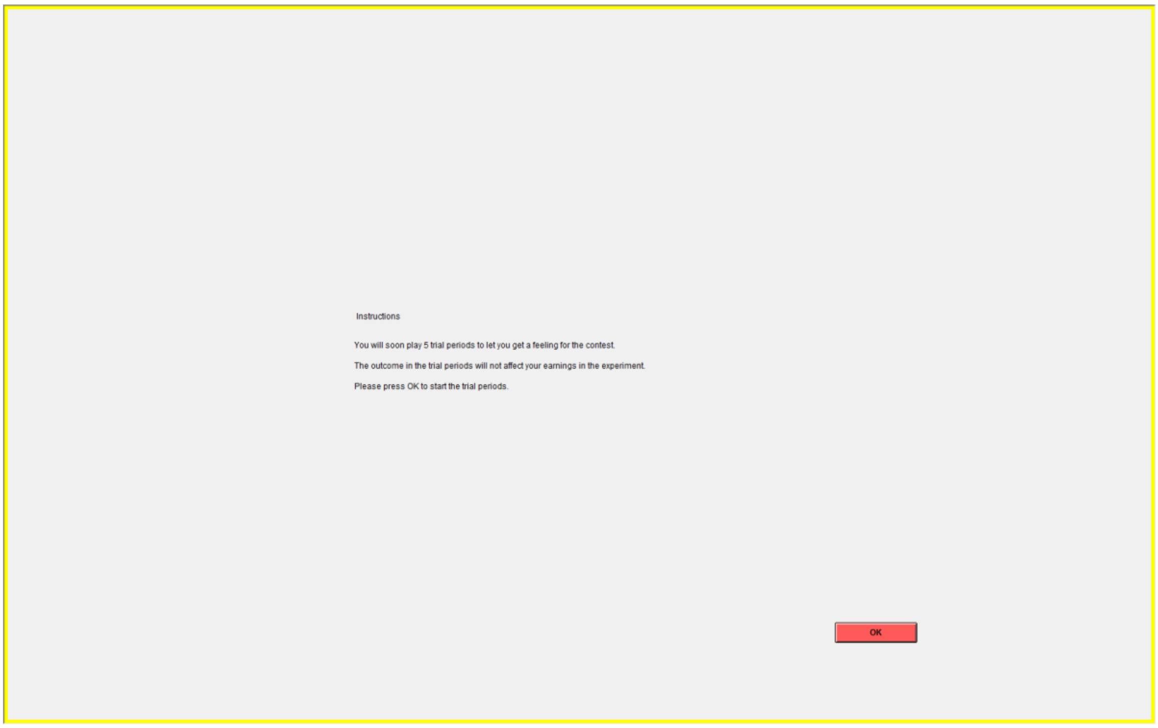
Units increase of spread	*0*	*10*	*20*	*30*
Cost	*0*	*5*	*20*	*45*

A2. Detailed instructions from treatments and screen shots from z-Tree

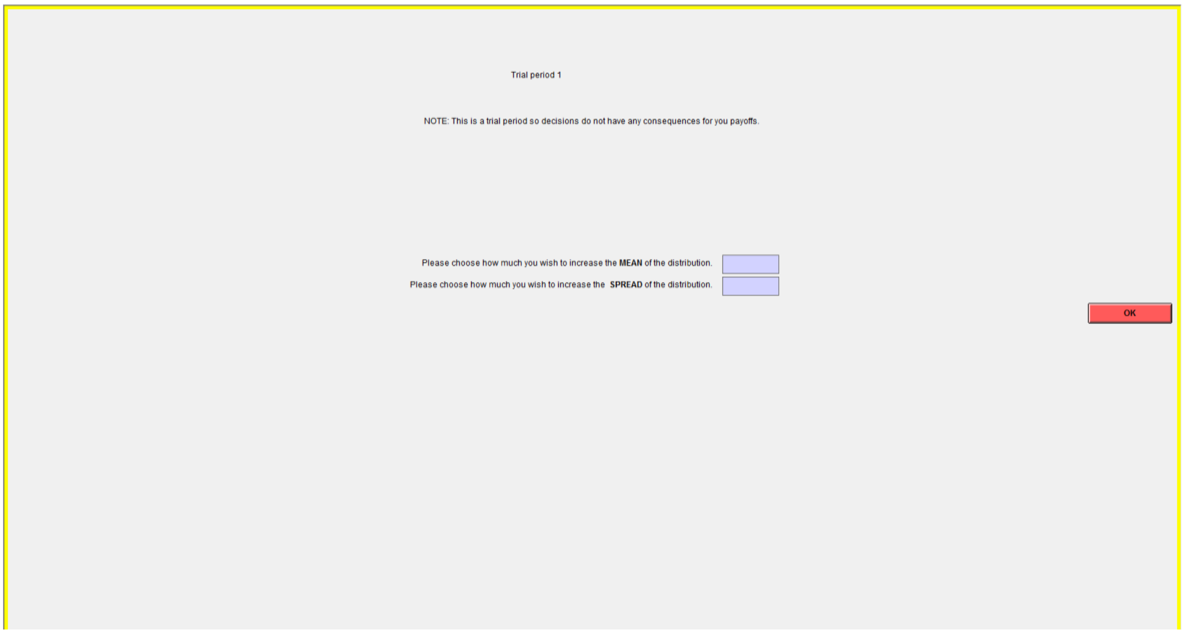
The rest of the instructions were embedded into the program and was displayed on screen. For that reason we now show the main features of the experiment by way of screen shots .



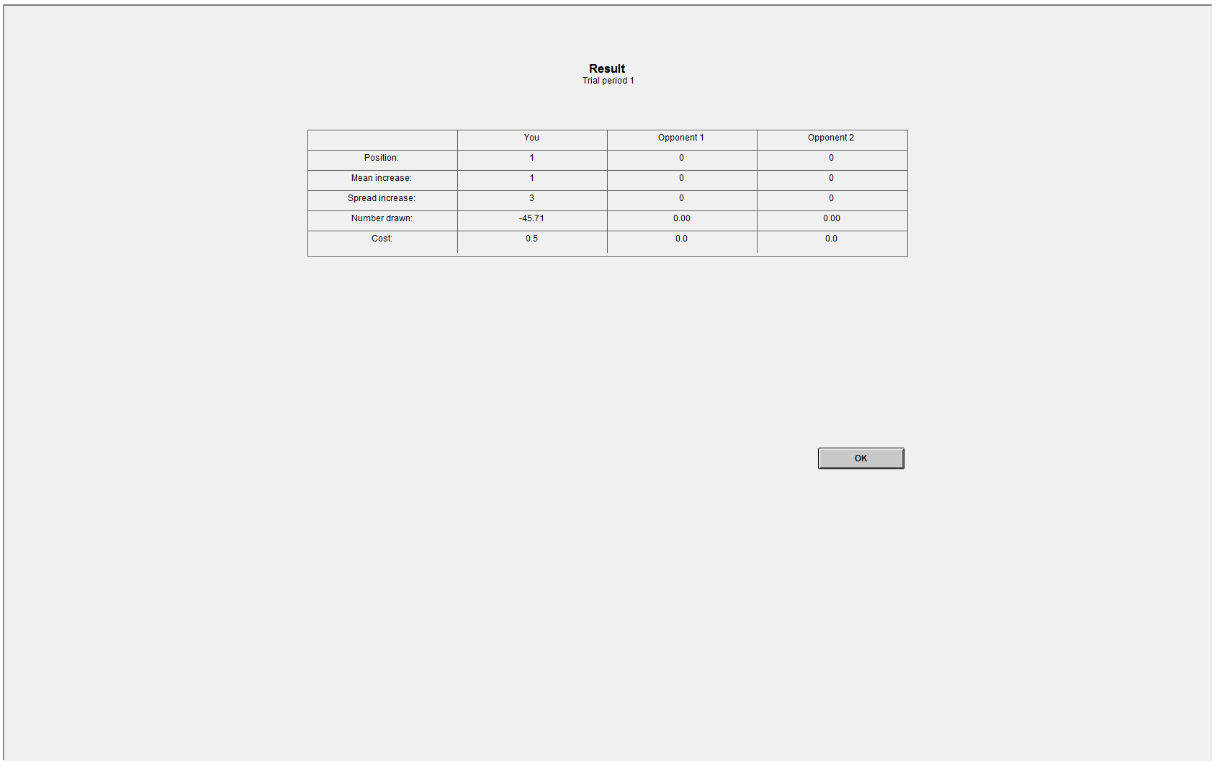
Screen shot 1: Control questions to check that the subject understood the General Instructions



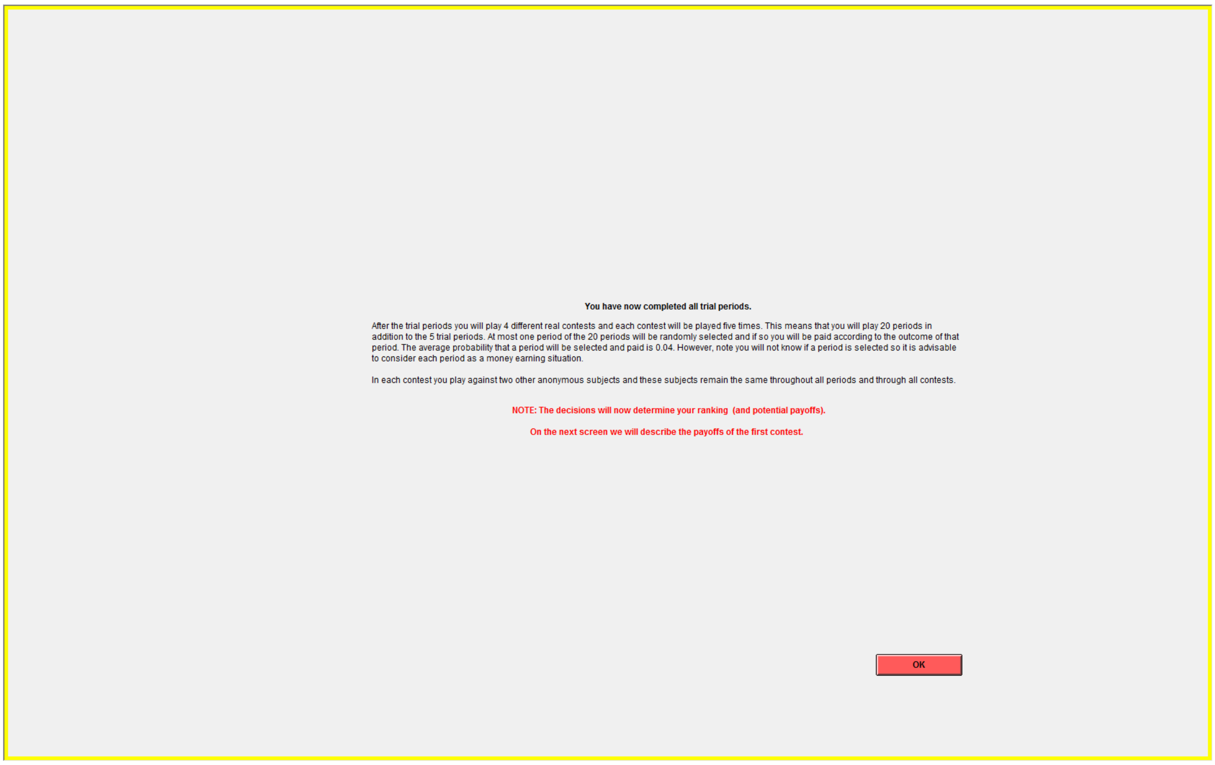
Screen shot 2: Information screen before trial round



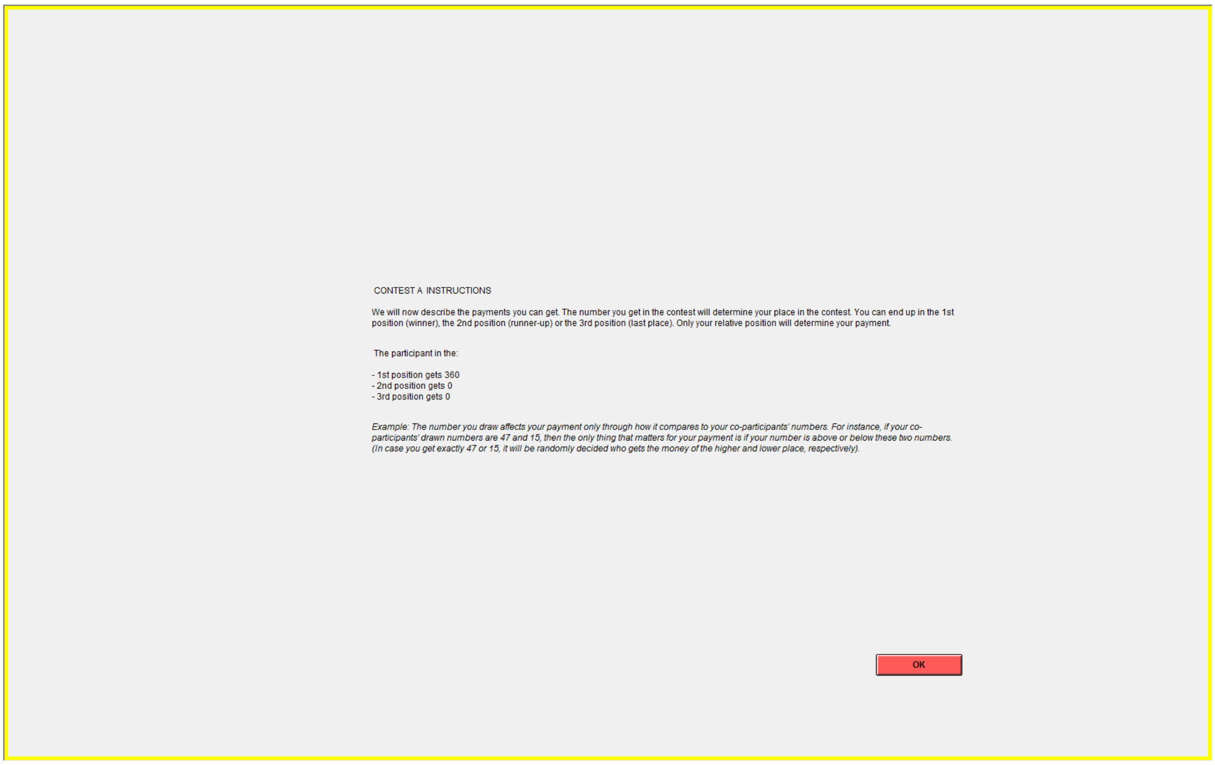
Screen shot 3: Effort and Spread choice in trial period.



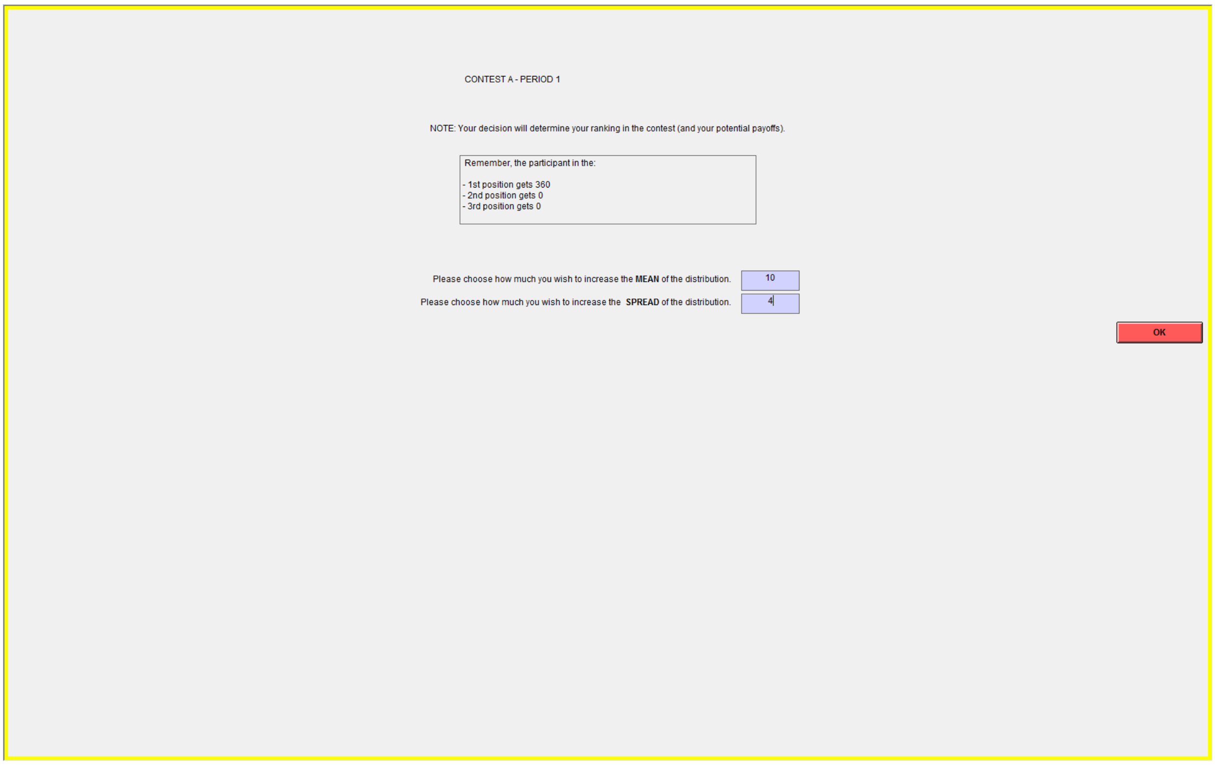
Screen shot 4: Information screen after realized choices



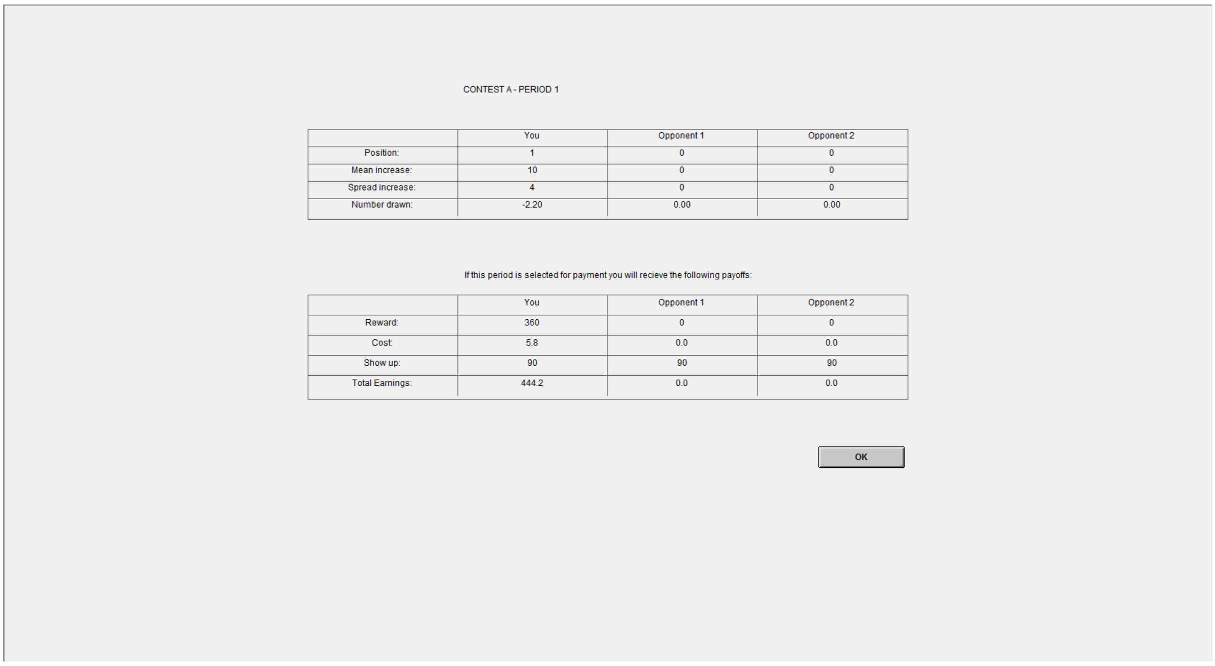
Screen shot 5: Information screen between trial period and first treatment



Screen shot 6: Information screen on WTA treatment



Screen shot 7: Effort and Spread choice in WTA treatment



CONTEST A - PERIOD 1

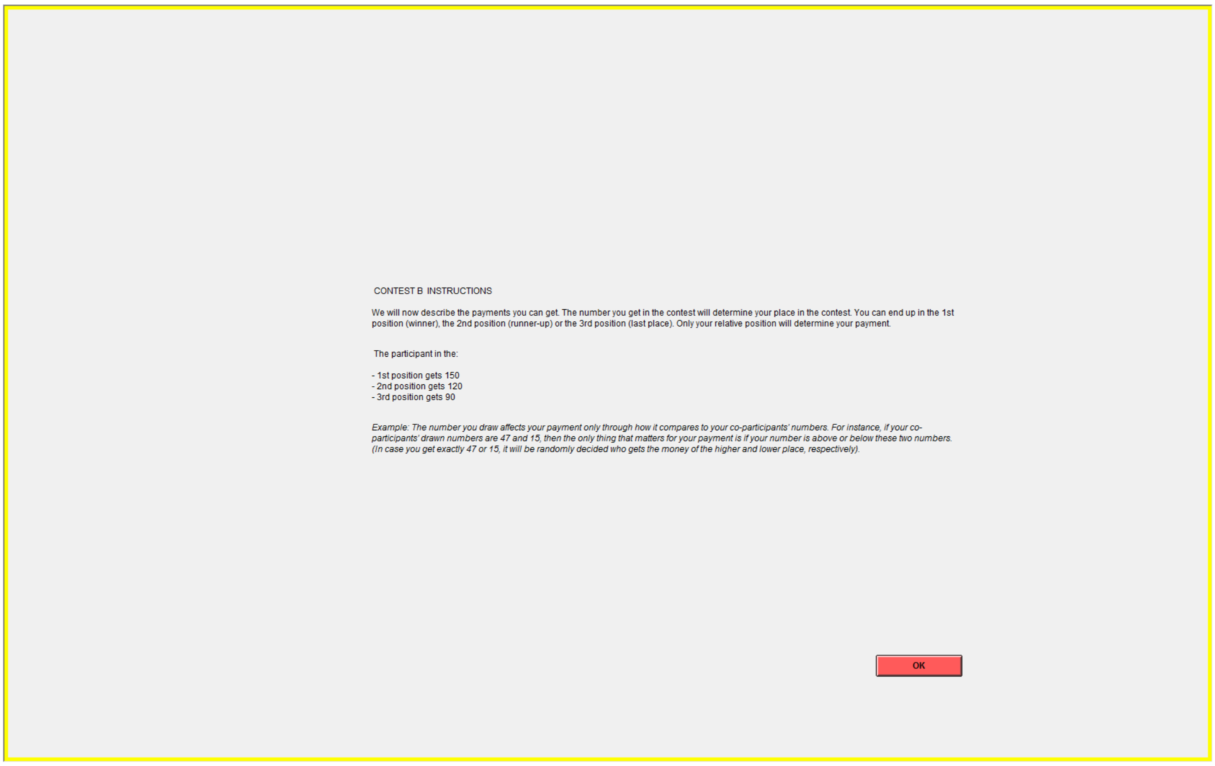
	You	Opponent 1	Opponent 2
Position:	1	0	0
Mean increase:	10	0	0
Spread increase:	4	0	0
Number drawn:	-2.20	0.00	0.00

If this period is selected for payment you will recieve the following payoffs:

	You	Opponent 1	Opponent 2
Reward:	360	0	0
Cost:	5.8	0.0	0.0
Show up:	90	90	90
Total Earnings:	444.2	0.0	0.0

OK

Screen shot 8: Information screen (note the extra information compared to the trial information screen)



CONTEST B INSTRUCTIONS

We will now describe the payments you can get. The number you get in the contest will determine your place in the contest. You can end up in the 1st position (winner), the 2nd position (runner-up) or the 3rd position (last place). Only your relative position will determine your payment.

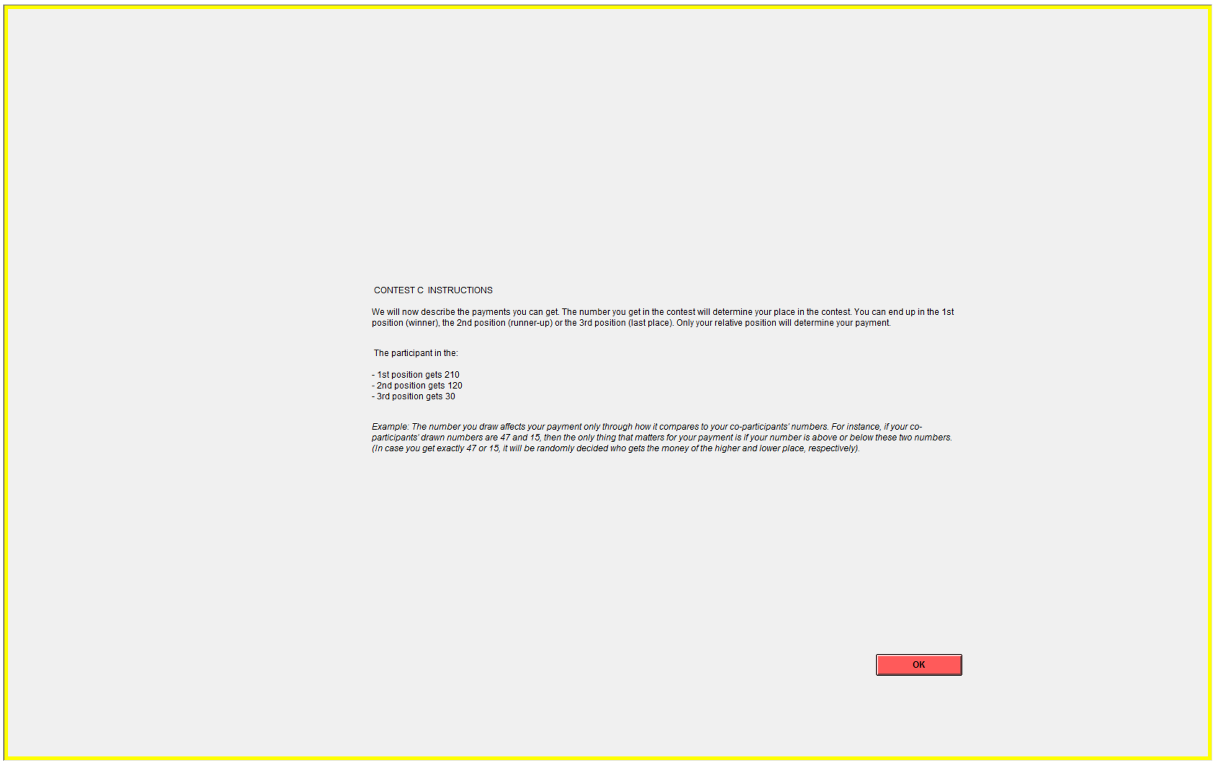
The participant in the:

- 1st position gets 150
- 2nd position gets 120
- 3rd position gets 90

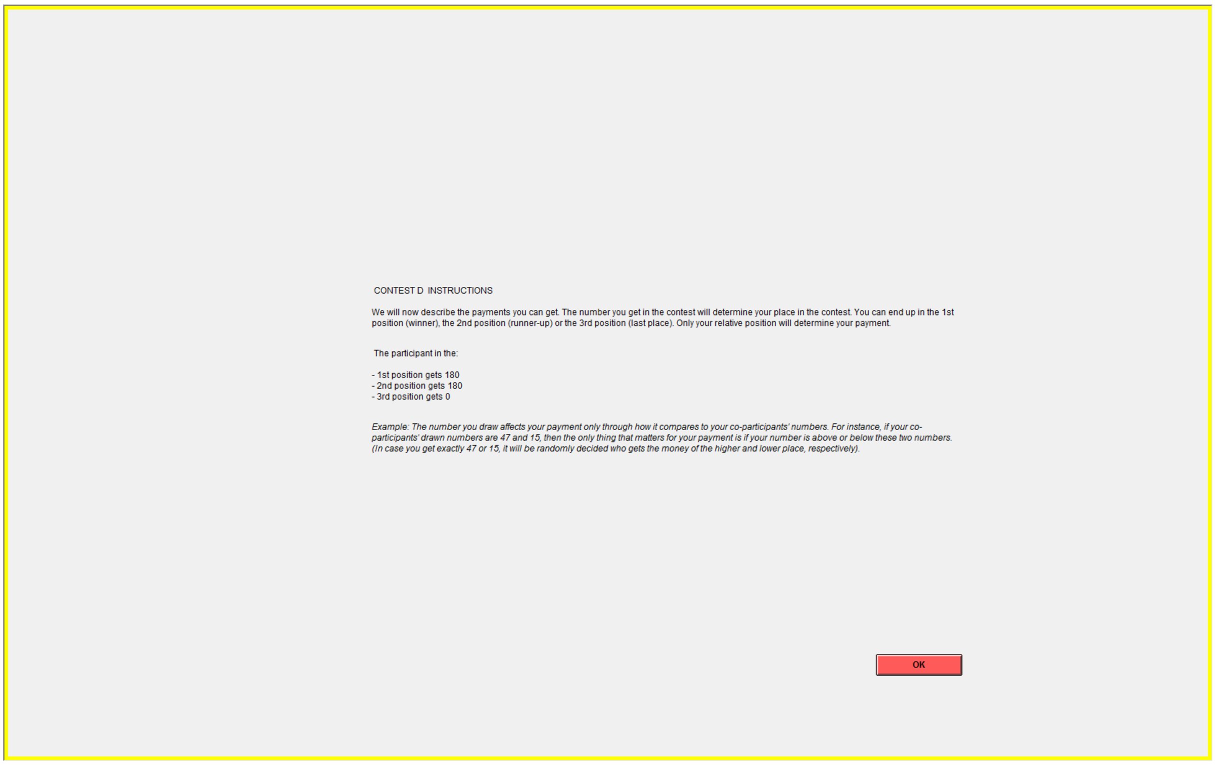
Example: The number you draw affects your payment only through how it compares to your co-participants' numbers. For instance, if your co-participants' drawn numbers are 47 and 15, then the only thing that matters for your payment is if your number is above or below these two numbers. (In case you get exactly 47 or 15, it will be randomly decided who gets the money of the higher and lower place, respectively).

OK

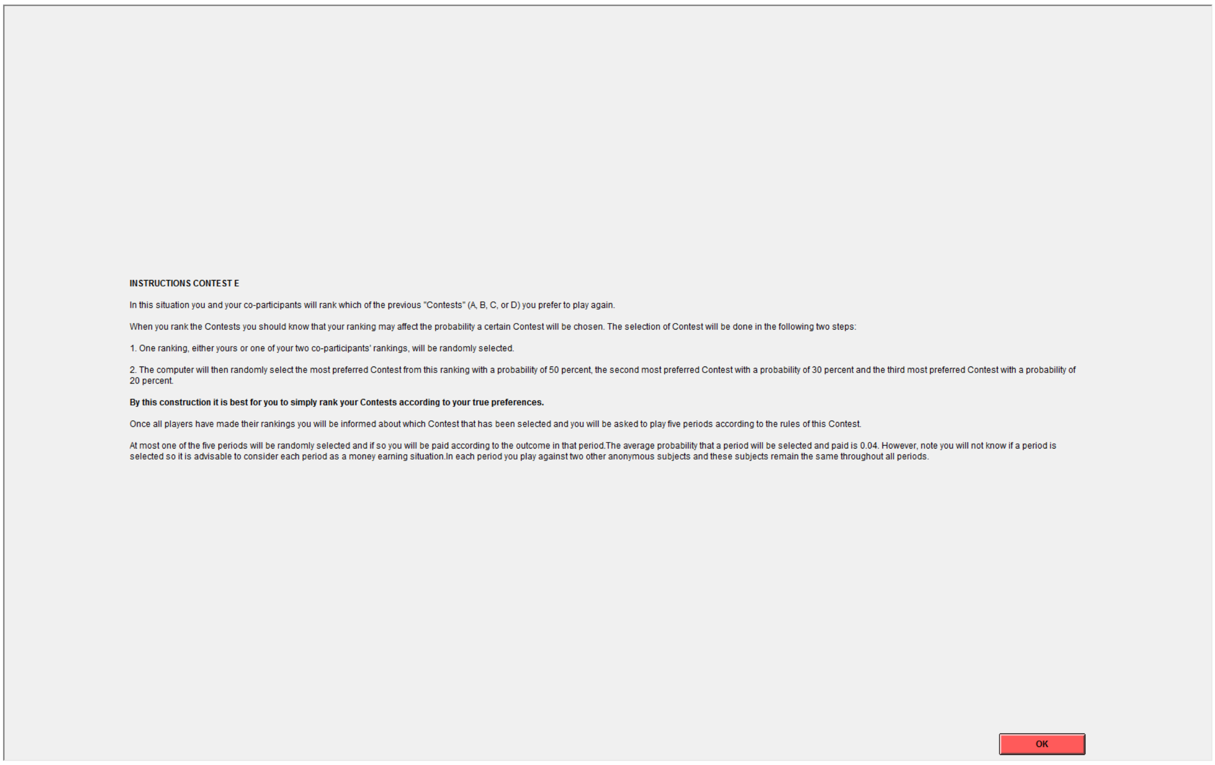
Screen shot 9: Information screen of SFAS



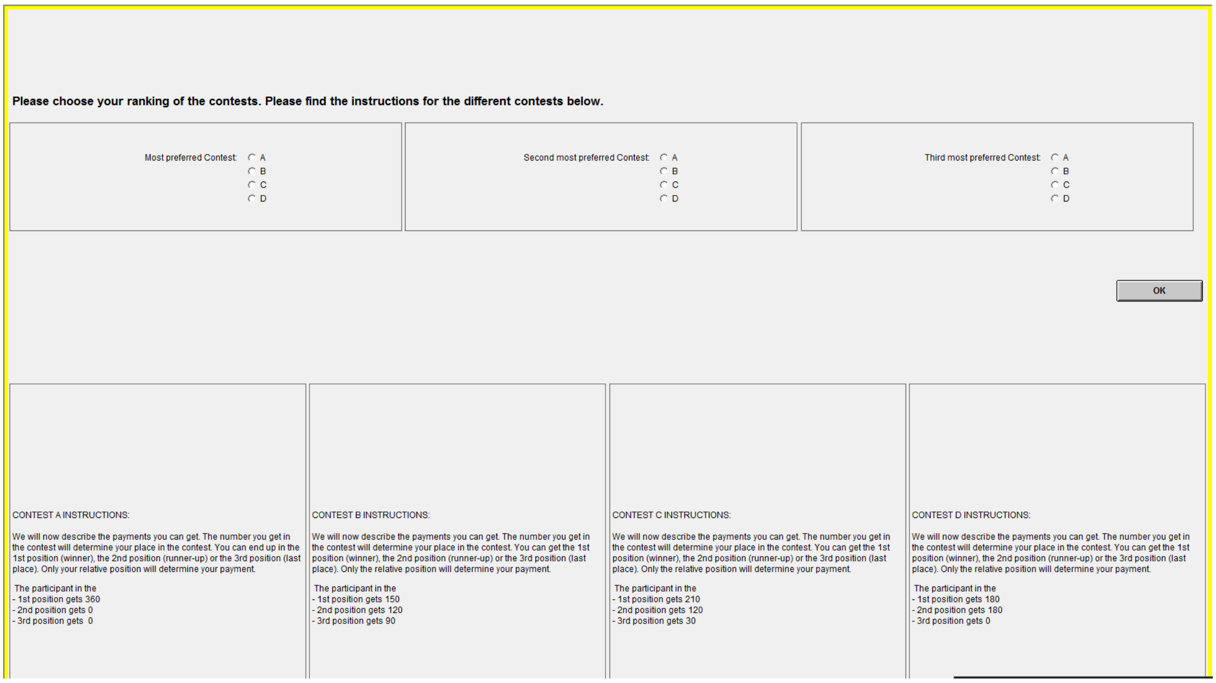
Screen shot 10: Information screen of SFAL



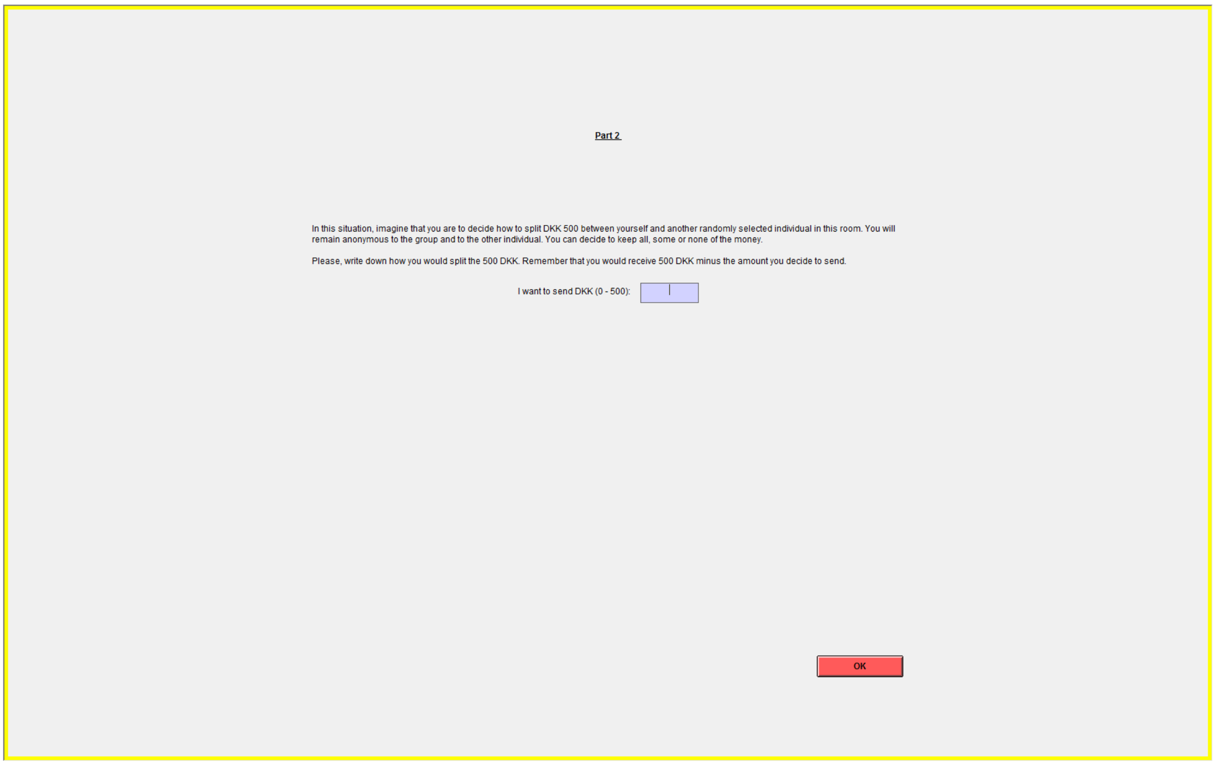
Screen shot 11: Information screen SL



Screen shot 12: Information screen CC



Screen shot 13: Ranking of treatments

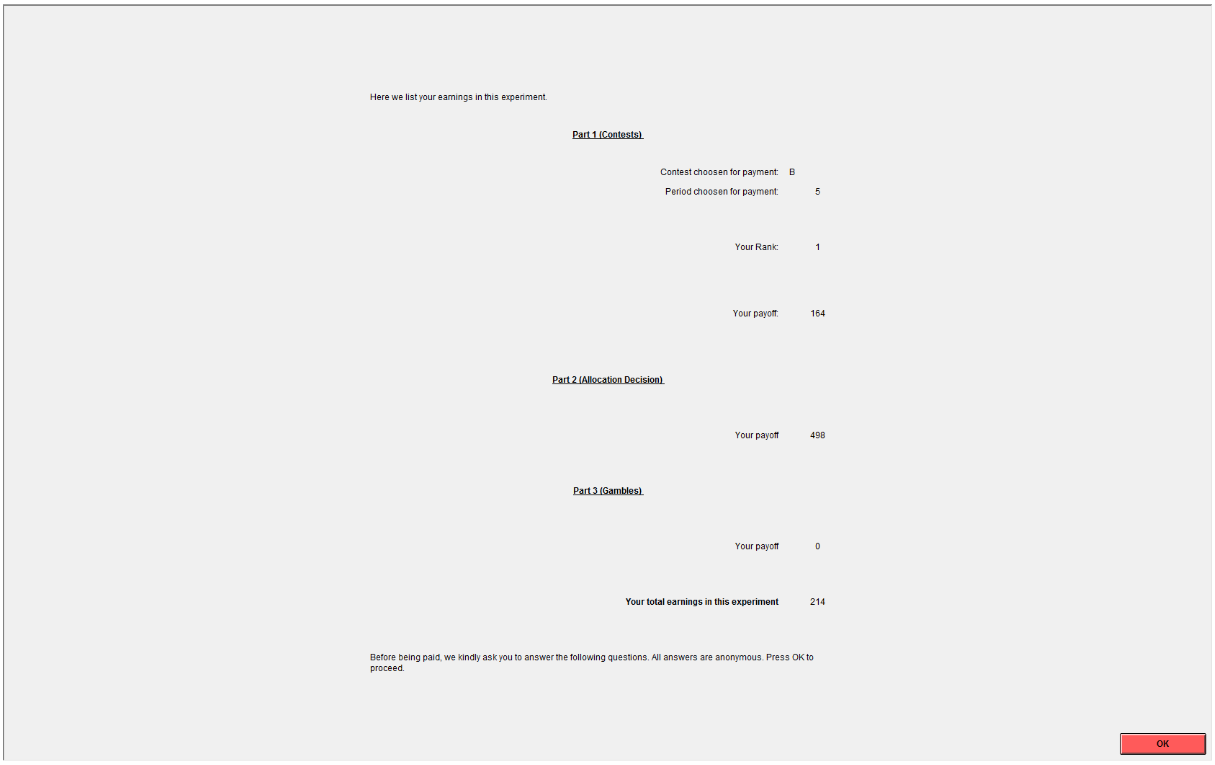


Screen shot 14: Dictator game

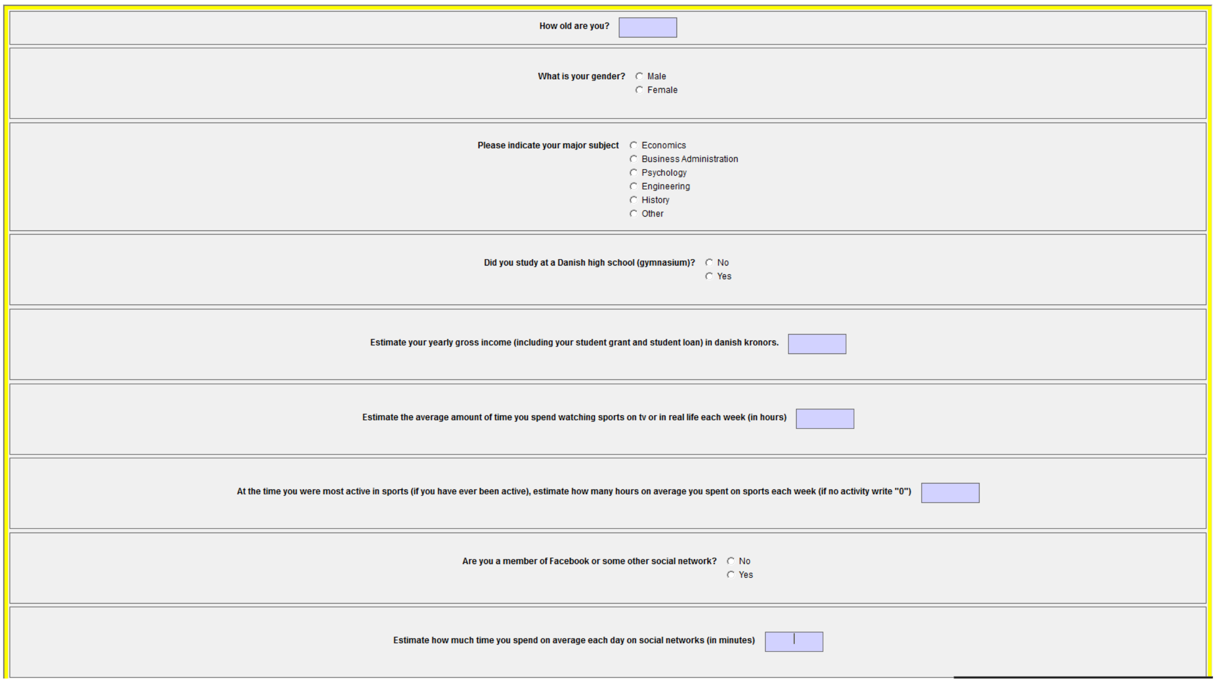
Screen shot 15: Information screen for Risk preference estimation

Please make your choice
240
240
Gamble 1
210
300
Gamble 2
190
340
Gamble 3
170
380
Gamble 4
Choose Gamble 1
Choose Gamble 2
Choose Gamble 3
Choose Gamble 4
150
420
Gamble 5
110
500
Gamble 6
70
540
Gamble 7
10
570
Gamble 8
Choose Gamble 5
Choose Gamble 6
Choose Gamble 7
Choose Gamble 8
View instructions again

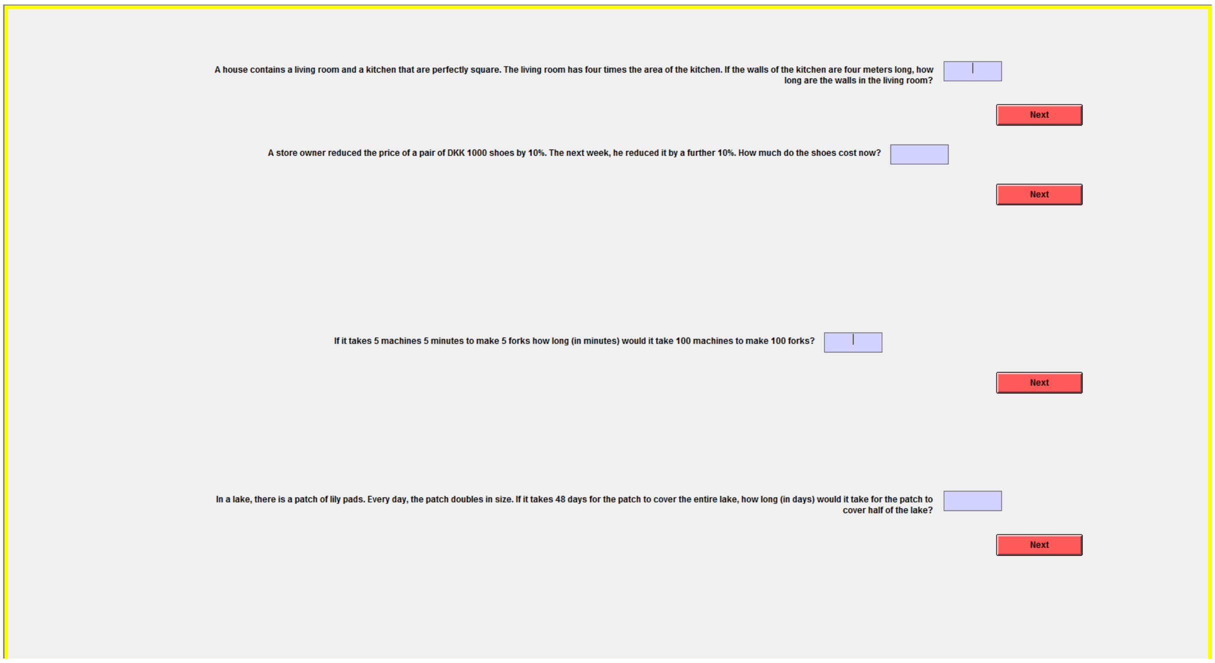
Screen shot 16: Spread choice for preference estimation



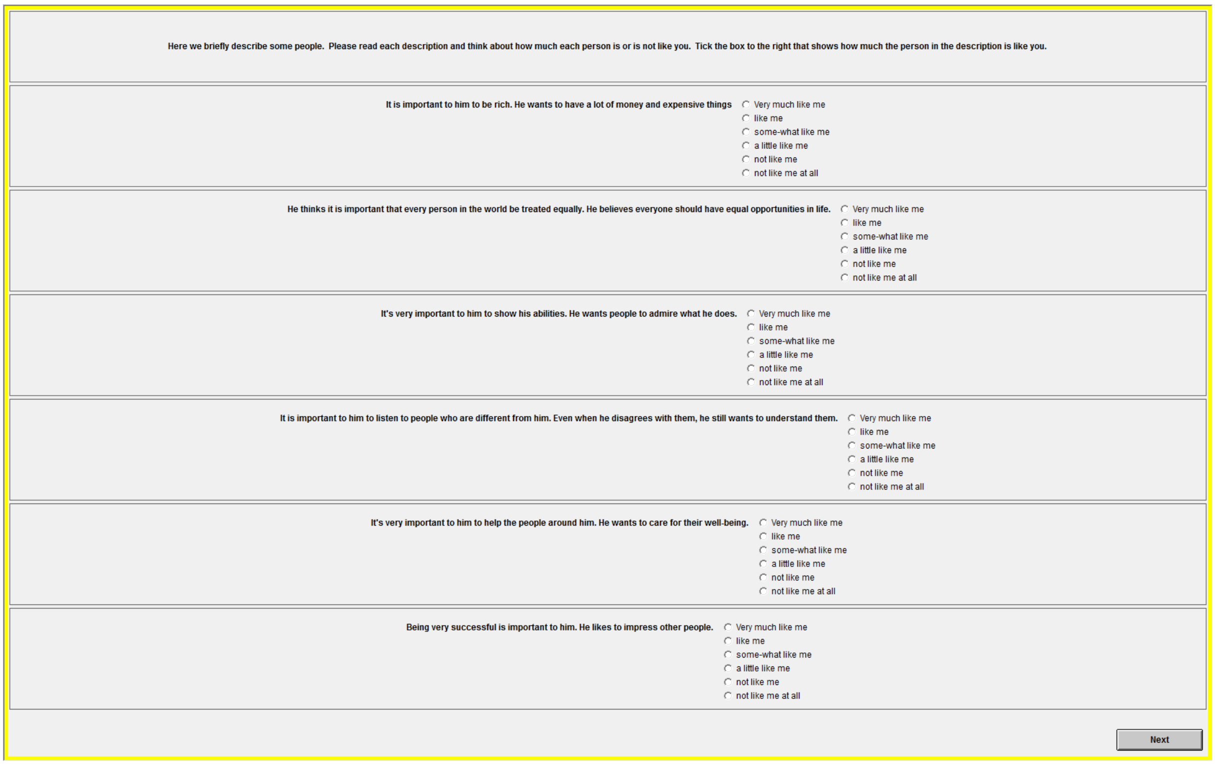
Screen shot 17: Information about payoffs selected for payment



Screen shot 18: Questionnaire



Screen shot 19: CRT questions. Note that in the experiment each question was shown on a separate page.



Screen shot 20: PVQ screen 1

Screen shot 21: PVQ screen 2

Screen 22: PVQ screen 3

Appendix B: Theoretical derivations

B1. The Model (Gilpatric 2009)

Following Gilpatric (2009) we model a contest with one risk neutral principal and three identical risk neutral agents. Each agent $i's$ action $(\mu_i, \sigma_i) \in R_+^2$ determines the mean and the variance of a random variable $Y_i = \mu_i + \varepsilon_i$ where $\varepsilon_i : U(-\sigma_i, \sigma_i)$ and iid. Let F be the associated cdf of the uniform distribution with $f = F'$. The cost associated with choosing μ_i and σ_i respectively is identical and governed by $C\ (\cdot)$ and which is assumed to be twice differentiable, strictly increasing and strictly convex with $C\ (0) = 0$.

We first a contest with one winning prize W_1 and two loosing prizes W_2 with $W_1 > W_2$. The agent with the highest (stochasic) output Y_i wins W_1 whereas the two agent with lower outupts each get W_2. Let $P(\mu, \sigma)$ represent the probability of winning W_1 for a given strategy vector (μ, σ). Then the (expected) payoff for agent i of choosing (μ_i, σ_i) is given by:

$$\pi_i(\mu, \sigma) = P_i(\mu, \sigma) \cdot W_1 + (1 - P_i(\mu, \sigma)) \cdot W_2 - C(\mu_i) - C(\sigma_i).$$

Before we proceed to equilibrium calculations let us first study how P is affected by changes in μ_i and σ_i. What is important for agent i is wheter $Y_i > Y_j = \mu_i + \varepsilon_i > \mu_j + \varepsilon_j$ for each j and $j \neq i$. The probability for this event is given by:

$$P_i = \int_{-\sigma_i}^{\sigma_i} f(\mu_i; \sigma_i)_{j \neq i} F(\varepsilon_i + \mu_i - \mu_j; \sigma_j) d\varepsilon_i. \qquad (1)$$

In what follows we restrict attention to symmetric equilibria. We are now ready to write down the first order conditions for optimality.

$$\frac{\partial \pi_i(\mu, \sigma)}{\partial \mu_i} = \frac{\partial P_i(\mu, \sigma)}{\partial \mu_i} \cdot W - \frac{\partial C(\mu_i)}{\partial \mu_i} = 0$$

$$\frac{\partial \pi_i(\mu, \sigma)}{\partial \sigma_i} = \frac{\partial P_i(\mu, \sigma)}{\partial \sigma_i} \cdot W - \frac{\partial C(\sigma_i)}{\partial \sigma_i} = 0,$$

where $W \equiv (W_1 - W_2)$.

Let us now also consider what happens if we introduce a prize W_3 for finishing last such that $W_1 > W_2 > W_3$. Now the expected profit of agent i reads

$$\pi_i(\mu,\sigma) = P_i(\mu,\sigma)\cdot W + W_2 - R_i(\mu,\sigma)\cdot L - C(\mu_i) - C(\sigma_i)$$

where R_i is the probability of having the lowest output and $L = W_2 - W_3$. We want to study how R_i is affected by changes in μ_i and σ_i. The first order conditions are given by

$$\frac{\partial \pi_i(\mu,\sigma)}{\partial \mu_i} = \frac{\partial P_i(\mu,\sigma)}{\partial \mu_i}\cdot W - \frac{\partial R_i(\mu,\sigma)}{\partial \mu_i}\cdot L - \frac{\partial C(\mu_i)}{\partial \mu_i} = 0$$

$$\frac{\partial \pi_i(\mu,\sigma)}{\partial \sigma_i} = \frac{\partial P_i(\mu,\sigma)}{\partial \sigma_i}\cdot W - -\frac{\partial R_i(\mu,\sigma)}{\partial \sigma_i}\cdot L - \frac{\partial z(\sigma_i)}{\partial \sigma_i} = 0,$$

Notice that, due to the symmetry of F we have

$$\frac{\partial R_1(\mu,\sigma)}{\partial \sigma_1} = \frac{\partial P_1(\mu,\sigma)}{\partial \sigma_1}$$

$$\frac{\partial R_1(\mu,\sigma)}{\partial \mu_1} = -\frac{\partial P_1(\mu,\sigma)}{\partial \mu_1}$$

at at point of symmetry. We can thus re-write the first order conditions as

$$\frac{\partial \pi_i(\mu,\sigma)}{\partial \mu_i} = \frac{\partial P_i(\mu,\sigma)}{\partial \mu_i}\cdot (W + L) = \frac{\partial C(\mu_i)}{\partial \mu_i}$$

$$\frac{\partial \pi_i(\mu,\sigma)}{\partial \sigma_i} = \frac{\partial P_i(\mu,\sigma)}{\partial \sigma_i}\cdot (W - L) = \frac{\partial C(\sigma_i)}{\partial \sigma_i}.$$

Notice that the first condition does not depend on W_2 whereas the second do.

B2. Experimental implementation

Let us now consider our experimental implementation. In order to assure that the second order conditions to be met the support of F cannot be below 100 so the minimal range of the noise is $[-50,50]$. Each contestant can add to the baseline spread by choosing some $\mu_i \in [0,30]$ so that there is zero probability that the support of one contestant is above any other contestant. Also let $\sigma_i \in [0,30]$. The costs of choosing μ_i and σ_i are given by the following function.

$$C(x_i) = \frac{x_i^2}{20}$$

B21. WTA

Under WTA condition we have the following set of prizes.

$$W_1 = 360$$

$$W_2 = 0$$

$$W_3 = 0$$

Which gives

$$W + L = 360$$

$$W - L = 360$$

We then get the following FOCs

$$\frac{1}{2}\frac{1}{(50+\sigma_i)} \cdot (360) = \frac{\mu_i}{10}$$

$$\frac{1}{6(50+\sigma_i)} \cdot (360) = \frac{\sigma_i}{10}$$

The symmetric Nash equilibrium strategy profile is given by:

$$\mu_i^* = 30$$

$$\sigma_i^* = 10$$

B22. SL

Under the SL condition we have the following set of prizes

$$W_1 = 180$$

$$W_2 = 180$$

$$W_3 = 0$$

Which gives

$$W + L = 180$$

$$W - L = -180$$

We get the following FOCs

$$\frac{1}{2}\frac{1}{(50+\sigma_i)}\cdot(180) = \frac{\mu_i}{10}$$

$$\frac{1}{6(50+\sigma_i)}\cdot(-180) = \frac{\sigma_i}{10}$$

The symmetric Nash equilibrium strategy profile is given by

$$\mu_i^* = 18$$

$$\sigma_i^* = 0$$

B23. SFAS

Under the SFAS condition we have the following set of prizes

$$W_1 = 150$$

$$W_2 = 120$$

$$W_3 = 90$$

Which gives

$$W + L = 60$$

$$W - L = 0$$

We get the following FOCs

$$\frac{1}{2}\frac{1}{(50+\sigma_i)}\cdot(60)=\frac{\mu_i}{10}$$

$$\frac{1}{6(50+\sigma_i)}\cdot 0=\frac{\sigma_i}{10}$$

Now clearly it will be optimal to set $a^* = 0$ (a boundary solution) so the solution reads

$$\mu_i^* = 6$$

$$\sigma_i^* = 0$$

B24. SFAL

Under the SFAL condition we have the following set of prizes

$$W_1 = 210$$

$$W_2 = 120$$

$$W_3 = 30$$

$$W + L = 180$$

$$W - L = 0$$

We get the following FOCs

$$\frac{1}{2}\frac{1}{(50+\sigma_i)}\cdot(180)=\frac{\mu_i}{10}$$

$$\frac{1}{6(50+\sigma_i)}\cdot 0=\frac{\sigma_i}{10}$$

Now clearly it will be optimal to set $a^* = 0$ from the second FOC so the solution reads

$$\mu_i^* = 18$$

$$\sigma_i^* = 0$$

Appendix C: Regression Results on Popularity

In this section we will analyze the popularity of the different prize schemes at the individual level. To study the effect of choosing incentive scheme (the last subsession) we estimate a multinomial logit regression model. In addition to the previously used control variables, we here also included variables capturing the subject's earnings in the different prize schemes. For each subject, we sum the earnings in each prize scheme and divide by the total earnings across all prize schemes; the variables are denoted WTA_earnings, SL_earnings, SFAS_earnings and SFAL_earnings. The reason behind the inclusion of these variables is that we hypothesize that prize schemes associated with high experience of previous earnings may be ranked more or less favorably by the subjects who are not fully rational. There are at least two opposing mechanisms that may be channeled through experience of earnings. The "gambler's fallacy" would cause subjects who have experienced good performances in a given treatment to think that the probability is higher for bad performances in the same treatment in the future. Contrary to this, the "hot hand fallacy" would make subjects think that the good performances in a given treatment are caused by his "hot hand" and hence that the probability for good performances in the future is reinforced by good performances in the past. For a discussion of these effects and empirical investigation, see e.g., Croson and Sundali (2005).

It should be noted that "experienced earnings" are potential earnings and only one of the periods was selected for payment in the end. We also control for in which order the four prize schemes occurred (Order 2 – Order 4). Table C1 reports marginal effects. Each column represents an Incentive scheme and the numbers give the change in probability of giving each incentive scheme the highest rank from a marginal change in that control variable (standard errors in parenthesis).

There is a significant effect in that subjects who donate more in the dictator game are less likely to rank WTA at the top. This makes sense if dictator donations reflect distributional preferences since WTA generates the most unequal payoff distribution. Except for this effect there are no strong convincing determinants of the subjects' rankings of the prize schemes. The order of the subsessions appears to play a role but not in an obvious consistent pattern. Furthermore, earnings in the prize scheme do not affect the subjects ranking of the prize schemes in a consistent and convincing way.

Table C1. Multinomial logit model, most preferred incentive scheme

	WTA	SL	SFAS	SFAL
WTA_earnings	0.006	0.507	-1.127	0.614
	[0.631]	[0.706]	[0.714]	[0.646]
SL_earnings	0.332	0.137	0.106	-0.575
	[0.873]	[0.994]	[0.928]	[0.907]
SFAS_earnings	-0.197	0.958	-2.163*	1.402
	[1.121]	[1.336]	[1.269]	[1.142]
SFAL_earnings	-0.169	1.168	-1.247	0.249
	[0.959]	[1.062]	[1.011]	[1.034]
Order 2 (SL_SFAL_SFAS_WTA)	0.035	-0.041	0.047	-0.041
	[0.073]	[0.075]	[0.074]	[0.091]
Order 3 (SFAL_WTA_SL_SFAS)	0.032	-0.177*	-0.057	0.203**
	[0.069]	[0.097]	[0.085]	[0.081]
Order 4 (SFAS_SL_WTA_SFAL)	0.140**	-0.022	-0.199**	0.081
	[0.065]	[0.082]	[0.094]	[0.090]
Riskchoice	-0.002	0.015	-0.019	0.007
	[0.013]	[0.015]	[0.014]	[0.015]
CRT	-0.003	-0.004	0.024	-0.017
	[0.016]	[0.018]	[0.020]	[0.019]
Dictator Offer	-0.001**	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]
N	237	237	237	237

Notes: Robust standard errors clustered at the group level in brackets. The _earnings variables measure the subject's earnings in the different prize schemes. For each subject, the earnings in each prize scheme summed and divided by the total earnings across all prize schemes. The Order variables capture the order in which the different contests were played. Riskchoice indicates the switch point in the risk elicitation task with a higher number indicating a higher degree of risk aversion. CRT gives the number of correct answers to the CRT questions. Dictator offer reports the offer made to an anonymous recipient in the dictator game.

*** p<0.01, ** p<0.05, * p<0.1