UNIVERSITÄT PASSAU

Wirtschaftswissenschaftliche Fakultät

Performance in the Beauty Contest: How Strategic Discussion Enhances Team Reasoning.

Caroline Baethge

Diskussionsbeitrag Nr. B-17-16

Betriebswirtschaftliche Reihe ISSN 1435-3539

PASSAUER DISKUSSIONSPAPIERE

Herausgeber:
Die Gruppe der betriebswirtschaftlichen Professoren
der Wirtschaftswissenschaftlichen Fakultät
der Universität Passau
94030 Passau

Performance in the Beauty Contest: How Strategic Discussion Enhances Team Reasoning.

Caroline Baethge

Diskussionsbeitrag Nr. B-17-16

Betriebswirtschaftliche Reihe ISSN 1435-3539

Adresse des Autors/der Autoren:

Universität Passau, Lehrstuhl für Betriebswirtschaftslehre mit Schwerpunkt Management, Personal und Information, Innstraße 27, 94032 Passau, E-Mail: caroline.baethge@web.de

PERFORMANCE IN THE BEAUTY CONTEST: HOW STRATEGIC DISCUSSION ENHANCES TEAM REASONING.

Caroline Baethge

This paper will analyze the influence that discussion has on the performance of two-player teams in an experimental beauty contest game. Teams were allowed to communicate via electronic chat. The findings suggest that teams are not more rational than individuals in terms of reasoning. However, strategically sophisticated discussion in teams such as the repetition of the experimental rules, as well as guesses at other players’ behavior, positively influences their level of reasoning. On the other hand, simple communication without strategic elements does not enable the teams to improve their reasoning level. Focusing on the performance of teams in terms of distance to the target number, I demonstrate that the amount of information exchanged between team members improves the teams' performance if and only if the teams strategically discuss their choice. The experimental results underline the assumption that communication per se does not foster strategic thinking, but that sophisticated discussion in teams enables the players to succeed in a strategic environment such as a beauty contest game.

Keywords

strategic discussion · team performance · reasoning · group decision-making · beauty-contest game

Highlights

- This study analyzes the influence of strategic discussion on the performance of two-player teams.
- Teams are not more or less rational than individuals in terms of reasoning.
- The amount of information exchanged improves team’s performance if and only if the teams strategically discuss their choice.
- Sophisticated discussion in teams enables players to succeed in a strategic environment.

1. Introduction

"Two souls, one thought" is an often used quotation when talking about the relationship between two people. This may be a true statement when persons in a team know each other very well and therefore sometimes have similar experiences, opinions and thought processes. But what about teams which come together ad hoc? Can individual members combine their abilities, resources and at times controversial opinions to form a successful team? In a simple guessing game like the beauty contest, the key to success is guessing the average opinion of others correctly. In this setting, team players have the advantage of being able to discuss their thoughts before deciding what their strategy to win will be. Individuals, on the other hand, have to reach a decision by themselves, which may lead to a less accurate guessing of others' behavior because they are not challenged by anyone. The exchange of information between team players could lead them to a better result than they would have achieved if deciding on their own. This being the case, are teams then able to take advantage of their numbers to be successful in an experimental beauty contest?

This paper describes an experimental beauty contest with two-person teams and individual decision-makers. The aim of the experiment is to examine the difference between the rationality of teams and individuals and more importantly to determine whether or not discussion in teams enhances their reasoning. Previous studies on team decisions in purely strategic settings such as the beauty contest game à la Keynes (1936) and Nagel (1995) were conducted by Kocher et al. (2005; 2006) and Sutter (2005). Their findings suggest that groups are not necessarily better than individuals per se, but that group size and learning through repetition leads to the superior performance of the former.

To expand upon the previous literature, the following experiment is constructed as two simple one-shot beauty-contest games. The first beauty contest game, in both the team and control group treatment, was designed to investigate the individual reasoning of participating subjects and consisted of eight single players in each decision making unit. In order to examine the possibly differential behavior of teams and individuals, the second beauty contest in the team treatment was designed to observe teams and individuals in direct interaction. Three teams of two members each competed with two individuals. However, the second beauty contest in the control treatment consisted again of eight single players such as in the first beauty contest (see Table 1 p. 10). The chosen design offers the possibility both to compare the behavior of teams and individuals and to observe whether individuals who are playing against teams behave differently than individuals playing against individuals only. Additionally, it allows for controlling of the individual ability of subjects which could be a decisive factor when comparing the performance of teams and individuals. Moreover, an electronic communication device was used to record the discussion within teams without influencing them by other factors such as gender, age or time limit. This novel element allows to answer the question as to whether or not discussion within a team has a positive influence on its performance. A conclusion could be drawn towards the efficiency of teamwork contributing to the still insufficient empirical literature about the decision process within teams.

My findings suggest that teams do not generally outperform individuals. But strategic discussion is a crucial element in determining whether or not teams enhance their reasoning and therefore their performance.

The remainder of this paper is structured as follows: section 2 reports the research questions, paying special attention to economic experimental evidence. Section 3 describes a basic beauty contest focusing on the game-theoretic solution, and presents the model used for evaluating the reasoning level for one-shot games. Section 4 describes the experimental design and procedure. The results of the experiment are shown in section 5. Section 6 suggests sources for future research and summarizes the results.

2. Hypotheses

Summarizing the experimental evidence on groups and individuals in different experimental settings[1], previous research in this area has arrived at contradicting results concerning the superiority of teams. Even though both individuals and groups differ from game-theoretic predictions, the majority of studies consider groups to be more rational in terms of payoff-maximizing behavior and selfishness than individuals (Blinder and Morgan 2005; Blinder 2008; Bornstein and Yaniv 1998; Bornstein et al. 2004; Cheung and Palan 2012; Cooper and Kagel 2005; Cox 2002; Gillet et al. 2009; Kocher and Sutter 2007; Kugler et al. 2007; Lombardelli et al. 2005; Luhan et al. 2009; Maciejovsky and Budescu 2007; Rockenbach et al. 2007; Sutter 2007). However, some studies suggest that groups behave less rationally (Cason and Mui 1997; Cox and Hayne 2005; Cox and Hayne 2006; Sutter et al. 2009) or do not find any difference in reaching superior decisions compared to individuals (Bone et al. 1999; Rockenbach et al. 2007). The findings of teams in experimental beauty contest games (Kocher and Sutter 2005; Kocher et al. 2006; Sutter 2005) suggest that not only

[1] For a detailed overview of previous research concerning group and individual decisions in the psychological literature see Cox and Hayne (2006).

do groups significantly improve their reasoning, but that they also outperform individuals in terms of payoff in the course of repetition.

To see whether the results hold for the smallest possible team size in a one-shot beauty contest, this study focuses on teams of two members. So far, only Sutter (2005) included two-person teams in his study on the beauty contest. He did not detect any differences in performance compared to individual players. Furthermore, Sutter (2005), as well as Kocher and Sutter (2005) could not control for the individual ability of team members due to their chosen experimental design. On the basis of previous experimental evidence I therefore set up hypothesis 1, which will be analyzed in section 5.

Hypothesis 1. *Teams are more rational than individuals.*

The superior rationality of teams could have an effect on two factors: Firstly, teams could be closer to the game-theoretic equilibrium (hypothesis 1.1). Secondly, they may outperform individuals in the beauty contest i.e. win more often (hypothesis 1.2). This possible advantage of teams could be attributed to the fact that members simply have to coordinate with teammates on a joint number.

Alternatively, a team could primarily improve its reasoning due to the quantity or quality of the discussion. Experimental studies on communication have so far focused on strategically relevant communication and its influence on coordination or other-regarding behavior (see Bochet et al. 2006; Brandts and Cooper 2007; Charness and Dufwenberg 2006, 2010; Cooper et al. 1989, 1992; Cooper and Kagel 2005; Dawes et al. 1977; Duffy 2002; Farrell and Rabin 1996; Frohlich and Oppenheimer 1998; Sutter and Strassmair 2009). They all found that communication had a positive impact, in the form of increased contribution rates or more trusting decisions.[2] Only two studies, namely Burchardi and Penczynski (2012) and Penczynski (2010), examined the effect of communication with one-way simultaneous messages on the outcome in the beauty contest game. Whereas Burchardi and Penczynski (2012) focused on a refinement of the level-*k* model of reasoning, i.e. one third of the subjects play non-strategically even if they choose numbers related to higher levels of reasoning, the later study examined the persuasiveness of messages. Penczynski (2010) found that strategically sophisticated messages are persuasive and induce less sophisticated subjects to

[2] Other studies examining strategically irrelevant communication also found a positive effect of communication on cooperation and trust (see Buchan et al. 2006; Brosig et al. 2003; Dawes et al. 1977; Fiedler 2009; Fiedler and Haruvy 2009).

correct their initial choice which could explain why teams are closer to the game theoretic equilibrium than individuals.

In contrast to my study, neither paper involved teams, focusing instead only on the individual reasoning level of subjects. The first paper, in which team players[3] could actually communicate via an instant messenger, was by Cooper and Kagel (2005). The team members had the possibility to chat for three minutes in order to arrive at a joint decision in a signaling game. If their choices did not coincide the computer randomly chose one decision. In my study teams could communicate freely via an electronic chat function and had to reach a joint decision with an equal payoff such as in Cooper and Kagel (2005). To my knowledge this paper is the first study on team communication in the beauty contest game.

Since the introduction of a communication device in a team beauty contest is the novel element of my experimental design, a closer look has to be given to the role played by discussion within teams and the effect it has on team reasoning. I investigated the chat protocol more closely to gain further insights into the reasoning of teams. I used a content analysis to examine the team conversations which was pioneered by psychologists and sociologists (see e.g. Krippendorff 2013; Houser and Xiao 2011). In a content analysis such as this, the researcher first develops categories based on the communication protocol and then recruits two or three research assistants who classify the messages according to the coded categories. The assistants should work independently so as to gain unbiased results compared to each other. Several experimental studies have applied this method for coding chat protocols (e.g. Brandts and Cooper 2007; Cooper and Kagel 2005; Sutter and Strassmair 2009).[4]

In this study, teams were separated into two categories: “discussion” and “no discussion”. Teams which only stated numbers and jointly accepted them were classified as “no discussion” teams, whereas teams which strategically discussed their choice were classified as “discussion” teams. But what makes a discussion strategic, thereby distinguishing it from other forms of internal team communication? Since the key in an experimental guessing game is to correctly anticipate others’ behavior and understand the strategic rules of the game several aspects account for strategic discussion. These include, 1) the repetition and discussion of the experimental rules, which underline the subjects’ understanding and interpretation of the strategic setting, 2) guesses about other players’ behavior within their group, which is crucial for winning the guessing game, 3) the well-

[3] To clearly separate the term “group” from “team” I use the latter one in the sense of subjects who jointly reach a decision and receive an equal payoff. A “group” on the other hand is a pairing of subjects who play against each other in the beauty contest.

[4] There is also non-experimental work using this method of content analysis (e.g. Pesendorfer and Koeszegi 2006).

founded proposition of a certain choice, and 4) statements about the dynamic or solution of the game with partial or complete explanation of the game-theoretic solution. For example, team 3 in group 18 stated that, "if everybody is thinking the number has to be smaller than 50, everybody is inserting 25 etc.". This shows that the subjects discussed the dynamic of the game and guessed about the behavior of others, illustrating that the conversation clearly involved strategic aspects and therefore differs substantially from regular or non-strategic communication within teams. By contrast, communication in teams which was classified as "no discussion" involved 1) simple statements of ill-founded choices such as "gut feeling" or "happy medium", 2) no discussion of the experimental rules, and 3) statements that the choice cannot be computed. For example, team 2 in group 5, which was classified as a no discussion team stated: "Player 1: 20? Player 2: good Player 1: OK". Appendix B provides a complete overview of the two exemplary original chat protocols.

Taken together, only 45% of the teams discussed their choice via chat, whereas 55% of the teams did not discuss their choice but only used the chat to agree on a joint number.[5] In fact, the latter quite often stated that the solution could not be calculated. This could be a hint for how to separate limited computation from limited reasoning (Camerer 2003). The "no discussion" teams obviously neither thought of themselves as being able to calculate the solution, nor did they try to rethink the instructions. The "discussion" teams on the other hand fully or at least partially understood the underlying reasoning process and adapted their choice based upon the expected behavior of others. The latter teams could therefore have a higher degree of reasoning because they strategically evaluate other players' behavior which is the most crucial element for being successful in the beauty contest. Building on this classification I derive the main hypothesis that strategic discussion has a positive impact on the performance of teams in the beauty contest.

Hypothesis 2. *Strategic discussion has a positive influence on the performance of teams.*

Strategic discussion could furthermore affect the performance of teams in the beauty contest in terms of distance to the target number, i.e. teams are more successful if they discuss their choice thoroughly and strategically (hypothesis 2.1), and in terms of team reasoning, i.e. strategic discussion enhances team reasoning (hypothesis 2.2).

[5] The two research assistants came to similar results concerning the classification of the teams as reported above (correlations significant on a $p < 0.01$ level). The further analysis of the data according to their evaluation resulted in the same results as reported in this paper.

3. The Beauty-Contest Game

3.1. *General Model and Game-Theoretic Solution*

The “p-beauty-contest” game was introduced by Moulin (1982) and first studied experimentally by Nagel (1995). It is called “beauty contest” after Keynes’s (1936) chapter about long-term expectation in “The General Theory of Employment, Interest, and Money”. Keynes compares professional investment with newspaper competitions in which people have to pick the six prettiest faces out of a hundred pictures of which they think that others will find the most beautiful. (Keynes 1936)

In a basic beauty contest game, n participants simultaneously choose a number from a closed interval between 0 and 100. The winner is the person i whose number x_i is closest to the target number, which is p times the mean of all chosen numbers where p is a predetermined and known number, e.g. $p = ½$. The target number can also be specified as $p * \sum_{i=1}^{n} \frac{x_i}{n}$. The winner receives a fixed amount of money whereas the other subjects receive nothing. If there is a tie, the prize is split between those subjects who tie. (Bosch-Domènech et al. 2002; Camerer 2003)

The beauty contest is a dominance solvable game. This means that by eliminating dominated strategies, a unique Nash equilibrium can be derived. The concept of iterated dominance is an important idea in game theory and especially in the beauty contest game as it enables the experimenter to study how many levels of iterated reasoning subjects actually apply. (Camerer 2003; Nagel 1999)

A rational player in the beauty contest does not simply choose a random number or his or her favorite one, nor does he or she choose a number above 100 times p because this would be dominated by any number smaller than $100\,p$. Given a $p = ½$ beauty contest, the average number chosen will never be above 100 due to the interval restriction. The target number, which is p times the average choice, will never be above 50. Choosing a number in the range of $50 < x_i \leq 100$ would violate dominance: no matter what other players do, a subject has a higher chance of winning the game if he or she chooses 50 or below. Moreover, if a player believes that the other subjects are also rational and obey dominance, hence never choose a number above 50, the player will not pick a number above 25 or $100\,p^2$ because the average will be no more than 50. All numbers in the range of $12.5 < x_i \leq 25$ therefore correspond to two steps of iterated dominance. Infinitely many steps of iterated elimination of dominated strategies finally lead to the unique Nash equilibrium in which all participants choose zero. In this way, the number choices of subjects reveal their level of iterated reasoning, which is why the behavior of participants in such a dominance-solvable game is an implicit measure of the extent to which subjects iteratively delete dominated strategies. Keynes (1936) already mentioned this quality of the beauty contest game to distinguish whether people "practice the fourth, fifth, and higher degrees" of reasoning. (Camerer 1997 pp. 178, 2003 pp. 199) But it is important to note that it would be a strategic mistake in this game to choose zero in the first round, because the goal is to be one step of reasoning ahead of the other players, not further. (Camerer 2003; Nagel 1999)[6]

There are several reasons why the beauty-contest game is an ideal tool to study reasoning processes and the rationality of different types of subjects, for instance teams and individuals. First of all, there is a clear distinction between bounded rationality and the game theoretic solution. Secondly, unlike mixed-motive games such as trust or ultimatum games, the beauty contest is generally a zero-sum game, so strategic factors can be separated from motivational ones such as fairness or cooperation, which are often invoked when payoff-maximizing solutions cannot describe the observed behavior. Nonstrategic aspects will therefore not matter when explaining violations of iterated dominance. (Bosch-Domènech et al. 2010; Nagel 1999) Thirdly, in contrast to other zero-sum games, different levels of reasoning can easily be structured and detected by measuring the number of iterated elimination of dominated strategies or by using other reasoning models such as the model of boundedly rational behavior by Nagel (1995). Yet another advantage of the beauty contest is the simplicity of the game. Not only are the rules very simple to explain, but the parameters of the game can easily be modified, for example by changing the underlying interval or the *p*-value to obtain a different equilibrium or by varying the payoff settings so that the beauty contest becomes a non-zero sum game. This flexibility in setting game parameters creates an ideal environment for researchers to explore the diverse questions at hand. Some of the obvious modifications of the basic game have already been explored (e.g. by Bosch-Domènech et al. 2002; Camerer 1997; Duffy and Nagel 1997; Güth et al. 2002; Grosskopf and Nagel 2008, 2009; Ho et al. 1998; Kocher and Sutter 2005; Kocher et al. 2006; Kocher and Sutter 2006; Nagel 1995, 1999, 2004; Slonim 2005; Sutter 2005).[7] Finally, this game can acutally demonstrate and explain how a false belief about other players' rationality can lead to a bad performance. If a subject expects others to only be "hyper-rational" players who have the same knowledge and expectations as the player himself, but instead faces bounded rational players, he will lose the game. (Nagel 2004)

[6] This connects game theory to other fields of social science, which study the beliefs people have about the beliefs and behavior of others. (Camerer 2003)
[7] For a more detailed overview of the beauty contest modifications see Güth et al. (2002) and Nagel (2004).

3.2. Static Model of Iterated Reasoning

The game-theoretic concept of iterated dominance is the simplest method to allocate the chosen numbers of subjects in a beauty contest to different levels of reasoning, as described in the previous section. Besides the game-theoretic solution, there are a number of methods which can be used to distinguish first-period choices in the beauty contest such as the model of boundedly rational behavior by Nagel (1995) and the model of iterated best reply based on Stahl and Wilson (1994; 1995), Ho et al. (1998) and Camerer et al. (2004). In this paper, the strategic sophistication of players is measured by using the model of iterated best reply such as that seen in Kocher and Sutter (2005). The model is applied as follows, using actual frequencies of chosen numbers. Players who apply zero levels of reasoning are assumed to choose numbers randomly from the given interval e.g. [0, 100] with the mean being 50. Level-1 players best reply to level-0 players by choosing numbers below $50\,p$. This means that numbers greater 25 are allocated to level-0 in the $p = ½$ beauty contest, whereas numbers from 25 to 12.5 are chosen by level-1 players. Players who apply higher levels of reasoning will best reply to players with lower depth of reasoning by choosing $50\,p^2$ or $50\,p^3$. For simplicity, numbers in the category from 0 to $50\,p^4$ are merged into level "> 3".

4. Design and Procedure

4.1. Experimental Design

Appendix C displays the complete experimental instructions.[8] The experimental design of both the team and control group treatment is reported in Table 1.

[8] The original instructions are in German and translated into English for the purpose of this paper.

Table 1

Experimental Design

	Beauty Contest 1	Beauty Contest 2
Team Treatment n=216	8 Individuals	2 Individuals 3 Teams à 2 Persons
Control Treatment n=48	8 Individuals	8 Individuals
Interval	[0, 100]	[0, 100]
Value of p	2/3	1/2
Target Number	$2/3 * (\sum x_i / n)$	$1/2 * (\sum x_i / n)$
Payoff	600 Taler = 6 €	600 Taler = 6 €

Notes. N=264.

In the first beauty contest each subject had to arrive at a decision on his or her own. Groups of eight subjects played together and the groups stayed the same for the second game. Each player had to choose a number from the interval [0, 100]. Decimal numbers were allowed but restricted to two decimal points.[9] The number closest to $p = ⅔$ times the average was the winning number. If there was a tie, the prize of 600 Taler (0.01 € = 1 Taler) was split between the candidates who tied. The first beauty contest was identical in both the team and the control group treatment. In the second beauty contest the subjects in the control treatment played again as individuals only, whereas the subjects in the team treatment either played as individuals or as members of a team. Six out of eight subjects were paired into teams of two players. As in Kocher and Sutter (2005) a small number of five decision makers in each beauty contest was chosen in order to obtain more observations on the performance of teams e.g. the frequency of winning. In contrast to the individual players, the teams had to choose a single number from the interval [0, 100] for both of them. For this purpose, they had the possibility to discuss via chat which number they would like to choose. If they both entered the same number on their individual screens, the game ended and they were immediately forwarded to the questionnaire. If, on the other hand, they disagreed and entered different numbers, they had two more possibilities to discuss and enter a new number.[10] Neither individuals, nor teams were faced with a time restriction when deciding upon a number. In the

[9] If only integers were allowed, several equilibria would be possible. For example, in the case of $p = ⅔$ there would also be an equilibrium with all choosing 1 in addition to the equilibrium with all players choosing 0. As ⅔ times 1 equals ⅔, a player choosing 0 instead of 1 would not win the game. (Bosch-Domènech et al. 2002)

[10] If the teams had not decided on a number after the third attempt, one of their numbers would have been randomly chosen for both of them by the computer. This element was adapted from Cooper and Kagel (2005). But only 11 out of 81 teams needed two trials to agree on a joint number and two teams needed three trials. Thus all chosen numbers by teams are numbers which both team members agreed on.

second beauty contest, either the single player or the team whose number was closest to $p = ½$ times the average number chosen won. In case of a tie, the prize was split between the subjects who tied. To keep the individual incentives constant among team and individual players, both team members received 600 Taler each in case of success, i.e. the payoff of a winning single player. With the exception of different p-values, the second beauty contest in the control group treatment was identical to the first one.

Before the subjects received feedback on their overall performance and payoff they were asked to fill in a short questionnaire which is displayed in Appendix C. The no-feedback condition was implemented to avoid possible learning effects between games and to prevent subjects from adapting their choice in the second beauty contest to the previously observed target number (Nagel 2004). To prevent any further influence from the first beauty contest game on the second, two different p-values were implemented. Both p-values were smaller than one for two reasons: Firstly, to be able to distinguish between different steps of iterated reasoning, which would not be possible with $p = 1$ (Duffy and Nagel 1997); secondly, a p-value smaller than 1 with a boundary equilibrium at 0 separates equilibrium from random play because subjects generally avoid extreme choices, and therefore only subjects who guess the game-theoretic equilibrium correctly will choose such a small number (Güth et al. 2002).

4.2. Experimental Procedure

The experiment was conducted over ten sessions. Each session lasted around 30 minutes, including an introductory talk by one of the experimenters. The subjects were recruited via e-mail, posters on campus, and advertisements in different courses. 264 subjects participated in the experiment. The students were mainly undergraduates (76.5%). The biggest groups comprised business administration and economics (26.6%), international cultural and business studies (19.7%) and education (12.5%). 91 male (34.5%) and 173 female students (65.5%) took part in the experiment. The experiment was computerized and conducted with the software z-Tree (Fischbacher 2007). One experimenter was in charge of controlling the server for the PCs while the other two experimenters acted as instructors. Upon arrival, the subjects had to draw a slip of paper from a box which either contained the letter “A” for room A or the letter “B” for room B. The sessions began with an introductory talk at the meeting point (“remain silent until completion”; “please raise your hand for technical support” etc.). The session started immediately after the subjects were seated in their allocated rooms. At the end of each session, the subjects received checks with their earned payoff which they could immediately turn into cash outside the rooms. The average payoff was 6.93 € including a show-up fee of 2 €. The whole experiment contained four different games: a dictator game, an ultimatum game and two beauty contests. The games were clearly separated by a welcome screen announcing that a new game was about to start. As already mentioned, the beauty contest has the ability to clearly separate strategic play from other regarding behavior such as trust and reciprocity. Moreover, subjects were randomly re-matched after each game and feedback was only provided after the whole experiment was completed. Therefore the behavior of subjects in the beauty contest games can be considered as not having been influenced by what happened in the previous games. 216 of the 264 subjects played in the team treatment and 48 in the control treatment. 162 out of the 216 subjects in the team treatment acted as teams of two in the second beauty contest and 54 out of the 216 subjects acted as individuals. Eight out of the overall ten sessions contained the team treatment and two contained the control group treatment with single players only. Thus 264 individual observations from the first beauty contest were gathered, while 81 team observations, 54 individual and 48 control group observations from the second beauty contest could be collected.

5. Experimental Results and Analysis

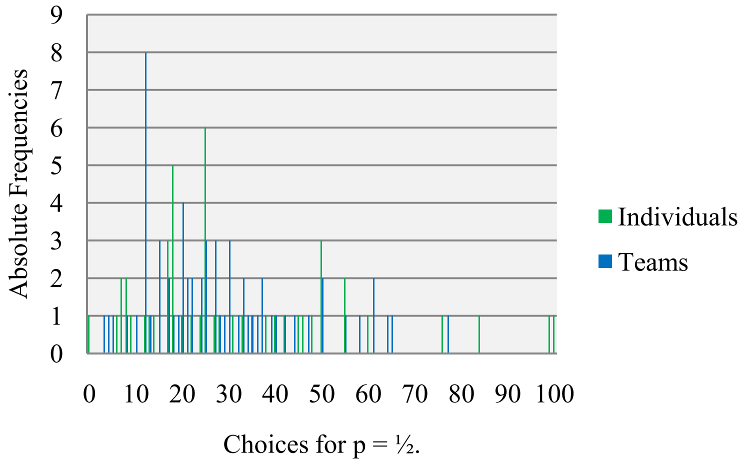
In the following section, the experimental results will be reported and it will be discussed whether or not they support the hypotheses. Beauty Contest 1 will be referred to as BCG 1, while Beauty Contest 2 will be referred to as BCG 2.

5.1. Are Teams More Rational than Individuals?

As previous studies such as Nagel (1995) and Ho et al. (1998) have shown, choices in the first round of a beauty contest are generally far away from the equilibrium choice of 0. The overall distribution of choices in BCG 2 is illustrated in Figure 1 which pictures the choices of teams and individuals.

Figure 1

Distribution of Choices in Beauty Contest 2



Notes. Teams chose numbers with a mean of 27.62, a median of 24.00 and a standard deviation of 16.17. Individuals chose numbers with a mean of 31.81, a median of 25.00 and a standard deviation of 22.42.

As illustrated by Figure 1, teams seem to choose numbers within a smaller range than individuals, with the latter choosing numbers from 0 to 100 whereas the maximum chosen number within the group of teams was only 77.08. Table A1 in Appendix A displays the average chosen numbers, the minimum, maximum and median numbers, as well as the standard deviation and the variance of chosen numbers. Whereas one individual player actually chose the Nash equilibrium 0.00, the minimum chosen number within the group of teams was 3.00. On the other hand, one individual player also chose 100.00 in BCG 2 while the maximum number within the teams was 77.08. Meanwhile, the median numbers are almost similar: 24.00 for teams and 25.00 for individuals. The variances show quite a different picture: teams had a variance of 261.35 in BCG 2, whereas the individuals had a much higher variance of 502.81. This obvious difference is confirmed by the Levene's test which results in a significant value of $p = 0.022$. Hence, the variance of team choices in BCG 2 is considerably smaller than the variance of individual choices. This could be an indication that communication in teams can lead to a minimizing of outlier effects or can balance biases, as suggested by Davis (1992). For example, subjects who previously chose very large numbers tend to, as a result of communication within teams, choose lower numbers. A comparison to the control group of individuals yields a different explanation: the control group choices range between 2.50 and 75.00. The variance within the control group was only 259.35, which is actually smaller than the variance of the teams. This leads to the conclusion that the individuals in the team treatment choose within an exceptionally wide range of numbers and that the smaller variance of teams cannot be attributed to the effect of communication within teams.[11] Compared to BCG 1 (see Figure A1 in Appendix A) the choices are more concentrated in both groups around 15 to 25, which is not surprising due to the different *p*-values. Nevertheless, it demonstrates that subjects are clearly influenced by the relevant value of *p* (Nagel 1995).

To answer the question as to whether or not teams are closer to the equilibrium value of zero, the average chosen numbers are taken into account. They are quite similar among the two groups: teams of two (N = 81) chose on average 27.62 and individuals (N = 54) 31.81. A Student's t-test was chosen to detect possible differences between teams and individuals. The t-test (two-sided) resulted in a non-significant value of $p = 0.240$ which indicates that teams of two did not choose significantly lower numbers in BCG 2. Consequently, hypothesis 1.1 cannot be confirmed: teams are no closer than individuals to the game-theoretic equilibrium.

Although teams and individuals do not differ with respect to their chosen numbers, the question remains as to whether or not teams significantly outperform individuals in the beauty contest. The absolute indicator for performance and success in the beauty contest is the frequency of winning. Having a closer look at the winning rates reveals the following picture: Teams won the beauty contest 16 times (57.14%), whereas individual players won the beauty contest 12 times (42.86%). A Student's t-test was chosen to analyze the difference in winning between teams and individuals. With a probability of 73.1%, the difference between individuals and teams is random and therefore the variables ("won" and "teams/individuals") are independent. This leads to the overall conclusion that hypothesis 1.2 cannot be confirmed. Teams are not considered to be more successful than individuals because there is no significant difference in winning the beauty contest.

Given the above results, hypothesis 1 cannot be confirmed because teams are neither more rational concerning their distance to the game-theoretic equilibrium, nor are they more successful than individuals.

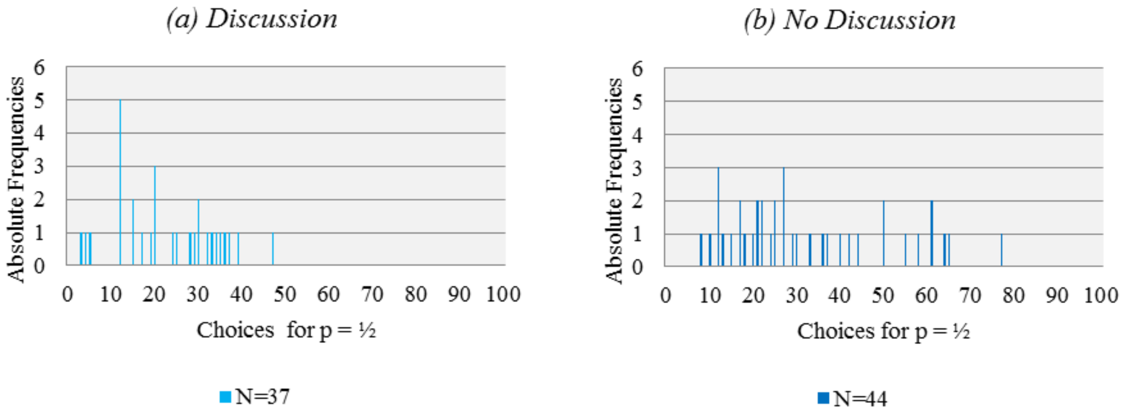
5.2. Strategic Discussion

To evaluate whether or not strategic discussion has an influence on the performance of teams the chosen numbers of teams and their distance to the target value have to be considered. Figure 2 displays the overall distribution of chosen numbers separately for strategically discussing teams (2a) and teams in which the choice was not discussed (2b).

[11] The means of chosen numbers did not significantly differ between any of the three groups of players. They also did not differ in their individual ability as measured by their individual reasoning level in BCG 1.

Figure 2

Distribution of Teams' Choices



Notes. "Discussion" teams chose numbers with a mean of 21.79, a median of 20.00 and a standard deviation of 10.80. "No Discussion" teams chose numbers with a mean of 32.53, a median of 26.00 and a standard deviation of 18.38.

As Figure 2 indicates, strategically discussing teams chose numbers in a much smaller range than the "no discussion" teams. The variance of chosen numbers within the latter group was 337.74, with a maximum chosen number of 77.08. The variance within the "discussion" group was 116.65, with a maximum chosen number of 47. A Levene's test confirms this difference to be significant on a $p < 0.01$ level. This indicates that it is not communication per se but rather strategic discussion within teams that reduces outlier effects and determines the performance of teams in a strategic setting.

In order to depict a clearer picture of the effect that strategic discussion has on the performance of teams the absolute distances to the target number for both strategically discussing teams and "no discussion" teams are shown in Table 2. The distance to the target number provides an indication for the performance in the beauty contest because it determines whether a subject is winning the beauty contest or at least how close a subject was to correctly anticipating the behavior of the other players. The values in Table 2 are grouped into six categories ranging from a distance to the target value smaller than 5 to a distance greater than 25. The results show that over 60% of the strategically discussing teams chose numbers that deviated by 0 to 10 from the target value whereas only 38.7% of the "no discussion" teams deviated by less than 10 digits.

Table 2

Performance in the Beauty Contest

Distance to Target Value	Discussion (N=37)	No Discussion (N=44)	Total (N=81)
00.00 - 05.00	0.351	0.205	0.272
05.01 - 10.00	0.270	0.182	0.222
10.01 - 15.00	0.162	0.159	0.160
15.01 - 20.00	0.135	0.091	0.111
20.01 - 25.00	0.027	0.068	0.049
> 25	0.054	0.295	0.185
			1.000
Mean	8.99	18.29	
Median	7.55	12.92	
Std. Deviation	7.47	15.34	

Notes. Percentages are calculated per group and in total for each category. The values are grouped into six categories ranging from a distance to the target value smaller than 5 to values greater than 25.

The mean values underline this difference: The mean distance to the target number within the "discussion" teams is 8.99; the mean distance within the "no discussion" teams is 18.29. A Student's t-test confirms the difference to be significant on a $p < 0.01$ level (t-test, df = 67.27, p = 0.001). Given this evidence, it is obvious that strategically discussing teams are closer to the target value and are therefore considered to be more successful in the beauty contest.[12]

The question now is if strategic discussion significantly influences the performance of teams, or if the amount of information exchanged[13] impacts the reasoning process within teams as well. There are grounds for the assumption that the quantity of discussion positively influences the performance if and only if the teams do qualitatively discuss their choice, i.e. it does not matter how much you talk but what you talk about. To detect the influences of both strategic discussion and the amount of information exchanged on the distance to the target value I ran two regressions which are displayed in Table 3.

[12] The same result holds true for the distance to the game-theoretic equilibrium: the mean distance to the equilibrium is 21.79 for the "discussion" teams and 32.53 for the "no discussion" teams (t-test, df = 71.224, p = 0.002).

[13] The amount of information exchanged is determined by the amount of letters (empty spaces included) exchanged in the chat.

Table 3

Strategic Discussion as Mediator

	Distance to Target Value (DV)			
Variables	Model 1		Model 2	
Constant	18.212 ***	(2.294)	19.527 ***	(2.509)
Amount of Information Exchanged	-0.017 ***	(0.006)	-0.008	(0.005)
Discussion			-7.321 **	(2.713)

Notes. Robust standard errors in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Model 1 in Table 3 predicts that the amount of information exchanged significantly influences the distance to the target value in the beauty contest, i.e. the performance (p = 0.009). Hence, the more teams wrote in the chat, the better they performed in terms of distance to the target value. Once strategic discussion is included in the model (see Model 2 in Table 3) the result changes; the influence of the amount of information exchanged vanishes. Strategic discussion thus fully mediates the influence of the amount of information exchanged on the performance in the beauty contest. To prove this mediating effect and to get further results on the direct and indirect effect of the amount of information exchanged on the performance I conducted a mediation test based on Preacher and Hayes (2004; 2008). Both the mediation test and the regressions yield the same result: the direct effect of the amount of information exchanged on the performance is insignificant (p = 0.225). But the total effect of the amount of information exerted through strategic discussion significantly influences the performance in the beauty contest (Coeff = -0.0175, se = 0.0065, p = 0.009). Hence, the amount of information exchanged positively influences the performance of teams if, and only if, the team members discuss their choice strategically. Hypothesis 2.1 can therefore be confirmed.

Before a conclusion can be drawn as to whether or not strategic discussion has an impact on the overall performance of teams (hypothesis 2), it is necessary to assess the influence of strategic discussion on team reasoning. Figure A2 (see Appendix A) displays both the individual reasoning level of subjects who later on played in teams (Figure 4a - BCG 1) and the reasoning level of teams in the second beauty contest (Figure 4b - BCG 2). The difference in depth of reasoning within the group of teams in BCG 2 is quite obvious: The mean reasoning level of strategically discussing teams was 1.06, whereas the “no discussion” teams had a mean of just 0.62. Moreover, none of the latter teams were rational to a degree greater than 2, whereas the strategically discussing teams were also represented in reasoning level 3 and greater than 3. A non-parametrical χ^2-test was chosen to assess whether this difference was significant. The χ^2-test produced a level of significance of p = 0.014, i.e. teams which use the chat to strategically discuss their choice are significantly different from the other teams in the sense that they have a higher reasoning level ($\chi^2 = 12.568$; df = 4; p = 0.014). Thus, hypothesis 2.2 can be confirmed: Strategic discussion enhances the reasoning of teams. The question remains as to whether or not this result can be attributed to strategic discussion only, or if the more rational or smarter individuals were randomly assigned into the strategically discussing teams. To control for the subjects' individual ability, or individual rationality, and to test this assumption, the first beauty contest was implemented. Figure 4a illustrates the results of BCG 1. The mean reasoning level of individuals who later on played in strategically discussing teams was 0.64; the mean of the other individuals was 0.66. Once more, a non-parametrical χ^2-test was chosen to analyze the differences in reasoning between the individual players. The test suggests the difference to be insignificant on a $p > 0.1$ level, i.e. the subjects were not per se significantly different ($\chi^2 = 2.702$; df = 4; p = 0.609) which basically means that the more rational individuals were not randomly assigned into "discussion" teams.

The stated results lead to the overall conclusion that hypothesis 2 can be confirmed: strategic discussion does indeed have a positive impact on the performance of teams in the beauty contest game.

6. Conclusion

This paper has addressed the question as to whether or not teams are more rational than individuals in an experimental beauty contest and whether strategic discussion within teams has any influence on their performance. Even though there have been various papers on the beauty contest game and on communication in experimental settings this is the first one which focuses on strategically relevant communication within teams in the beauty contest.

My findings concerning the comparison of teams and individuals in one-shot games underline Kocher and Sutters' (2005) results as I too did not find any difference between teams and individuals in first-round behavior: teams are neither one step ahead in reasoning, nor do they outperform individuals in terms of payoff or distance to the target value. There is simply no significant difference between the behavior of teams and individuals. Kocher and Sutter did, however, find weak evidence that the variance of groups is smaller than that of individuals. The team treatment in this paper yielded the same result concerning the variance of teams, but when compared to the control group proved that the difference was a random occurrence.

However, having a closer look at the group of teams I found that strategic discussion within teams not only improves their performance in the beauty contest but also enhances their team reasoning. Teams who used the chat to strategically discuss their choice were more successful in correctly guessing the average chosen number which is equivalent to succeeding in the beauty contest. So even if they share more than one thought, they can discuss their controversial view points and possibly decide on a strategy that benefits both. The experimental results have shown that strategic discussion is indeed the key to success in the beauty contest.

The results deliver a valuable contribution for the design of adequate support systems for team work and could benefit teams in other strategic situations and in real life. Working teams should not only have the possibility to communicate per se but they should be trained and incentivized to focus on strategic aspects when communicating to employ the full range of their abilities. Additionally, research on the quality of communication in groups by Schoop et al. (2010) notes that studies on negotiations often forget to include the complex aspects of communication and assess negotiations based only on their economic outcome. However, they argue that the quality of communication on its own affects economic outcomes in the short- and long-term. This study addresses this argument perfectly. Firstly, it demonstrates that it is indeed the qualitative aspect of communication – that is the strategic discussion – which determines the performance of teams in terms of the distance to the target value. And secondly, it shows that strategic discussion helps teams to actually increase their reasoning level, something which is not necessarily achieved through a simple exchange of words.

My findings yield interesting results with regard to teams that deliver potential for additional research. As Sutter (2005) pointed out, teams of four members perform significantly better than individuals or teams of two. Hence, it would be interesting to combine varying group sizes and discussion to examine whether communication is a helpful tool to improve the reasoning process only up to a certain group size. Additional work should also be done concerning the team behavior in repeated games to see whether repetition allows teams to employ the full range of their joint knowledge. Other future research could be conducted regarding the advantage of communication in other strategic scenarios and the decision process within teams. For example: do teams agree on a joint number because they have to compromise or because they can exchange diverging tendencies?

Appendix A: Additional Tables and Results

Figure A1 displays the distribution of choices in the first beauty contest game with individual players only.

Figure A1

Distribution of Choices in Beauty Contest 1

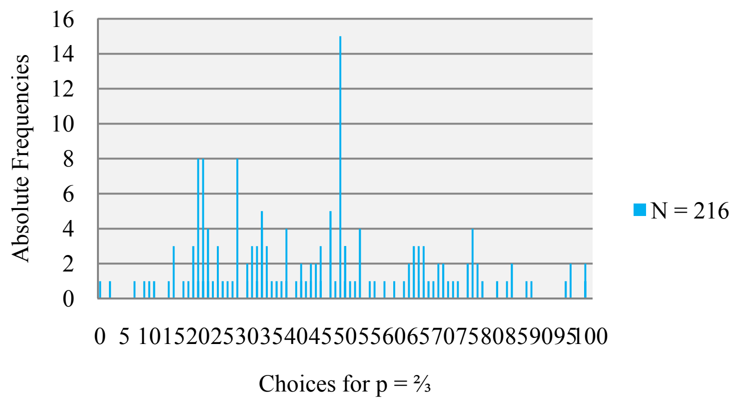


Figure A2 illustrates both the individual and team reasoning for the 162 team players in BCG 1 (a) and BCG 2 (b). The first beauty contest was administered to control for the individual ability of participants who acted in teams of two in the second beauty contest.

Figure A2

Levels of Reasoning

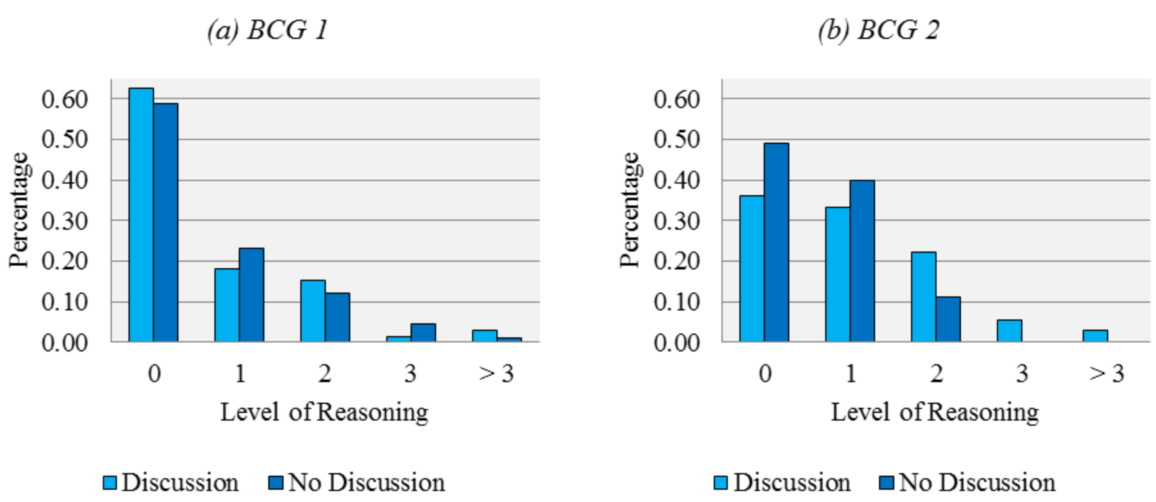


Table A1 displays the overall results for the team (a) and control group (b) treatment both for the first (BCG 1) and second (BCG2) beauty contest game.

Table A1

Overall Results

(a) Team Treatment

	BCG 1			BCG 2		
	Team	Individuals	Total	Team	Individuals	Total
Mean	43.74	44.22	43.86	27.62	31.81	28.67
Median	42.01	40.35	41.92	24.00	25.00	24.80
Minimum	2.55	9.30	2.55	3.00	0.00	0.00
Maximum	99.10	100.00	100.00	77.08	100.00	100.00
Std. Deviation	20.85	25.62	22.08	16.17	22.42	17.97
Variance	434.67	656.23	487.31	261.35	502.81	322.96

Notes. N=216.

(b) Control Group Treatment

	BCG 1	BCG 2
Mean	45.50	27.02
Median	48.77	25.25
Minimum	0.00	2.50
Maximum	90.00	75.00
Std. Deviation	22.55	16.10
Variance	508.63	259.35

Notes. N=48.

Appendix B: Chat Protocols

Two examples of the chat protocols are displayed below, corresponding to the two possible classifications of strategic discussion (“no discussion” and “discussion”). The chats are translated into English and include the original spelling and typos of the subjects as far as possible in order to illustrate the original conversations. The first of the displayed conversations was coded as “no discussion” by the two research assistants. The second one, classified as “discussion”, shows one of the sophisticated conversations since the subjects talked about every aspect of the game in the sense of repeating the rules, assuming what other players might do and choosing their own number correspondingly. Since subjects could simultaneously write to each other the messages sometimes overlapped in the time being sent which is displayed in chat protocol 2.

Chat Protocol 1 – “No Discussion”

Player’s ID	Message
1	20?
2	good
1	OK

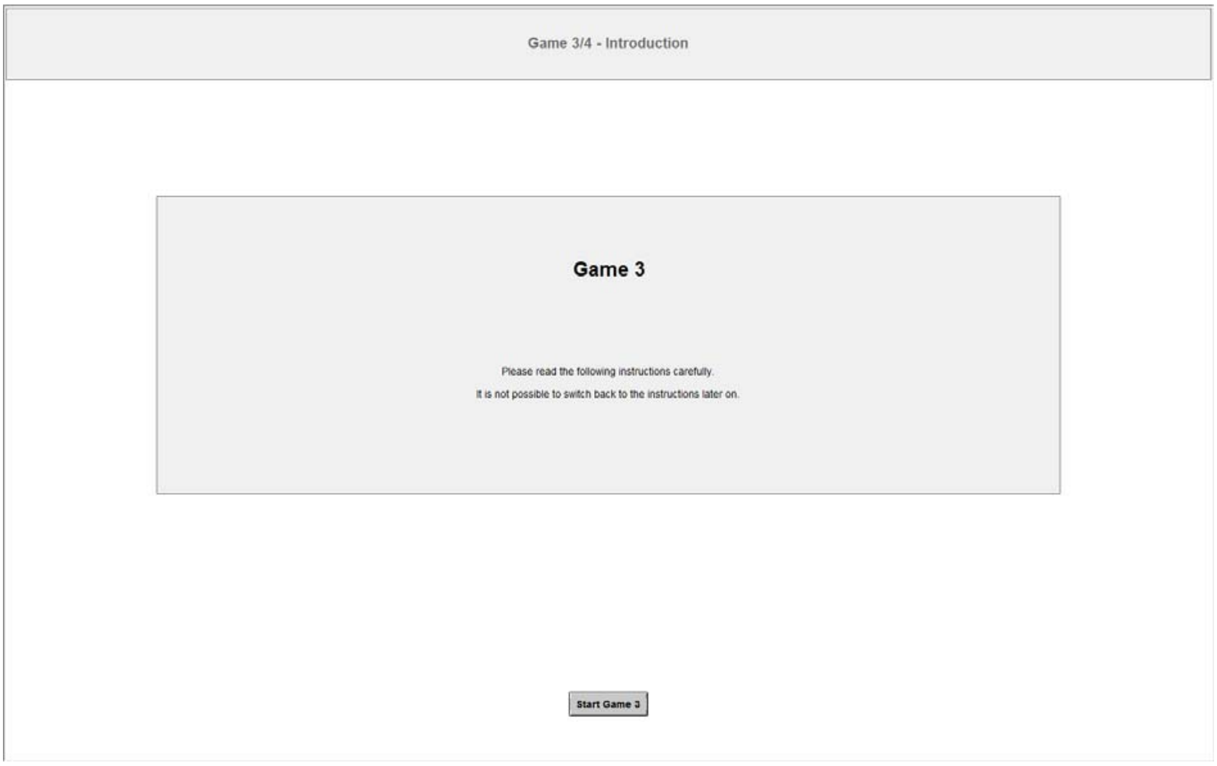
Chat Protocol 2 – “Discussion”

Player’s ID	Message
1	hello?
2	hi
1	my first thoughts are…
1	in an experiment with mathematicians only I would suggest that the solution would be 0
2	the number has to be smaller than 50
1	but there are certainly not only mathematicians here… so, any ideas? :)
1	yep
1	if everybody is thinking the number has to be smaller than 50, everybody is inserting 25 etc.
2	smaller than 35?
1	yes, for sure
1	it depends on how often the others think outside the box
1	if the average is at 50, then the half is at 25
2	how many are going to enter a high number?
2	sure
1	if everybody is saying 25, then the half is 12.5
1	i don’t know how many are entering a high number
1	what do you think about 6?
2	too small
1	your suggestion?
2	maybe even something higher than 12.5, say 15?
1	ok, we can do that
1	so 15?
1	fine by me
2	good
1	ok, was nice with you :)
2	i am entering it
1	me too
2	mkt you too :_)

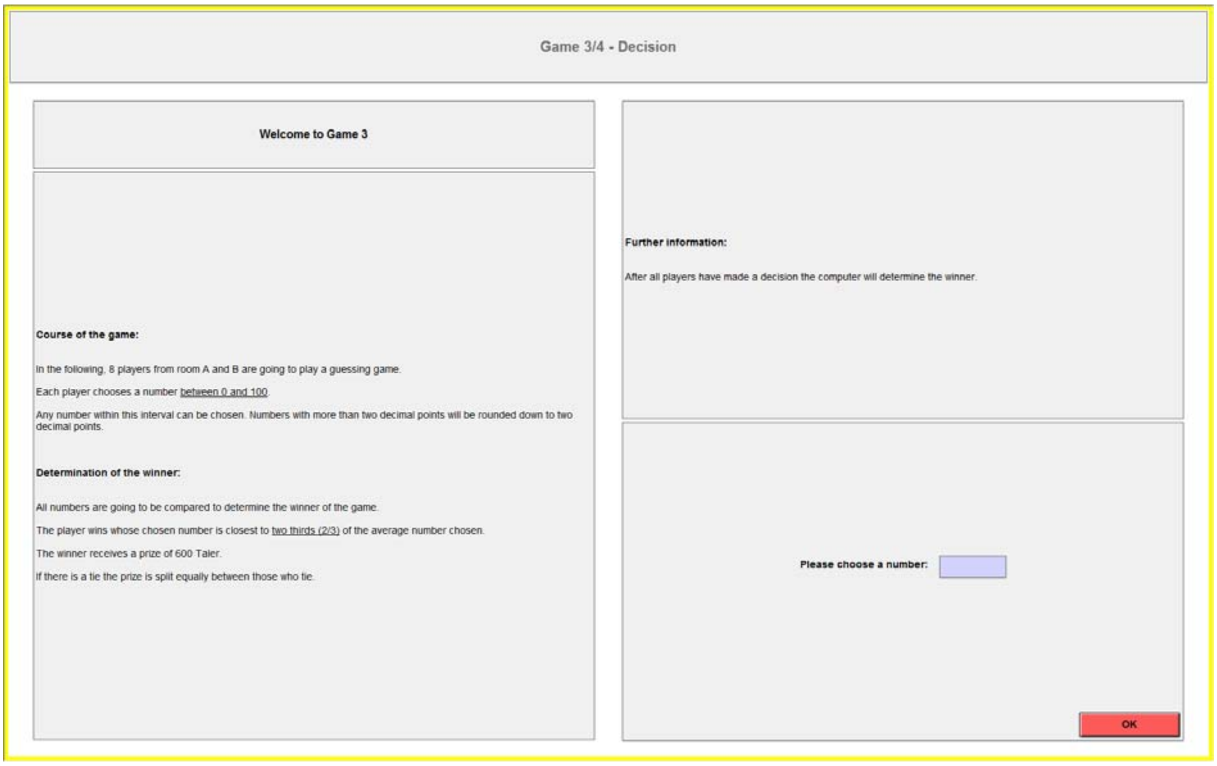
Appendix C: Experimental Instructions

Team Treatment – Instructions for Team Players

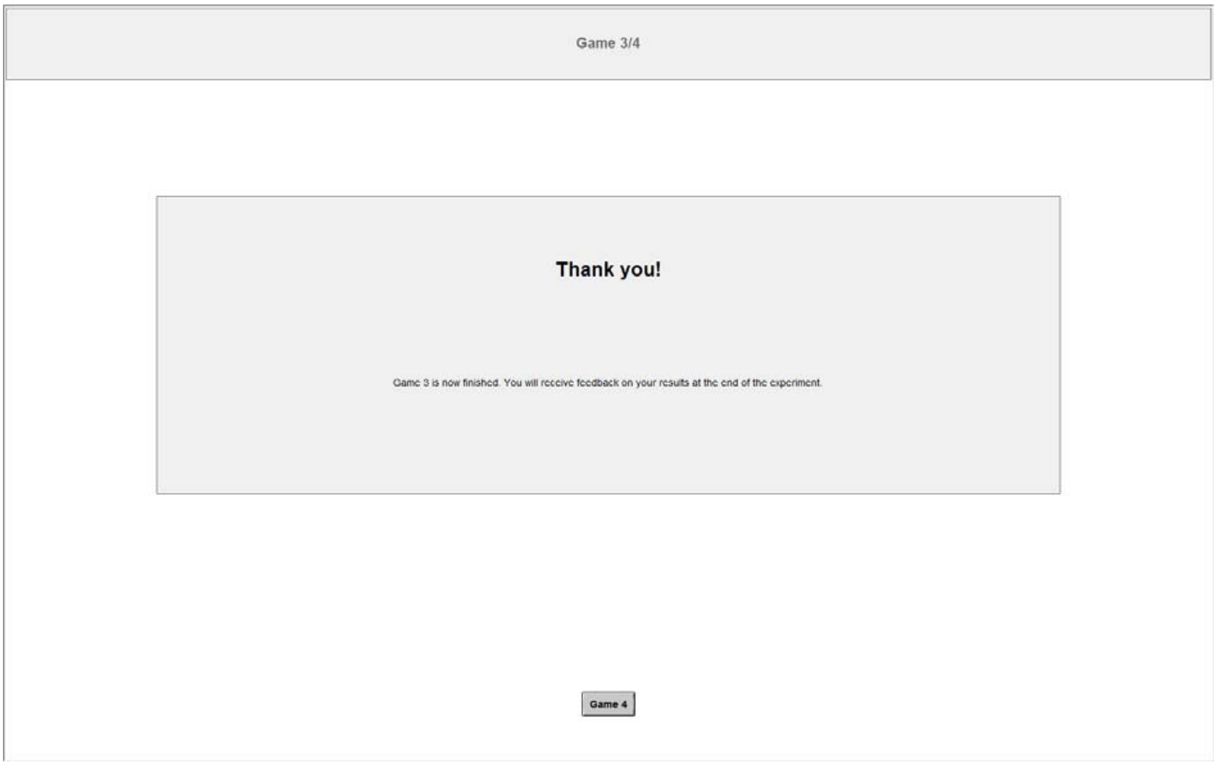
Screen Preceding Beauty Contest 1



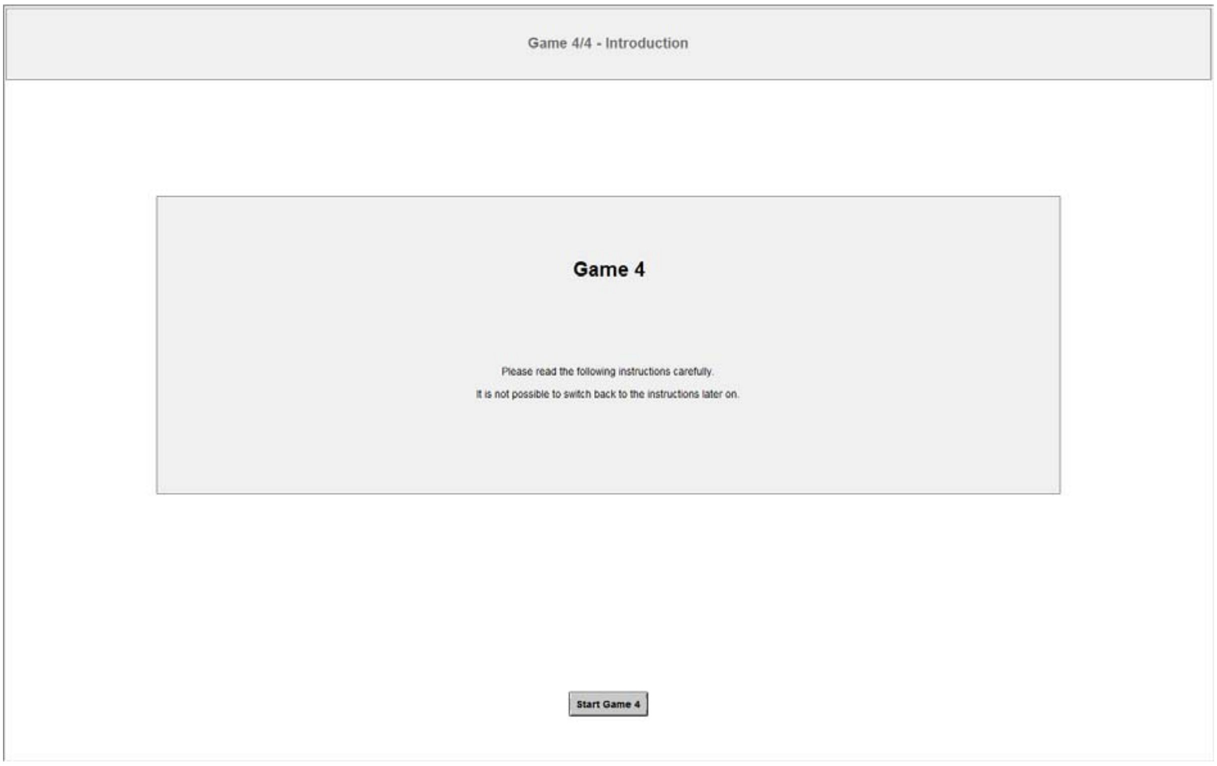
Beauty Contest 1 – Decision



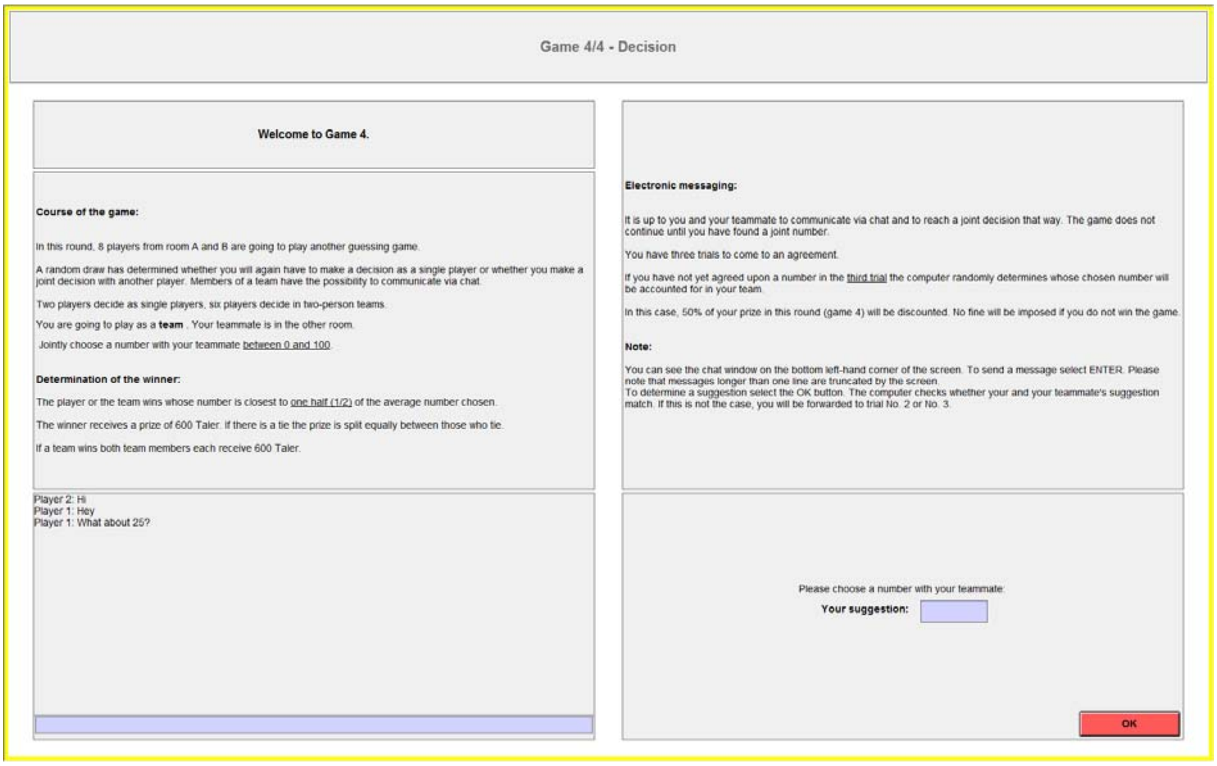
Screen after Beauty Contest 1



Screen Preceding Beauty Contest 2



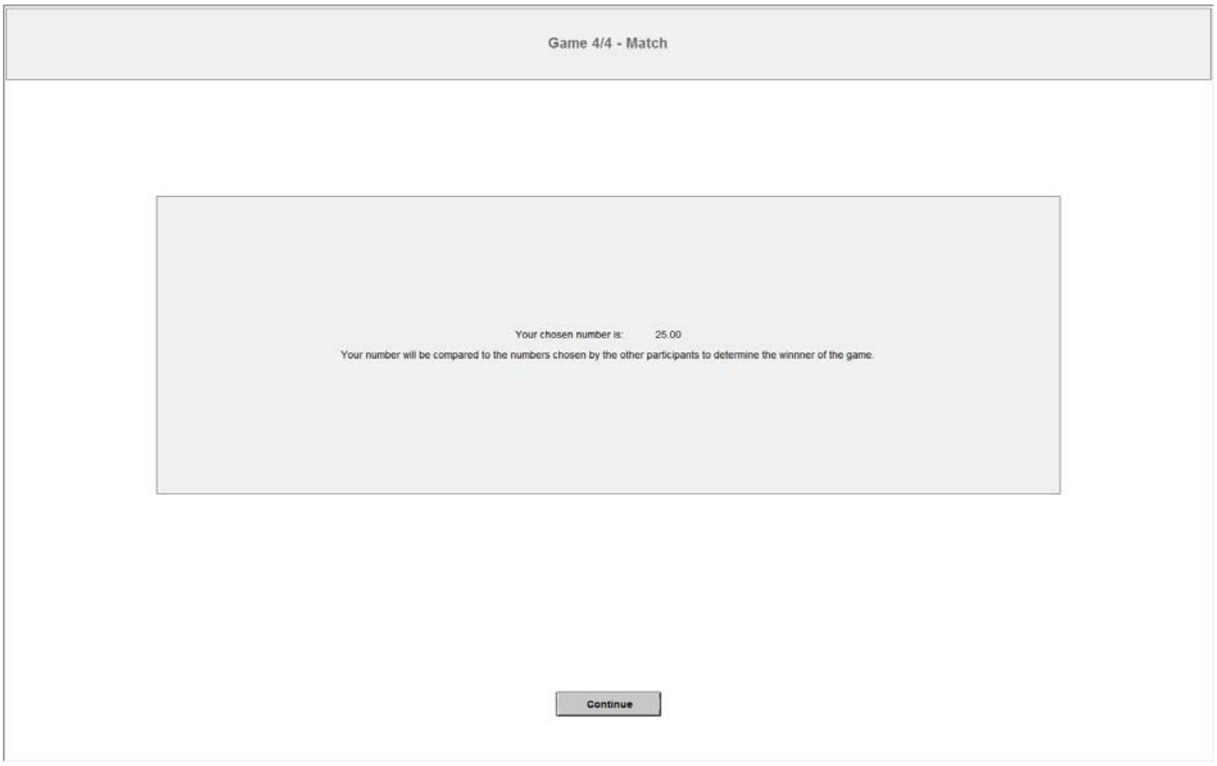
Beauty Contest 2 – Team Decision



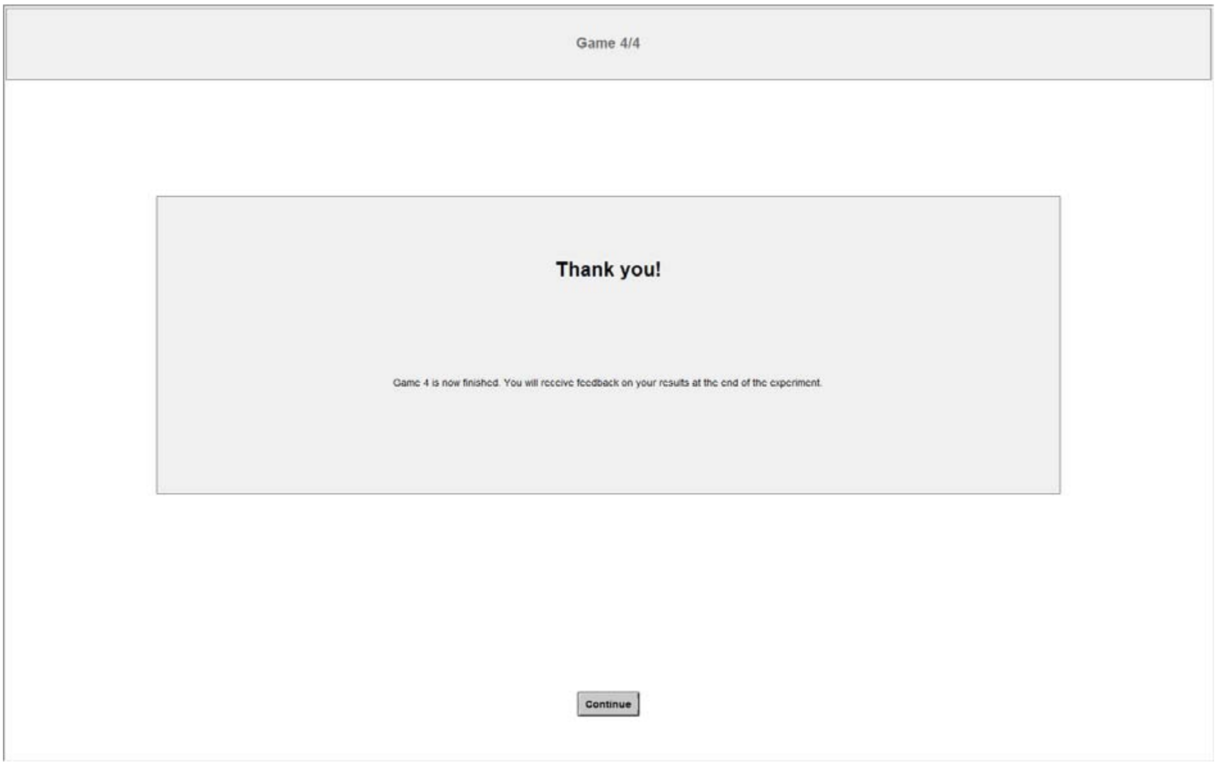
Beauty Contest 2 – Trial No. 2

Beauty Contest 2 – Trial No. 3

Beauty Contest 2 – Team Choice



Screen after Beauty Contest 2



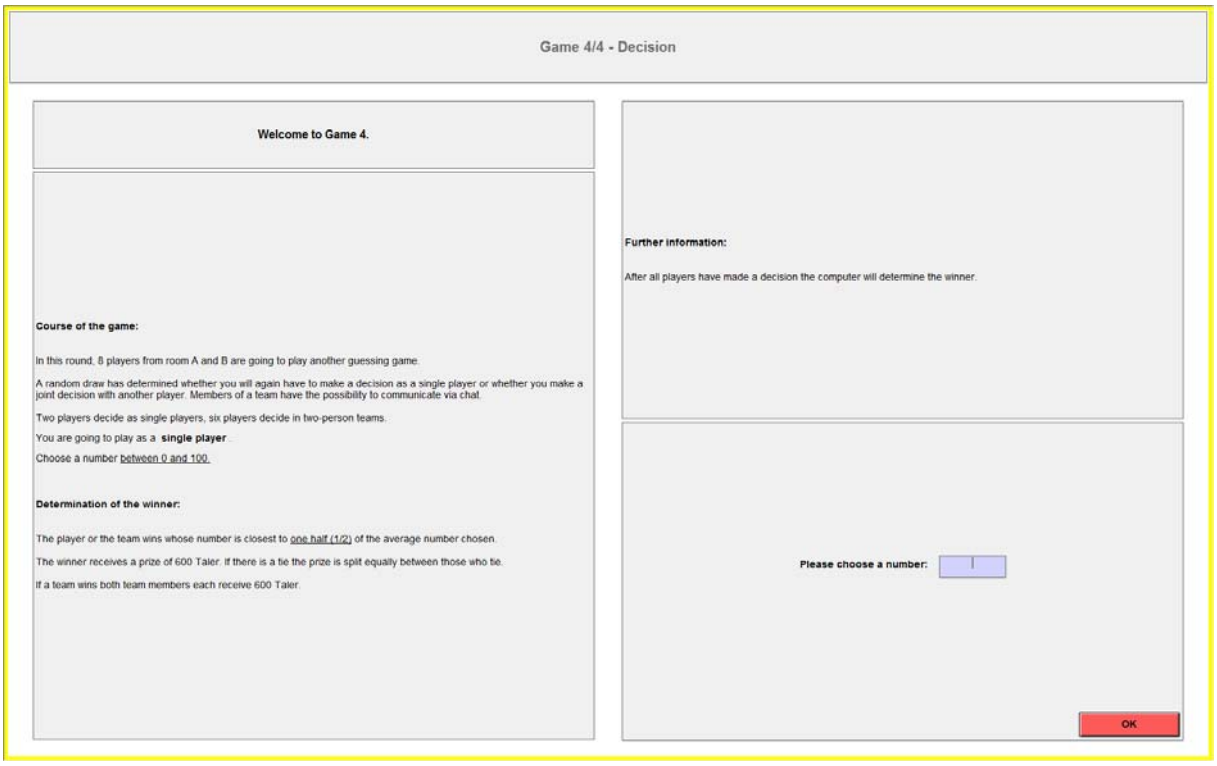
Questionnaire

Results

Team Treatment – Instructions for Single Players

Note: All other screenshots are equivalent to instructions for the team players.

Beauty Contest 2 – Individual Decision



Control Treatment – Instructions

Note: The instructions for the control treatment are the same as in the team treatment. They differ only concerning the group composition in the second beauty contest game. The instructions in the control treatment informed the subjects that eight individual players are going to play another guessing game.

References

Blinder, AS (2008). Leadership in groups: A monetary policy experiment. International Journal of Central Banking, 4 (4), 117–150.

Blinder, AS & Morgan, J (2005). Are two heads better than one? Monetary policy by committee. Journal of Money, Credit and Banking, 37 (5), 789–811.

Bochet, O, Page, T & Putterman, L (2006). Communication and punishment in voluntary contribution experiments. Journal of Economic Behavior & Organization, 60 (1), 11–26.

Bone, J, Hey, J & Suckling, J (1999). Are groups more (or less) consistent than individuals? Journal of Risk and Uncertainty, 18 (1), 63–81.

Bornstein, G, Kugler, T & Ziegelmeyer, A (2004). Individual and group decisions in the centipede game: Are groups more ‘‘rational’’ players? Journal of Experimental Social Psychology, 40 (5), 599–605.

Bornstein, G & Yaniv, I (1998). Individual and group behavior in the ultimatum game: Are groups more "rational" players? Experimental Economics, 1 (1), 101–108.

Bosch-Domènech, A, Montalvo, JG, Nagel, R & Satorra, A (2002). One, two, (three), infinity,…: newspaper and lab beauty-contest experiments. The American Economic Review, 92 (5), 1687–1701.

Bosch-Domènech, A, Montalvo, JG, Nagel, R & Satorra, A (2010). A finite mixture analysis of beauty-contest data using generalized beta distributions. Experimental Economics, 13 (4), 461–475.

Brandts, J & Cooper, DJ (2007). It's what you say, not what you pay: An experimental study of manager–employee relationships in overcoming coordination failure. Journal of the European Economic Association, 5 (6), 1223–1268.

Brosig, J, Weimann, J & Ockenfels, A (2003). The effect of communication media on cooperation. German Economic Review, 4 (2), 217–241.

Buchan, NR, Johnson, EJ & Croson, RTA (2006). Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. Journal of Economic Behavior & Organization, 60 (3), 373–398.

Burchardi, KB & Penczynski, SP (2012). Out of your mind: Eliciting individual reasoning in one shot games. Working Paper.

Camerer, CF (1997). Progress in behavioral game theory. Journal of Economic Perspectives, 11 (4), 167–188.

Camerer, CF (2003). Behavioral Game Theory. Princeton: Princeton University Press.

Camerer, CF, Ho, T & Chong, J (2004). A cognitive hierarchy model of games. Quarterly Journal of Economics, 119 (3), 861–898.

Cason, TN & Mui, V (1997). A laboratory study in group polarisation in the team dictator game. The Economic Journal, 107 (444), 1465–1483.

Charness, G & Dufwenberg, M (2006). Promises and partnership. Econometrica, 74 (6), 1579–1601.

Charness, G & Dufwenberg, M (2010). Bare promises: An experiment. Economics Letters, 107 (2), 281–283.

Cheung, SL & Palan, S (2012). Two heads are less bubbly than one: team decision-making in an experimental asset market. Experimental Economics, 15 (3), 373–397.

Cooper, DJ & Kagel, JH (2005). Are two heads better than one? Team versus individual play in signaling games. The American Economic Review, 95 (3), 477–509.

Cooper, R, DeJong, DV, Forsythe, R & Ross, TW (1989). Communication in the battle of the sexes game: Some experimental results. The RAND Journal of Economics, 20 (4), 568–587.

Cooper, R, DeJong, DV, Forsythe, R & Ross, TW (1992). Communication in coordination games. Quarterly Journal of Economics, 107 (2), 739–771.

Cox, JC (2002). Trust, reciprocity, and other-regarding preferences: groups vs. individuals and males vs. females. In R Zwick, A Rapoport (Eds.): Experimental Business Research, 331–350. Boston MA: Springer US.

Cox, JC & Hayne, SC (2005). When does an incentive for free riding promote rational bidding? In A Rapoport, R Zwick (Eds.): Experimental Business Research, 133–149. Boston MA: Springer US.

Cox, J & Hayne, SC (2006). Barking up the right tree: Are small groups rational agents? Experimental Economics, 9 (3), 209–222.

Davis, JH (1992). Some compelling intuitions about group consensus decisions, theoretical and empirical research, and interpersonal aggregation phenomena: Selected examples 1950–1990. Organizational Behavior and Human Decision Processes, 52 (1), 3–38.

Dawes, RM, McTavish, J & Shaklee, H (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. Journal of Personality & Social Psychology, 35 (1), 1–11.

Duffy, J (2002). Do actions speak louder than words? An experimental comparison of observation and cheap talk. Games and Economic Behavior, 39 (1), 1–27.

Duffy, J & Nagel, R (1997). On the robustness of behaviour in experimental `beauty contest' games. The Economic Journal, 107 (445), 1684–1700.

Farrell, J & Rabin, M (1996). Cheap talk. Journal of Economic Perspectives, 10 (3), 103–118.

Fiedler, M (2009). Cooperation in virtual worlds. Schmalenbach Business Review, 61, 172–194.

Fiedler, M & Haruvy, E (2009). The lab versus the virtual lab and virtual field—An experimental investigation of trust games with communication. Journal of Economic Behavior & Organization, 72 (2), 716–724.

Fischbacher, U (2007). z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics, 10 (2), 171–178.

Frohlich, N & Oppenheimer, J (1998). Some consequences of e-mail vs. face-to-face communication in experiment. Journal of Economic Behavior & Organization, 35 (3), 389–403.

Gillet, J, Schram, A & Sonnemans, J (2009). The tragedy of the commons revisited: The importance of group decision-making. Journal of Public Economics, 93 (5-6), 785–797.

Grosskopf, B & Nagel, R (2008). The two-person beauty contest. Games and Economic Behavior, 62 (1), 93–99.

Grosskopf, B & Nagel, R (2009). Unraveling in two–person guessing games. Working Paper.

Güth, W, Kocher, MG & Sutter, M (2002). Experimental 'beauty contests' with homogeneous and heterogeneous players and with interior and boundary equilibria. Economics Letters, 74, 219–228.

Ho, T, Camerer, CF & Weigelt, K (1998). Iterated dominance and iterated best response in experimental "p-beauty contests". The American Economic Review, 88 (4), 947–969.

Houser, D & Xiao, E (2011). Classification of natural language messages using a coordination game. Experimental Economics, 14 (1), 1–14.

Keynes, JM (1936). The general theory of employment, interest, and money. London: Macmillan.

Kocher, MG, Strauß, S & Sutter, M (2006). Individual or team decision-making - Causes and consequences of self-selection. Games and Economic Behavior, 56 (2), 259–270.

Kocher, MG & Sutter, M (2005). The decision maker matters: individual versus group behaviour in experimental beauty-contest games. The Economic Journal, 115 (500), 200–223.

Kocher, MG & Sutter, M (2006). Time is money—Time pressure, incentives, and the quality of decision-making. Journal of Economic Behavior & Organization, 61 (3), 375–392.

Kocher, MG & Sutter, M (2007). Individual versus group behavior and the role of the decision making procedure in gift-exchange experiments. Empirica, 34 (1), 63–88.

Krippendorff, K (2013). Content analysis. Thousand Oaks California: Sage Publications.

Kugler, T, Bornstein, G, Kocher, MG & Sutter, M (2007). Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy. Journal of Economic Psychology, 28 (6), 646–657.

Lombardelli, C, Proudman, J & Talbot, J (2005). Committees versus individuals: an experimental analysis of monetary policy decision making. International Journal of Central Banking, 1 (1), 181–205.

Luhan, WJ, Kocher, MG & Sutter, M (2009). Group polarization in the team dictator game reconsidered. Experimental Economics, 12 (1), 26–41.

Maciejovsky, B & Budescu, DV (2007). Collective induction without cooperation? Learning and knowledge transfer in cooperative groups and competitive auctions. Journal of Personality & Social Psychology, 92 (5), 854–870.

Moulin, H (1982). Game theory for social sciences. New York: New York University Press.

Nagel, R (1995). Unraveling in guessing games: an experimental study. The American Economic Review, 85 (5), 1313–1326.

Nagel, R (1999). A survey on experimental beauty contest games: bounded rationality and learning. In DV Budescu (Ed.): Games and human behavior, 105–142. Mahwah NJ: Erlbaum.

Nagel, R (2004). Games. In D Friedman, A Cassar, R Selten (Eds.): Economics lab, 115–132. London, New York: Routledge.

Penczynski, SP (2010). Persuasion and strategic thinking. Job Market Paper.

Pesendorfer, E & Koeszegi, S (2006). Hot versus cool behavioural styles in electronic negotiations: the impact of communication mode. Group Decision and Negotiation, 15 (2), 141–155.

Preacher, KJ & Hayes, AF (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behavior Research Methods, 36 (4), 717–731.

Preacher, KJ & Hayes, AF (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40 (3), 879–891.

Rockenbach, B, Sadrieh, A & Mathauschek, B (2007). Teams take the better risks. Journal of Economic Behavior & Organization, 63 (3), 412–422.

Schoop, M, Köhne, F & Ostertag, K (2010). Communication quality in business negotiations. Group Decision and Negotiation, 19 (2), 193–209.

Slonim, RL (2005). Competing against experienced and inexperienced players. Experimental Economics, 8 (1), 55–75.

Stahl, DO & Wilson, PW (1994). Experimental evidence on players' models of other players. Journal of Economic Behavior & Organization, 25 (3), 309–327.

Stahl, DO & Wilson, PW (1995). On Players' Models of Other Players: Theory and Experimental Evidence. Games and Economic Behavior, 10 (1), 218–254.

Sutter, M (2005). Are four heads better than two? An experimental beauty-contest game with teams of different size. Economics Letters, 88 (1), 41–46.

Sutter, M (2007). Are teams prone to myopic loss aversion? An experimental study on individual versus team investment behavior. Economics Letters, 97 (2), 128–132.

Sutter, M, Kocher, MG & Strauß, S (2009). Individuals and teams in auctions. Oxford Economic Paper, 61 (2), 380–394.

Sutter, M & Strassmair, C (2009). Communication, cooperation and collusion in team tournaments—An experimental study. Games and Economic Behavior, 66 (1), 506–525.