Are four heads better than two?

An experimental beauty-contest game with teams of different size

Matthias Sutter[#]

Max Planck Institute for Research into Economic Systems, Jena

Abstract

We examine the influence of team size on decision making in a beauty-contest experiment. Teams with four members outperform teams with two members and single persons significantly, whereas the latter two types of decision makers do not differ.

JEL-classification: C72, C91

Keywords: team decision making, team size, beauty-contest experiment

This version: 16th March 2004

[#] Max Planck Institute for Research into Economic Systems. Strategic Interaction Group. Kahlaische Strasse 10. D-07745 Jena. Germany. Phone: +49 3641 686 642. e-mail: sutter@mpiew-jena.mpg.de

I. Introduction

In recent years, economists' interest in team decision making has been increasing steadily, predominantly because many important economic decisions are rather taken by teams (such as central bank boards, legislatures, political committees, or families) than by individuals. In this paper, we are going to explore the influence of team size on economic decision making. Even though the question 'Are two heads better than one?' has been addressed repeatedly and has been given an affirmative answer (Blinder and Morgan, 2004; Cooper and Kagel, 2003), the natural extension whether four heads are better than two has not been addressed so far. Given that the size of decision-making teams within commercial or political organizations varies considerably and can often easily be changed, it seems of high relevance how teams should be organized with respect to number of members and whether more heads are better in economic decision making than less heads.

Social psychology has long been interested in the question whether individual and team decisions differ significantly, yielding rather inconclusive results, also with respect to the influence of team size (Davies, 1992; Levine and Moreland, 1998). Experimental economics has not explored the effects of different team sizes on decisions so far, since the main interest has been put on general differences in the decisions of individuals and teams (of typically three subjects), but it has been able to establish some stylized facts (for a more extensive overview see Kocher and Sutter, 2005). For instance, teams have been found to behave more in line with payoff maximization in strategic bargaining games like the ultimatum game (where teams demand more than individuals in the role of proposer and accept lower offers in the role of responder; Bornstein and Yaniv, 1998) or the trust game (where teams return significantly smaller amounts than individuals do; Cox, 2002). In signaling games, teams consistently play more strategically than individuals (Cooper and Kagel, 2003). In auctions, teams seem to bid more competitive than individuals, which leads to more overbidding and worse results of teams than of individuals (Cox and Hayne, 2002; Kocher et al., 2004). Concerning the performance of teams in (non-interactive) investment decisions, teams accumulate a given level of expected payoffs at a significantly lower risk than individuals (Rockenbach et al., 2001) or have higher expected payoffs at a constant level of risk (Sutter, 2004).

In this paper, we will use an experimental beauty-contest game to examine the influence of the size of a team on its decisions. In section 2 we will introduce the game and our experimental design. Section 3 is devoted to the experimental results. Section 4 concludes.

II. The beauty-contest game

In our version of the beauty-contest game – which was likened by Keynes (1936) to professional investment activity – there are three decision makers: one individual, one team with two members, and one team with four members. All decision makers choose simultaneously a real number from the interval $I \equiv [0,100]$. The winner of the game is the decision maker whose number is closest to two thirds of the average number chosen. Hence, the target number is $x^* = 2/3 \cdot \bar{x}_t$, where $\bar{x}_t$ denotes the mean of all choices in round *t*.

The process of iterated elimination of dominated strategies leads to the game's unique equilibrium at which all decision makers choose zero. The game-theoretic structure makes the beauty-contest game an ideal tool to study how many iterations of eliminating dominated strategies a decision maker actually applies. Textbook rationality would require the decision maker to iterate infinitely. However, previous studies (see Nagel, 1995; Duffy and Nagel, 1997; Ho et al., 1998; Bosch-Domenech et al., 2002; Güth et al., 2002; Kovac et al., 2004) found that decision makers iterate only a few steps, i.e. the depth of reasoning is rather limited and there is only gradual convergence to the game-theoretic equilibrium.

The beauty-contest game has several attractive characteristics for studying differences in the quality of decision making between individuals and teams. In contrast to most of the other games that have been applied to analyse differences between team and individual decision making, it does not confound effects of rationality and learning with the effects of social preferences, like inequality aversion, fairness or reciprocity that always play a role in simple bargaining games, but which are not the focus of our study. Furthermore, loss or risk aversion cannot occur in the beauty-contest game.[1] Despite these advantages the beauty-contest game is still interactive (which is not the case in the experiment of Blinder and Morgan, 2004, for instance), has a clear economic interpretation and is very simple to explain.

Kocher and Sutter (2005) have been the first to study team decisions in the beauty-contest game. They have found that teams of three subjects each win the beauty-contest significantly more often than individuals, thereby outperforming individuals in terms of payoffs. The advantage of teams stems in particular from their ability to better anticipate the dynamics of the game in the course of repetition. However, Kocher and Sutter (2005) have not varied the size of a team to see whether larger teams are better than smaller teams and whether any kind of team is better (i.e. more successful in winning the game) than an individual. We have shaped our experimental design in order to address these questions.

Before proceeding to the design and the results, we would like to form some expectation about the possible behavior of teams of different size. According to information load theory (Chalos and Pickard, 1985), teams have higher decision consistency and are able to process high information load better than individuals. In the context of the beauty-contest game, information load consists of previous rounds data (distribution of numbers, mean, winning number, development of winning numbers over rounds etc.) and expectations about current round behaviour of other players. Given the need to process this information, information load

[1] Compared to the experiments of Bone et al. (1999, 2004) or Rockenbach et al. (2001), the beauty-contest game is also not concerned with meeting or deviating from the axioms of expected utility theory.

theory would predict that four heads are better than two, and that two heads are better than one.

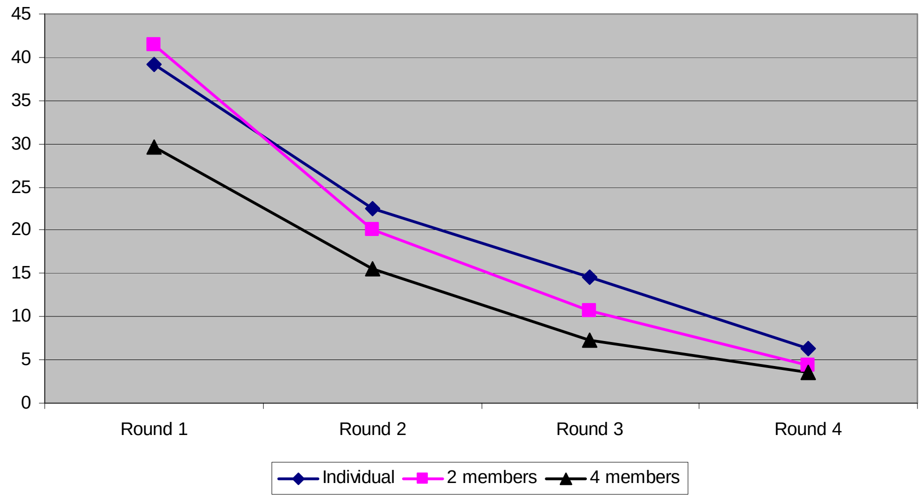
In order to test this prediction, we let three decision makers – one individual, a team with 2 members, and a team with 4 members – interact for four rounds in an experimental beauty-contest game as described above (see the Appendix for the instructions). We invited 210 subjects from the University of Jena to participate, which yielded 30 independent observations. The experiment was run at the Max Planck Institute for Research into Economic Systems Jena in December 2003. Teams (as well as individuals) were seated in sound-proof cabins so that they could discuss their decision without being heard by other decision makers. Decisions were entered on a computer in each cabin, using the software z-Tree (Fischbacher, 1999).

The winner of each round received 5€ (about 6 US$ at that time). Note that if a team won, *all* team members received 5€ in order to keep the marginal incentives for each participant constant. After each round, subjects were informed about the other choices, the mean number, the target number (i.e. two thirds of the mean) and whether they had won or not. Chosen numbers of other decision makers could not be linked to the type of decision maker (either individual or team with 2, respectively 4, members). The experiment lasted about 40 minutes, with average payoffs of 11.7€ (including a show-up fee of 4€).

III. Experimental results

Figure 1 shows the average chosen numbers by type of decision maker. In each of the four rounds, teams with 4 members choose on average the lowest number, whereas from round 2 on individuals choose the highest numbers. A non-parametric Friedman test shows that the numbers chosen by the different decision makers (individual vs. team with 2 members vs. team with 4 members) are significantly different in each single round ($p < 0.05$; $N = 30$ in each round). In particular, the numbers from teams with 4 members are smaller than those from individuals or teams with 2 members from round 1 to round 3 ($p < 0.05$ in pairwise tests in each round). In round 4, numbers from individuals are significantly higher than those from teams of either 2 or 4 members ($p < 0.05$).

Figure 1. Average guesses



Of course, choosing the lowest number in a beauty-contest game with three decision makers need not necessarily lead to winning the game, because the minimum number might be too low, resulting in the median number to win. Only the maximum number can never be the winning number. Table 1 reports in the top row the total number of wins in the whole experiment. Teams with 4 members win in 58 cases, individuals only in 31 cases and teams with 2 members in 32 cases ($\chi^2 = 5.33$, $p = 0.07$, $df = 2$).[2] The number of wins in each single round is displayed in Figure 2, showing that teams with 4 members win most often in every round.

[2] The total sum of 121 instead of 120 wins is due to a single draw between a team with 2 members and a team with 4 members, in which case each team member received 2.5€.

Table 1. Overall frequency of winning and chosen numbers

Decision maker	Individuals	Teams with 2 members	Teams with 4 members
Frequency of winning	31	32	58
Minimum number chosen (won with minimum)	35 (27)	28 (19)	58 (44)
Median number chosen (won with median)	25 (4)	44 (13)	47 (14)
Maximum number chosen	60	48	15
Number (round 2) / mean number (round 1)	0.60	0.53	0.41
Number (round 3) / mean number (round 2)	0.94	0.51	0.34
Number (round 4) / mean number (round 3)	0.46	0.33	0.28

Figure 2. Absolute win frequency

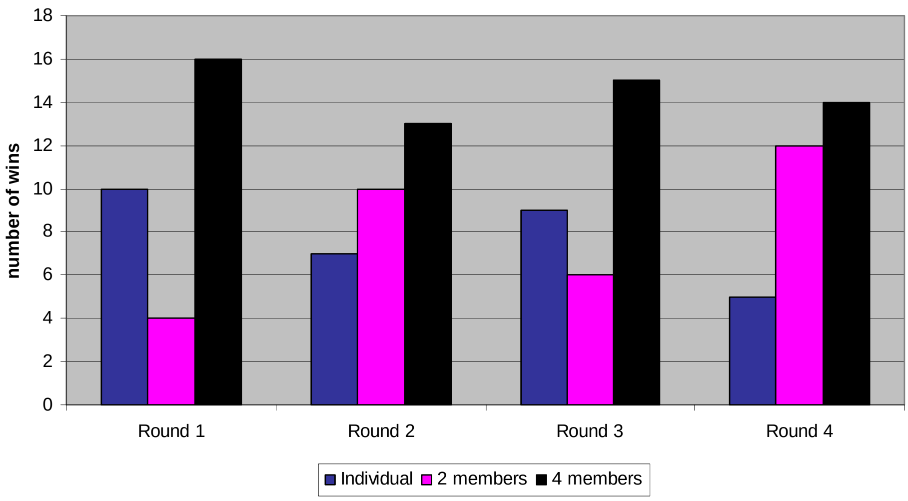


Table 1 also shows the overall frequency of each type of decision maker choosing the minimum, median or maximum number in the set of three decision makers. Teams with 4 members choose the minimum number with 58 times most often, of which it was 44 times the winning number. The other 14 wins are due to choosing the median number. Teams with 2 members win relatively often with the median number (namely in 13 out of 32 winning cases). Individuals choose most often the maximum number, but win – if at all – almost always with the minimum number.

In the bottom part of Table 1 we show the average relation of a decision maker's number in round t to the mean number in round t–1. If this relation were two thirds, that would mean that a decision maker chooses in round t the target number of round t–1. Teams with 4 members have the lowest relation in any round, falling from 0.41 in round 2 to 0.28 in round 4. That means that teams with 4 members have the fastest convergence towards the game-theoretic equilibrium (which the relation being weakly significantly smaller than the one of individuals, $p < 0.1$ in any round, based on a Friedman test for general differences across types of decision makers).

IV. Conclusion

Team decision making has caught considerable interest in economic research in recent years. So far, economics has concentrated on the differences between individual and team decision making, summarized in the question whether two heads are better than one. However, economics has devoted no attention to the possible effects of team size on team decision making.

We have shown that four heads are better than two heads or one head in an experimental beauty-contest game. Thus, the affirmative answer to the title question of this paper indicates that team size has, indeed, an effect on team decision making, at least in a context where information processing is an important aspect of the decision.

We have also found that two heads are not better than one head. However, this finding should not be misinterpreted as a contradiction to the previous findings that two heads are better than one (Blinder and Morgan, 2004; Cooper and Kagel, 2003). In particular, it is no contradiction to the earlier findings of Kocher and Sutter (2005) who had found that teams of three subjects outperform individuals in a beauty-contest game. In their experiment, they only had teams of size three, meaning that they were not able to judge the performance of teams of different size. In our experiment, we have confirmed the basic message that teams are better than individuals. However, only the large teams with four members have been more successful than individuals, but not the small two-person teams. Whether this pattern holds also in other economically interesting situations, such as when teams of different size bargain with each other, is an open question for future research.

References

Blinder, A. S., Morgan, J., 2004. Are two heads better than one? An experimental analysis of group vs. individual decision making. *Journal of Money, Credit and Banking*, forthcoming.

Bone, J., Hey, J., Suckling, J., 1999. Are groups more (or less) consistent than individuals?. *Journal of Risk and Uncertainty* 8, 63-81.

Bone, J., Hey, J., Suckling, J., 2004. A simple risk sharing experiment. *Journal of Risk and Uncertainty* 28, 23-38.

Bornstein, G., Yaniv, I., 1998. Individual and group behaviour in the ultimatum game: Are groups more 'rational' players?. *Experimental Economics* 1, 101-108.

Bosch-Domènech, A., García-Montalvo, J., Nagel, R., Satorra, A., 2002. One, two, (three), infinity, …: Newspaper and lab beauty-contest experiments. *American Economic Review* 92, 1687-1701.

Chalos, P., Pickard, S., 1985. Information choice and cue use: An experiment in group information processing. *Journal of Applied Psychology* 70, 634-641.

Cooper, D. J., Kagel J. H., 2003. Are two heads better than one? Team versus individual play in signalling games. Working Paper, Case Western Reserve University, Ohio State University.

Cox, J. C., 2002. Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. in: Zwick, R., Rapoport, A. (eds.), *Advances in Experimental Business Research*, pp. 331-50, Dordrecht: Kluwer Academic Publishers.

Cox, J. C., Hayne, S. C., 2002. Barking up the right tree: Are small groups rational agents?. Working paper, University of Arizona.

Davis, J., 1992. Some compelling intuitions about group consensus decisions, theoretical and empirical research, and interpersonal aggregation phenomena: Selected examples, 1950-1990. *Organizational Behaviour and Human Decision Processes* 52, 3-38.

Duffy, J., Nagel, R., 1997. On the robustness of behaviour in experimental 'beauty contest' games. *Economic Journal* 107, 1684-1700.

Fischbacher, U., 1999. Z-tree: Zurich toolbox for readymade economic experiments. Working Paper No. 21, Institute for Empirical Research in Economics, University of Zurich.

Güth, W., Kocher, M., Sutter, M., 2002. Experimental 'beauty contests' with homogeneous and heterogeneous players and with interior and boundary equilibria. *Economics Letters* 74, 219-228.

Ho, T.-H., Camerer, C. F., Weigelt, K., 1998. Iterated dominance and iterated best response in experimental "p-beauty contests". *American Economic Review* 88, 947-969.

Keynes, J. M., 1936. *The General Theory of Interest, Employment and Money*, London: Macmillan.

Kocher, M. G., Sutter, M., 2005. The decision maker matters: Individual versus group behavior in experimental beauty-contest games. *Economic Journal*, forthcoming January 2005.

Kocher, M. G., Strauß, S., Sutter, M., 2004. The 3G/UMTS auctions – Differences between bidding behavior of individuals and groups. Working Paper, University of Innsbruck.

Kovac, E., Ortmann, A., Vojtek, M., Brazdik, F., 2004. Guessing games with homogeneous and heterogeneous players: An experimental reconsideration. Working Paper. CERGE, Charles University Prague.

Levine, J. M., Moreland, R. L., 1998. Small groups. in: Gilbert, D. T. et al. (eds.), *The Handbook of Social Psychology*. 4th edition, vol. 2, pp. 415-469.

Nagel, R., 1995. Unravelling in guessing games: An experimental study. *American Economic Review* 85, 1313-1326.

Rockenbach, B., Sadrieh, A., Mathauschek, B., 2001. Teams take the better risks. Mimeo. University of Erfurt, 2001.

Sutter, M., 2004. Are teams prone to myopic loss aversion? An experiment study on individual versus team investment behavior. Discussion Paper, Max Planck Institute for Research into Economic Systems Jena.

Appendix (NOT for publication – Instructions will be made available on a webpage)

In this experiment, three decision makers will interact with each other in a way described below. The three decision makers are seated in separate cabins. In each of the three cabins, there will be a different number of subjects, either a single person, two persons, or four persons. The persons inside a cabin will have to come up with a single decision. I.e. if there are two or four persons in a cabin, they can discuss their decision and have to agree on a single decision.

Each decision maker i has to decide on a number x_i in the interval [0,100]. Zero and one hundred are also possible choices. The number need not be integer. The winning number is the one who is closest to a target number $\bar{x}$, the latter being defined as:

$$\bar{x} = \frac{2}{3} \cdot \frac{\sum_{i=1}^{3} x_i}{3}$$

I.e., winner is whose number is closest to two thirds of the average of all chosen numbers. The winner receives a prize of 5€. Note that if a team of two or four persons wins, ***each*** person will get 5€. If there are two or three decision makers equally close to $\bar{x}$, then the prize is split equally among these decision makers.

After all cabins have entered their decision on the computer screen, you will be informed about the three numbers (in random order; that means that you don't know which number was chosen by which decision maker), the average number, the target number and whether your cabin has won or not.

There will be **four rounds**, so that each cabin has to make four separate decisions on x_i. After each round you will be informed about the other number, the target and the winning number and about your accumulated payoffs.

In each round, you should come up with a decision within 8 minutes. In case you have not entered your decision after 8 minutes, there will be a red sign on the screen which asks you type in a decision soon.

Please stay in your cabin until the experimenter enters your cabin at the end of the experiment, when you will be paid your earnings from the experiment (plus a show-up fee of 4€).