Teaching Nash Equilibrium and Dominance: A Classroom Experiment on the Beauty...
Virtudes Alba-Fernández; Pablo Brañas-Garza; Francisca Jiménez-Jiménez; Javie...
Journal of Economic Education; Summer 2006; 37, 3; ABI/INFORM Global
pg. 305

Teaching Nash Equilibrium and Dominance: A Classroom Experiment on the Beauty Contest

Virtudes Alba-Fernández, Pablo Brañas-Garza, Francisca Jiménez-Jiménez, and Javier Rodero-Cosano

Abstract: The authors' aim in this article was to show how the use of classroom experiments may be a good pedagogical tool to teach the Nash equilibrium (NE) concept. The basic game is a version of the beauty contest game (BCG), a simple guessing game in which repetition lets students react to other players' choices and converge iteratively to the equilibrium solution. The authors perform this experiment with undergraduate students with no previous training in game theory. After four rounds, they observe a clear decreasing tendency in the average submitted number in all groups. Thus, the findings show that by playing a repeated BCG, students quickly learn how to reach the NE solution.

Key words: beauty contest game, classroom experiments, Nash equilibrium, teaching
JEL codes: A22, C99, D83

Generally, the teaching of game theory begins with the study of strategic interactions among players considering only pure strategies. Once players understand that each player's payoffs depend not only on their own actions but also on those of other players, the next step is to explain the process of elimination of dominated strategies. To do this, only two basic assumptions are required: rationality (maximization) and common knowledge (of rationality). Both concepts allow one to solve (dominance solvable) games and also predict some particular behavior. However, some games are not dominance solvable even if they have a Nash equilibrium (NE). In this case, the NE requires an additional hypothesis: common

Virtudes Alba-Fernández *is a professor of statistics, University of Jaén and researcher of centrA, Spain.* ***Pablo Brañas-Garza*** *is a professor of economics, University of Grenada, Spain, and researcher of centrA, Spain.* ***Francisca Jiménez-Jiménez*** *is a teaching assistant of economics, University of Jaén and researcher of centrA, Spain, and* ***Javier Rodero-Cosano*** *is a professor of economics, University of Mélaga, Spain (e-mail: jrodero@uma.es).* *The authors thank Rosemarie Nagel, Curro Martínez Mora, Quique Fatás, and the participants at the Economic Science Association Meeting in Tucson (2003), especially David Reiley. They thank three anonymous referees for valuable comments. They thank Samuel Hernández for his contribution to the data gathering and processing and Martha Gaustad for the language revision. Pablo Brañas acknowledges the hospitality of IESA-CSIC during this research. They gratefully acknowledge the financial support provided by the University of Jaén R&D program (# 20210/148).* ***Copyright © 2006 Heldref Publications***

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

knowledge of players' beliefs about their rivals' actions. Hence, the best response becomes the suitable mechanism for reaching the equilibrium solution. *Best response* is the best strategy that players may choose given the strategies chosen by their rivals. The NE is reached when all agents play (some of) their best response. Thus, the NE is self-enforcing because no player has incentives to deviate from it. Clearly, the use of the NE eliminates circular reasoning (player 1 thinks that player 2 thinks that player 1 thinks...) in games with at least one pure NE, not in a game with a unique mixed NE.

We propose a pedagogical tool for simplifying the teaching of the NE: successive repetitions of a dominance solvable game, the beauty contest game (BCG) introduced by Moulin (1986). This game is very useful for showing, in an intuitive manner, the two procedures described above. The game is played four times in the form of a classroom experiment, prior to teaching a game theory class.

Our findings clearly illustrate that students inexperienced in game theory apply both iterated elimination of dominated strategies and best-reply behavior in each round to the previous choices. On average, we observed a recursive approximation to the Nash prediction, which we call the *learning effect*. Therefore, by repeating this static game, we ensure that students learn how to solve it in some way.

THEORETICAL BACKGROUND

Since the introduction of the concept by Nash (1951), the NE has not only become the standard tool for the economic scientific community but also the basis for the systematic teaching of the discipline.

In game theory, it is common practice to assume that individuals magically choose a set of actions such that all the (infinitively recursive) predictions come true. To reach this hard-to-believe solution, theorists can follow one of the two original Nash interpretations: either agents are perfectly rational or an evolutionary equilibrium exists.[1] The first interpretation relies on the assumption that all agents are able to compute the game equilibria and reach one of them. In contrast, the second interpretation presupposes the existence of a large population of ordinary people that play the game in an evolutionary framework: that is, they pick up one strategy randomly, if the outcome is good, they repeat it again, or if the strategy is bad, it will disappear. After some time, the game will converge to one of the NE. The BCG is a good example of the underlying convergence to both interpretations. From a rational point of view, it is not credible that the subjects will solve this problem (few people can at the first try!), but if it is repeated, the solution will converge to equilibrium most of the time.

The beauty contest game is a simple guessing game that makes it easy to evaluate individuals' level of reasoning. The basic BCG is as follows: A certain number of subjects are invited to play a game in which all of them must simultaneously choose a real number from an interval (generally between 0 and 100). The winner is the player who chooses the number that is closest to p times the mean of all the numbers chosen, where $1 > p > 0$. The winner receives a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

prize, the losers get nothing. Under these rules, the unique NE is 0 for all players.[2] Other parameter configurations may produce different or even multiple equilibria (Bosch-Domènech et al. 2002).

As Stahl (1996) pointed out, the distribution of chosen numbers lets us analyze the depth of reasoning of the agents. In this case, level 1 includes people who expect that the other players behave randomly so they choose p * mean (being mean = 50 if the choice distribution is uniform); level 2 contains people expecting that the other's depth of reasoning is level 1, and, choosing then p^2 * mean, ...; generalizing, at level K are people who choose p^K * mean because they believe that the other people are at level $K - 1$. If K is big, p^K * mean $\simeq$ 0, so if the process is repeated ad infinitum ($K = \infty$), the theoretical solution 0, is reached, that is, the highest level of reasoning (Figure 1 represents this process with $p = 2/3$). Random answers are called level 0 of reasoning. A level of reasoning higher than 3 is rare in BCG experiments (Bosch Domènech et al. 2002), although this iterated best replay behavior is quite commonly found (Ho, Camerer, and Weigelt 1998).

In Figure 2, taken from Ho, Camerer, and Weigelt (1998, 951), we show the convergence to the 0 theoretical solution from a dominance iterated point of view.[3] Given $p = 2/3$, any number chosen between 66.6 and 100 is dominated by 66.6 = 100 * 2/3. Hence, the interval [66.6–100] corresponds to irrational behavior, which we call $R(0)$. Rational individuals will always choose a number in the [0–66.6] interval. Applying the same reasoning, $R(1)$ players will choose a number below 66.6 but above 44.4 = 2/3 * 66.6, whereas $R(2)$ players will choose a number in the interval [29.6–44.4]. Following this iterated reasoning level process ad infinitum, one reaches the unique Nash equilibrium (0, with $R(\infty)$). Thus, this game is called dominance solvable. Camerer (2003, ch. 5) described some other games that also tend to unravel and would be fun to teach in class. These include the patent race game, in which iteration eliminates some strategies but not others, centipede games, price competition, or traveler's dilemma, in which the dominance reduces

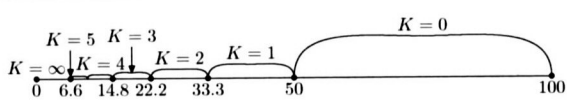


FIGURE 1. An example of different reasoning levels.

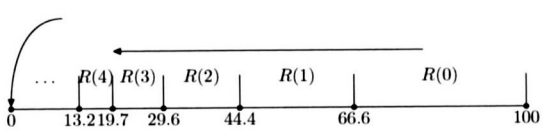


FIGURE 2. Iterated reasoning of individuals by eliminating dominated strategies.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the collective payoffs to players, and the dirty faces game, in which an increase in the number of steps of iterated deletion increases payoffs.

Nevertheless, this game makes a much more important point in practice: If a subject plays the NE (0), in the majority of the cases, he or she will not win. Note that the key to this game is to take one more step of deletion of dominated strategies than the other players (but not too many).[4] Hence, the NE is useful for knowing where the adaptive process leads, but it is not a useful normative theory as a basis for advice.[5]

The original idea behind the BCG was first put forward by Keynes (1936, 155) to show that clever investors have to "anticipate the basis of conventional valuation a few months hence, rather than... over a long term of years," if they are to act in the stock market before other investors do.

The unique equilibrium of the formal game model is 0 for $p < 1$ and is obtained by iterated elimination of weakly dominated strategies.[6] After this basic framework was laid down, some experimental researchers began to investigate BCG or *p-beauty* (Ho, Camerer, and Weigelt 1998). The first experimental study can be found in Nagel (1994, 1995).[7] Other studies have been done by Bosch-Domènech and Nagel (1997a, 1997b); Bosch-Domènech et al. (2002) and Duffy and Nagel (1997). Nagel (1998) provided a survey of the literature.

A wide variety of BCG experimental designs can be found.[8] In these, sometimes participants are students, at other times they are professors or newspaper readers. Some designs are one-shot, others are repeated, with communication versus no communication or laboratory versus field experiments.

Generally, BCG is run with individual subjects. Kocher and Sutter (2001), however, compared individual versus group behavior in this type of game. They found that although groups did not apply deeper levels of reasoning, they did, in fact, learn faster.

Repetition permits individuals to learn dynamically from other people's expected behavior. Our experiment was very similar to one in Ho, Camerer, and Weigelt (1998). In their design, individuals were given information about previous period choices and, therefore, the learning process was based on an evolutive game. In contrast, Weber (2003b) argued that learning with the BCG can happen even without feedback. He found that there was convergence toward the NE in the game, even when subjects did not receive any information between periods.

Over the past years, research on the BCG has returned to Keynes' (1936) original idea. Hirota and Sunder (2002) experimentally explored whether price bubbles in security markets are generated by a beauty-contest mechanism: If dividends are paid beyond investors' personal investment horizons, investors must create beliefs about others' beliefs (second-order beliefs) that depend on third-order beliefs, which in turn depend on fourth-order beliefs, and so on. They concluded that when the realization of the dividend is distant and well beyond investors' investment horizon, investors found it difficult to induce the fundamental value of securities from the future to the present (backward induction). This difficulty gives rise to price bubbles because, in this case, investors adjust their expectations on the basis of observed prices (forward induction).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Up to now, the BCG has primarily been used to study the depth of individuals' level of reasoning. However, we propose a new application of this framework: classroom experiments with a pedagogical aim.[9]

THE EXPERIMENT

The experiment involves using a repeated version of the beauty contest game with several groups of students to teach the NE; specifically, the iterated elimination of dominated strategies. Given that the NE is a general topic in both intermediate microeconomics and industrial organization (and, obviously, in game theory), we expect this pedagogical tool to be of great help to educators.

Three fundamental issues must be taken into account when running the experiment.

1. The group should be large enough to reduce the effect of any one person's guess on the average. On the other hand, a very large group may be too cumbersome for the instructor.[10] Nagel (1995) used 12 and 17 subjects in each group; Ho, Camerer, and Weigelt (1998) reduced the group size to 7 subjects and 3 subjects (like Kocher and Sutter 2001). Following these guides, we used small groups (5–6 subjects) and large groups (10–11 subjects).[11] Each group played its own game in which each subject played a repeated BCG against the other members of the group.
2. To motivate students, some type of reward should be given. In our classroom experiment, the winner of each round was awarded 0.25 extra-credit points applicable toward the final exam. In the case of several winners in the same round within the same group, the prize was split among them.
3. Another important issue is the number of rounds that experimental subjects must play in the BCG. Because our aim was for students to learn the NE through the iterated elimination of dominated strategies, several rounds were needed to observe this learning effect. Kocher and Sutter (2001) proved that four rounds are enough to approximate the theoretical prediction.

We performed our experiment in an intermediate microeconomics theory course during the spring semester of the 2002–2003 academic year with three different classroom groups: business majors (morning and afternoon groups, B–1 and B–2) and business + law majors (one single group, B + L). All groups had an identical individual answer handout, instructor, and grading system (ranging from 0 to 10). This subject is a requisite in both majors.

The experiment lasts about one hour and the following post-experimental session half an hour (it is worth running both sessions consecutively, approximately 1.5 hours).

Procedures

1. The instructor defined the size of each group. However, it is not important if one group differs slightly in size from the other groups. Nevertheless, for purposes of comparison, it is desirable to have similarly sized groups.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. After determining group size, the instructor randomly selected one monitor for each group from the student pool. Each monitor was awarded 0.25 points. The monitor's task was to help the instructor to record and monitor the experiment in his or her group. Each monitor was given a calculator and a monitoring sheet correlatively numbered (see appendix). The monitor kept track of the individual choices in the group and calculated the mean and the winning number. The monitoring sheets were crucial to explaining the results after the experimental session as these findings would easily illustrate basic concepts of game theory, namely the iterated deletion of dominated strategies.
3. The third task consisted of creating groups, that is, students were randomly assigned to different groups.
4. To make the monitor's tasks easier, the instructor sat each group in a single row (or column).
5. The instructor then gave the monitors their instructions and explained the procedures to them. When this was done, the monitors handed out the instruction sheets and the individual answer sheets to the experimental subjects (see appendix).
6. The instructor explained the instructions out loud. It is important to avoid numerical examples as the generation of focal points was immediate in this kind of game. Any questions were answered aloud. Experimental subjects were told that speaking was absolutely forbidden.[12] When everyone fully understood the rules, the experiment began.
7. Round 1: Experimental subjects filled in their answer sheet. The monitor then collected all guessed numbers and calculated the mean and 2/3 * *mean*. By comparing this value to those reported by the experimental subjects, the monitor determined the winner. The group was told the mean, 2/3 * *mean*, and the winner's guess (but not who the winner is). Nagel (1999) also gave the students the complete set of guesses.[13]
8. Round 2: Following round 1, the students were informed that they would play another round, independent of the previous one. The procedure for round 2 was identical to round 1. Groups remained the same and did not communicate.
9. Rounds 3 and 4 were the same as round 2.

POSTEXPERIMENTAL DISCUSSION

Immediately following the experimental session, the postexperimental discussion was begun. In this second part of the experiment, we introduced the theoretical background of the NE. This step involved three topics:

1. We present the iterated elimination of dominated strategies underlying the BCG by means of Figure 2. By explaining this process, students can understand that if rationality is common knowledge, nobody will choose a number within the interval [66.6–100] because this subset of numbers is dominated by [0–66.6]. Following this reasoning *ad infinitum*—students fully capture this idea.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

At this moment, it is a good idea to show students the first round results of the groups involved in the session. Alternatively, a simple histogram of the pooled data may be used to cover up any embarrassing answers (for more details, see the Results section). Students may compare their own chosen numbers with those reported in Figure 2. Moreover, it could be fun and interesting to show other experimental results to offer students the opportunity to compare themselves with others (for example, high school students, 80-year-olds, corporate CEOs, or even game theorists). Nagel (1999) and Camerer, Ho, and Chong (2003b) reported a high number of BCG experiments with several contexts and subject pools.[14]

2. Once common knowledge rationality was clearly defined, it was time to explain *best response* behavior. Students observed that their own performance depended on their beliefs about other players' actions. Furthermore, they understood that other players' behaviors also depended on their beliefs about others and so on.

   As an example, we presented some preliminary results of their own responses. To do this, some monitoring sheets (see appendix Table A1) were used to illustrate the average behavior of the groups. Because the average number chosen decreased through successive rounds, students realized that game theory approximately predicted their performance.

3. We now had all the necessary ingredients to "cook up" the NE concept. Let $\overline{x}$ be player i's forecast of the mean of all numbers submitted. Player i chose x_i with the objective of $min_{x_i \in [0,100]} |x_i - p\overline{x}|$. All players behaved identically. Common knowledge of beliefs, together with mutual best response, implied that in any equilibrium, $x = px$. Hence, the unique equilibrium must be that $x = 0$, the number that is equal to p times itself.

   Because students had experienced the recursive dynamics of the NE, we explained what the abstract concept implied and how this could be reconciled with the real world. In the theoretical model, we supposed that everything was instantaneously adjusted. The BCG had to be explained as if it were a slow motion picture: Repetition lets individuals reach the theoretical solution step by step. Perfectly rational agents do not need this, but people (even economists) are not fully rational. This is the second remarkable point: Repetition could be a good substitute for rationality. In fact, the modern view in game theory is that equilibrium arises from adaptation, evolution, communication, or imitation. Because these processes take time, equilibration should not occur instantly in one-shot games.

   Nagel (1998), Stahl (1996), and, most recently, Camerer, Ho, and Chong (2003b) presented very precise formal models of limited rationality that accurately described the experimental results of these games. Specifically, Camerer, Ho, and Chong (2003a) developed a nonequilibrium model for one-shot games with decision rules for players doing different steps of thinking.[15] Their prediction for dominance solvable games will converge to the NE as a limit result.

Finally, as an anecdote, we mentioned *A Beautiful Mind*, the movie about John Nash's life. The film tries to explain the NE to ordinary people using an allegorical

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sequence. In this scene, John is in a bar with some friends when a group of girls walks in. If all the guys try to pick up only the prettiest girl, all the girls would leave, and all the guys would be out of luck.[16] The BCG can be seen as a similar symbolic game: Individuals have to compete to outguess the rest of the participants, but without cooperation, costly errors can occur. Knowing game theory concepts could reduce these costs.[17]

RESULTS

The experimental BCG was played in three different sessions, using samples B–1, B–2, and B + L. Table 1 shows the distribution of subjects participating in the experimental sessions by group. Our entire sample consisted of 139 subjects assigned to 6 groups per session (5 in B + L). All groups played four rounds of the BCG.

Although our interest was specifically in the process of learning the NE, we considered it meaningful to show first-period behavior. In this way, we were able to relate our experimental results to the theoretical reasoning levels shown in Figure 2. The relative distribution of first-round iterated steps is shown in Figures 3–5. As seen, the behavior of subjects in sessions B–1 and session B + L were quite similar, whereas the subjects in session B–2 seemed to have a higher level of reasoning than the other two.[18]

Table 2 shows the aggregate level of the average chosen number and the standard deviation across sessions and rounds. The mean of the chosen numbers decreased throughout successive rounds in all cases.[19] However, we observed that the path to convergence was not identical. In Figure 6, we illustrate each trend in the experimental session.

At first sight, sessions B–1 and B + L behaved similarly, but the trend in B–2 showed a flatter slope. As our main interest was to study the learning effect, we analyzed the average variation more accurately round after round. We defined *speed of convergence* as: $(\mu_i - \mu_{i-1})/\mu_{i-1}$ where μ_i is the average guessed number in round i ($i = 1, \ldots, 4$) (shown in Table 3).

TABLE 1. Distribution of Group Size, Number of Subjects per Group and Session

	Session		
Group	B − 1	B − 2	B + L
1	10	6	10
2	11	6	8
3	9	6	9
4	9	5	8
5	10	6	10
6	10	6	—
Session pop.	59	35	45

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

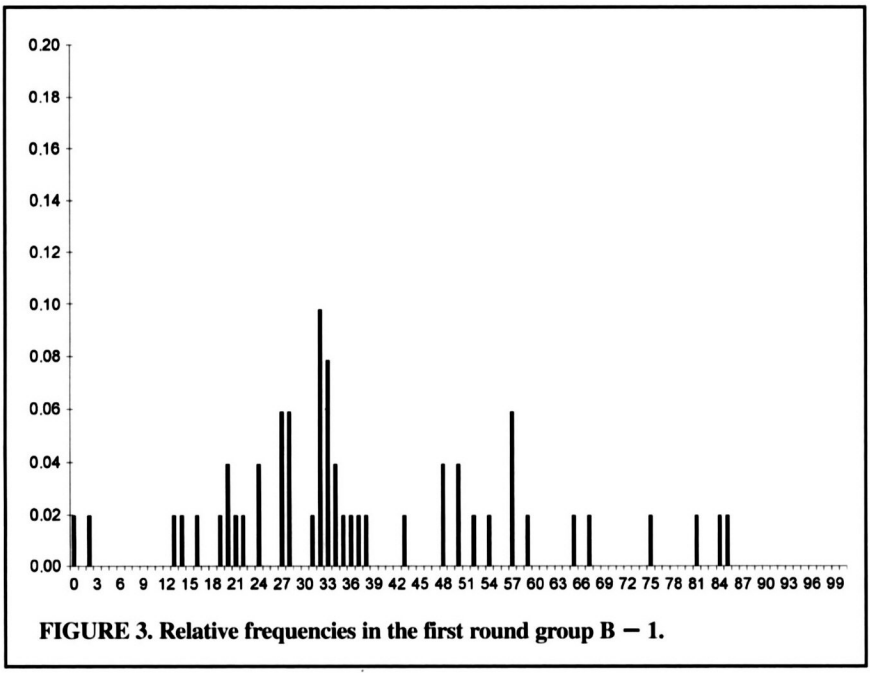


FIGURE 3. Relative frequencies in the first round group B − 1.

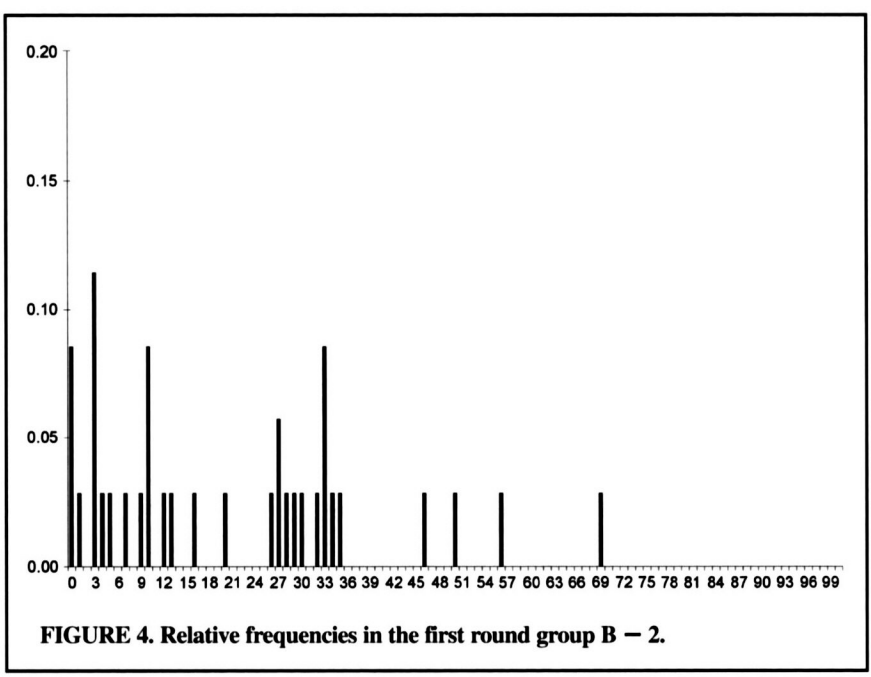


FIGURE 4. Relative frequencies in the first round group B − 2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

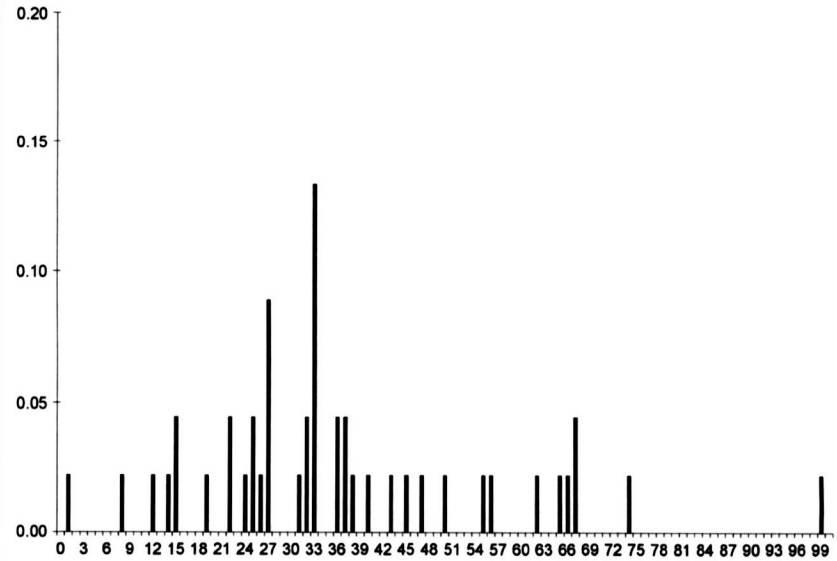


FIGURE 5. Relative frequencies in the first round group B + L.

TABLE 2. Average Number and Standard Deviation per Round and Session

	Session					
	B − 1		B − 2		B + L	
Round	Average	S.D.	Average	S.D.	Average	S.D.
1	38.885	19.348	20.782	17.611	37.067	18.891
2	32.087	14.235	17.060	12.362	29.008	11.554
3	22.660	16.750	13.042	7.318	22.712	11.577
4	14.419	13.360	9.748	7.503	19.142	13.944

Note that all the sessions shared a similar initial speed of convergence close to 20 percent. However, this speed varied dramatically within and between groups from round 2 to 3 and from round 3 to 4. Thus, the learning effect was clearly observable in all rounds and sessions.

After this descriptive approximation, we statistically analyzed the results obtained. To be more specific, we were interested in testing if students were induced to modify their behavior toward the NE by repeatedly playing a BCG a set number of times. Therefore, we expected the average number in a round to be different from the following round and, furthermore, that the difference between one round and the following round would be greater than 0, that is, that it would decrease. To verify this, we formulated the following null hypothesis:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

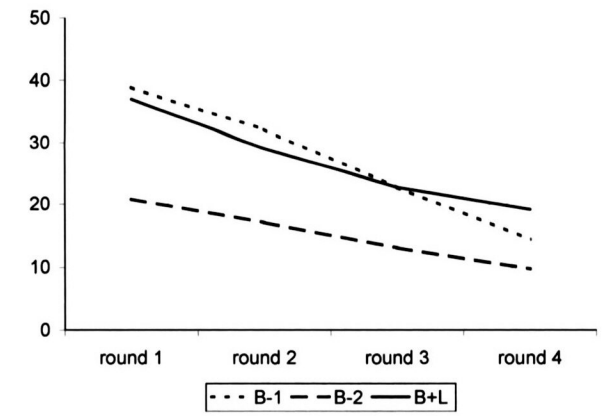


FIGURE 6. Evolution of session averages.

TABLE 3. Speed per round and session

	Session		
Speed of Convergence for round	B–1	B–2	B + L
$(\mu_2 - \mu_1)/\mu_1$	0.17	0.18	0.22
$(\mu_3 - \mu_2)/\mu_2$	0.29	0.24	0.22
$(\mu_4 - \mu_3)/\mu_3$	0.36	0.25	0.16

$H_0 : \mu_{ij} - \mu_{kj} = 0, i < k$ and $i, k = 1, \ldots, 4$

and alternative hypothesis:

$H_1 : \mu_{ij} - \mu_{kj} > 0, i < k$ and $i, k = 1, \ldots, 4,$

where i, k represent two consecutive rounds, and j represents the session number (1 for B–1, 2 for B–2, and 3 for B + L).

The statistical analysis is developed in depth in the appendix. Briefly, we checked if each population followed the normal distribution. If one of them did not follow the normal distribution, we used a parametric and a nonparametric test to contrast our null hypothesis. Our main results showed that, in sessions B–1 and B + L, the learning effect was observed in the entire experiment, and, in session B–2, the learning effect was observed in the last two rounds.

To sum up, these results indicate that this is a powerful tool to aid students in reaching the NE.

VARIATIONS OF THE EXPERIMENT

Multiple possible variations of the BCG could be used as a classroom experiment. We describe some of the most important of these here.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Two-Person BCG

This version of the game is interesting because, unlike the general BCG in which there are more than two players, in a two-person game, the player who picks the lowest number wins (if $p < 1$). Thus, the subject who plays the NE strategy (0) always wins. Therefore, 0 is not only a weakly dominant strategy but also the best response for all the choices of the other person. For this reason, this variant is a good example of the attractive power of the equilibrium point. As Camerer, Ho, and Chong (2003a) pointed out, this special game can be solved by only one step of weak dominance. In their model, all players using one or more thinking steps will choose 0. However, it is important to recall that the most important point of the experiment is to show students the strength of the equilibrium concept, even for complicated games.

However, Grosskopf and Nagel (2001) showed that usually subjects do not choose the 0 in the one-shot game and then they do not reach the equilibrium immediately. Even under full information, students need some repetitions to realize that 0 is the NE. They concluded that "convergence toward equilibrium is driven by imitation and adaptation rather than self-initiated rational reasoning" (2001, 1).

Nonetheless, an interesting option could be to perform both the $n = 2$ and $n > 2$ experiments in the same class. Some groups could play the first version, while the others play the second one. It would then be interesting to compare both versions in the post-experimental session and show students that the NE sometimes fails to describe the reality if the convergence process is too long (Grosskopf and Nagel 2001).

The Entire-Class Experiment

If the class is not large enough, an interesting alternative is to perform the experiment using the whole class as a single group. The main advantage to this approach is that no student needs to sit out as a monitor. Also, the individual effect on the population average would decrease slightly. In contrast, if the group is too large, the process can become a bit cumbersome for the teacher. The most straightforward way to run this BCG is for the instructor (or any ad hoc volunteer) to write the submitted numbers on the board while another volunteer calculates the mean and p * *mean*. Note that all guessed numbers should be collected beforehand.

Web-based Experiments

In his Web page at the University of Virginia,[20] Charles Holt provides some excellent internet software for playing the BCG (among many other games). Not only can students play the BCG on this Web site, but they can also check their own results and performance online as well as compare their outcomes with other previous experiments. However, running the experiment in this way requires having a computer room with Internet access and banning its use in the classroom. Except for this minor problem, Veconlab is an amazing tool for teaching. Another point to take into consideration, in our case, is the fact that Veconlab is available only in English at this time.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Other Variations

There are several additional variations on the game. A good idea would be to run some of these simultaneously with different groups in the same class to compare the different outcomes arising from different rules.

- When $p > 1$, Nagel (1998) showed that this variation of the game was not trivial. She considered two cases: (*a*) under $p > 1$ and the interval [0–100], there are two equilibria: everybody chooses either 0 or 100. The latter is the perfect equilibrium.[21] (*b*) For $p > 1$ and the interval [100–200], the only equilibrium is 200; the upper bound of the interval. The iterated deletion process starts from the lower bound $100p$, $100p^2$, and so on.[22]
- Using the median instead of the mean, for example, can help to illustrate to students that the game does not change and the convergence toward the equilibrium is basically the same under different conditions (Nagel et al. 1999).
- By using large and small groups, identical results may be found in the one-shot game, but faster convergence to the equilibrium occurs in larger groups. Duffy and Nagel (1997) studied the influence of a single player on aggregated performance. They use three treatments changing the order statistic from mean to maximum or median for $p = 1/2$, so the extreme choices had very different weights.
- Giving students different feedback and when comparing feedback-free BCG with normal BCG, we observed that convergence was slower when subjects were not informed about the numbers chosen by their partners. The lack of information constrained subjects' learning to their own experience (Weber 2003a, 2003b). In our case, we gave students very little information (only aggregated data). Given that one of the goals of the experiment was to show students the attractive power of the equilibrium, we preferred a low information variant.[23]

FINAL REMARKS

We have formulated an interesting classroom experiment: a repeated version of the beauty contest game. Our experience is that this sort of game is a good way to introduce the complexity of the equilibrium concept. Not only does it permit students to see the mechanism in action but also to comprehend the difference between a theoretical abstraction and real-world dynamics toward equilibrium. The BCG is one of those tricky games worth playing. Although it is plainly simple in its formulation and the solution is always obvious *ex post*, students realize the difficulties of outsmarting other people while having a good time doing it.

The classroom experiment needs some careful preparation to make conveniently mixed groups. However, once everything is ready to go and the monitors are well trained, rounds should run quite smoothly and increasingly fast. Although for our purposes, 4 rounds are enough, it would be easy to play as many as 10 rounds, if necessary. Continuous repetitions would be very rewarding for students, as this would let most of them reach the theoretical solution on their own before the postexperimental session.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Once the BCG is done and the postexperimental session is concluded, it is possible to run other experiments so students realize how game theory training helps them to solve interesting (and sometimes lucrative) puzzles.

NOTES

1. Because of space constraints Nash used only the rationality explanation in the published version.
2. This is not as obvious as it looks, Grosskopf and Nagel (2001) tried a one-shot BCG with game theorists. The answers were very different from 0 but better than the usual ones!
3. See Rapoport and Amaldoss (2000) for an experimental study of iterated elimination of (strongly) dominated strategies.
4. Obviously, this process of reasoning involves knowing who you are playing with. In this case, an individual who knows the 0 solution could also guess that most people will not achieve the 0 solution. Thus, clever individuals will be able to link their answer to their estimation of the average rationality level. We call this rationality level ∞ plus (Keynes' "clever investor"). Alba-Fernández et al. (2004) reported some examples of ∞ plus behavior arising from internet BCG experiments. Bosch-Domènech et al. (2002, 1693) found strong evidence of this behavior, reporting some comments from students who followed this strategy. Grosskopf and Nagel (2001) argued that most individuals think that other people are not fully rational, thus explaining why the equilibrium is not reached immediately.
5. This is one of the main advantages of using the BCG for teaching purposes. Nagel (1999), using the same graph shown in Figure 2, illustrated how to explain iterated elimination of dominated strategies in a classroom setting.
6. It can also be obtained by using a best-response argument (Nagel 1999).
7. Nagel's (1994, 1995) main purpose was to contrast an iterated best-reply dominance model.
8. See Camerer, Ho, and Chong (2003b) for a list of 24 different BCG experiments.
9. See Nagel's 1999 study in which she described this application.
10. Ho, Camerer, and Weigelt (1998) showed that smaller groups (three subjects in their research) need more time to converge.
11. It is difficult to have identically sized groups in a class. Usually, the instructor does not have the exact number of students required.
12. Subjects were required to keep a maximum level of confidentiality for their own sake. If any subjects were to know their rival's guess, they used the best reply rule.
13. To simplify the process, we did not give our students the entire set of numbers. Nonetheless, the type of information provided in our experiment affected the convergence speed: Weber (2003a, 2003b) analyzed the effect of different feedback conditions on learning. He found that convergence toward the equilibrium occured under all circumstances, although it was faster with feedback. See also the Variations section to follow.
14. In general, these studies reflect a substantial regularity across very different groups, except some highly trained ones.
15. The link between thinking steps and the iterated deletion of dominated strategies lies in the fact that players who are not thinking strategically randomize, whereas players doing k steps of thinking accurately predict what lower-level players do and best respond accordingly.
16. This "equilibrium" (all the men including Nash, ignore the pretty blonde girl), in fact, is not an equilibrium at all because any man in the group who is willing to score her could do it given the strategies of the others, as Anderson and Engers (2002, 2) point out.
17. It is important to remind students about the importance of their intellectual achievement: John Nash won the Nobel Prize thanks to this simple discovery. See Varian (2002) for a nice explanation about Nash's brilliant and beautiful contribution. In contrast to what viewers of the film might think, it is clearly not a new strategy to pick up girls. Myerson (1999) gave an interesting historical perspective to this important innovation.
18. Camerer, Ho, and Chong (2003a) estimated the average number of thinking steps at 1.61 across 24 one-shot beauty contest games. They found quite low estimates ($0 - 1$) when $p > 1$ and high estimates ($3 - 5$) in games in which the equilibrium was within the interval.
19. Although the initial average observed in our experiment was similar to that of Kocher and Sutter (2001), our final average was very different! After running four rounds as well, their final averages were around 7, whereas ours were clearly higher (from 9.7 to 19.1).
20. Found at http://www.people.virginia.edu/~cah2k/.
21. Nevertheless, there are no dominated strategies in this version and, thus, no process of iterated elimination of dominated strategies leading to equilibrium.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

22. In this case, the number of eliminations is finite. That is the reason why these games are called finite threshold games. See also Ho, Camerer, and Weigelt (1998).
23. As Weber (2003b) pointed out, even if you do not give any feedback, convergence will be achieved, although the pace will be slower. It will take more time to make the calculation, although the rounds will be faster. It is up to the instructor to choose the desired mix of information and convergence speed.

REFERENCES

Alba-Fernández, V., P. Brañas-Garza, F. Jiménez-Jiménez, and J. Rodero-Cosano. 2004. Communication, coordination and competition in the beauty contest game: Eleven classroom experiments. *Central Europe Journal of Operations Research* 12 (4): 369–87.

Anderson, S. P., and M. Engers. 2002. A beautyful blonde: A Nash coordination game. Working Paper 359. University of Virginia, Charlottesville.

Bosch-Domènech, A., and R. Nagel. 1997a. El juego de adivinar el número *x*: Una explicación y la proclamación del vencedor. *Expansión* June 16:42–43.

———. 1997b. Guess the number: Comparing the FT's and expansión results. *Financial Times* June 30:14.

Bosch-Domènech, A., J. García Moltalvo, R. Nagel, and A. Satorra. 2002. One, two, (three), infinity, ...: Newspaper and lab beauty-contest experiments. *American Economic Review* 92 (5): 1687–1701.

Camerer, C. F. 2003. *Behavioral game theory: Experiments in strategic interaction.* Princeton, NJ: Princeton University Press.

Camerer, C. F., T. H. Ho, and J.K. Chong. 2003a. A congnitive hierarchy theory of one-shot games and experimental analysis. Working paper, Caltech, http://www.hss.caltech.edu/camerer.

———. 2003b. Models of thinking, learning, and teaching in games. *American Economic Review* 93 (2): 192–95.

Duffy, J., and R. Nagel. 1997. On the robustness of behavior in experimental beauty contest games. *Economic Journal* 107: 1684–1700.

Grosskopf, B., and R. Nagel. 2001. Rational reasoning or adaptative behavior? Evidence from two-persons beauty contest games. Harvard NOM Research Paper 01-09. Harvard Business School, Cambridge, MA.

Hirota, S., and S. Sunder. 2002. Stock market as a "beauty contest" investor beliefs and price bubbles sans dividend anchors. Working Paper 02-42, Yale International Center for Finance (ICF), New Haven, CT.

Ho, T.H., C. F. Camerer, and K. Weigelt. 1998. Iterated dominance and iterated best response in experimental "p-beauty contests." *American Economic Review* 88 (4): 947–69.

Keynes, J. M. 1936. *The general theory of interest, employment and money.* London: MacMillan.

Kocher, M. G., and M. Sutter. 2001. The "decision maker" matters: Individual versus group behavior in experimental "beauty-contest" games. Working paper, University of Innsbruck, Austria.

Moulin, H. 1986. *Game theory for social sciences.* New York: New York Press.

Myerson, R. B. 1999. Nash equilibrium and the history of economic theory. Unpublished Manuscript.

Nagel, R. 1994. *Reasoning and learning in guessing games and ultimatum games with incomplete information: An experimental study.* PhD diss., University of Bonn, Germany.

———. 1995. Unraveling in guessing games: An experimental study. *American Economic Review* 85 (5): 1313–326.

———. 1998. A survey on experimental "beauty-contest games": Bounded rationality and learning. In *Games and human behavior, essays in honor of Amnon Rapoport*, eds., D. Budescu, I. Erev, R. Zwicke. 105–42. New Jersey: Lawrence Erlbaum.

———. 1999. A Keynesian beauty contest in the classroom. *Classroom Expernomics* 8 (Fall). http://www.marietta.edu/ delemeeg/expernom/f99.html#nagel.

Nagel, R., A. Bosch-Domènech, A. Satorra, and J. Garcia-Montalvo. 1999. One, two, (three), infinity: Newspaper and lab beauty-contest experiments. Working Paper 438, UPF, Barcelona, Spain.

Nash, J. F. 1951. Non-cooperative games. *Annals of Mathematics* 54 (2): 286–95.

Rapoport, A., and W. Amaldoss. 2000. Mixed strategies and iterative elimination of strongly dominated strategies: An experimental investigation of states of knowledge. *Journal of Economic Behavior & Organization* 42 (4): 483–521.

Stahl, D. O. 1996. Boundedly rational rule learning in a guessing game. *Games and Economic Behavior* 16 (2): 303–30.

Varian, H. R. 2002. What, exactly, was on John Nash's beautiful mind? Economic Scene. *New York Times*, April 11.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Weber, R. A. 2003a. Learning and transfer of learning with no feedback: An experimental test across games. Working paper, Carnegie Mellon, University, Pittsburgh, PA.
———. 2003b. "Learning" with no feedback in a competitive guessing game. *Games and Economic Behavior* 44 (1): 134–44.

APPENDIX

Experimental Instruments

Instructions

You are going to participate in a microeconomic experiment. You have been randomly assigned to a 6–10 person group. In this game you will have to make decisions repeatedly in four rounds. These people will be your partners throughout the four stages. Moreover, your group has been assigned a monitor that will oversee the procedure.

The rules of the game in each period are as follows: you should choose a (integer or decimal) number in the interval [0–100]. You are allowed to choose 0 and 100. Once the monitor has collected your group's choices, the winner will be determined. The winning number is the number closest to 2/3 of the average of all the numbers chosen by your group:

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

The winner of each round will receive a prize of 0.25 extra points valid toward the Micro II final exam. If two or more people are equally close to x, the prize will be split equally among them. A person so lucky as to guess the right number all four times will get a whole point!

TABLE A1. Monitoring Sheet

Monitor:_____ Group:_____ Session:_____

Subject	Full name	Round 1	Round 2	Round 3	Round 4
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
	Average				
	2/3 of average				

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE A2. Individual Answer Sheet

Code: _____ Group: _____

First round

Your choice: [ ]

Group average: _____
2/3 of the average _____
Winning number: _____

Second round

Your choice: [ ]

Group average: _____
2/3 of the average _____
Winning number: _____

Third round

Your choice: [ ]

Group average: _____
2/3 of the average _____
Winning number: _____

Fourth round

Your choice: [ ]

Group average: _____
2/3 of the average _____
Winning number: _____

TABLE A3. Kolmogorov-Smirnof Goodness-of-Fit Results

	Session					
	B–1		B–2		B + L	
Round	z	p	z	p	z	p
1	1.116	0.165	0.937	0.344	1.135	0.152
2	1.063	0.208	0.713	0.690	0.681	0.742
3	1.579	0.014*	0.384	0.998	0.792	0.556
4	1.285	0.074	0.709	0.696	0.867	0.440

* Significant at 0:05 Type I error level.

First you must write down the group and code you were given at the top of the attached sheet. Then in each round, you must select your number. When everyone has finished, the monitor will collect your responses. Afterwards, you will be given back your sheet with some additional information: the average of your group, 2/3 of this average, and the winning number. This procedure will be repeated four times.

If you have any questions, please raise your hand and the instructor will come to you. You are not allowed to speak during the experiment.

Statistical Analysis

Let X_{ij} be the number chosen by a student at round i in session j, where $i = 1, \ldots, 4$ and $j = 1, 2, 3$. Let $F_{ij}(X)$ be the distribution function associated with each variable X_{ij}.

First, verify if each $F_{ij}(X)$ follows a normal distribution with mean μ_{ij} and standard deviation σ_{ij}. To do so, use the Kolmogorov-Smirnof goodness-of-fit test. The results are shown in Table A3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE A4. Parametric Test for Mean Equality

Round Comparison	Session B–1		Session B–2		Session B + L	
	t	*p*	*t*	*p*	*t*	*p*
1 vs. 2	2.315	0.012	1.252	0.109	2.951	0.002
2 vs. 3	—	—	2.107	0.021	3.148	0.001
3 vs. 4	—	—	2.859	0.003	1.758	0.042
2 vs. 4	7.586	0	—	—	—	—

TABLE A5. Wilcoxon Nonparametric Test for Mean Equality

Round Comparison	Session B–1		Session B–2		Session B + L	
	z	*p*	*z*	*p*	*z*	*p*
1 vs. 2	−2.378	0.008	−1.077	0.1414	−2.467	0.006
2 vs. 3	−4.612	0	−2.001	0.022	−2.871	0.002
3 vs. 4	−5.225	0	−3.459	0	−3.45	0

As seen, individual choices adjust to a normal distribution except in the third round of the first session (B–1) of the game.

Using a parametric test for the equality of means between paired rounds, the results are given in Table A4.

Observe that it is not possible to apply this test to the third round of the first session (B–1) because it does not follow a normal distribution.

From the *p* values shown in Table A4 one can say that the difference between successive average choices is statistically positive in all cells except for the first two rounds in the second session (B–2). In this last case, the null hypothesis of mean equality is accepted.

To include the third round of the first session in the statistical analysis, apply the Wilcoxon nonparametric test using the same hypothesis as above. Results are summarized in Table A5.

Given these *p* values, only the means of the first two rounds in the second session are equal, and the null hypothesis is rejected for the remainder of cases.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.