Guessing Games Experiments in School Education and Their Analysis

Oleksii Ignatenko[0000-0001-8692-2062]

Institute of Software Systems of NAS of Ukraine,
40/5 Academic Hlushkov Ave., Kyiv, 03187, Ukraine
o.ignatenko@gmail.com

Abstract. The paper deals with experimental game theory and data analysis. The research question is how different pools of players understand strategic situations, they never faced before. We examine data from a number of strategic interactions (games) and present players progress in finding solutions in competition with other players. We propose four "pick a number" games, all with similar-looking rules but very different properties. These games were introduced (in the body of scientific popular lectures) to very different groups. In this paper we present data gathered during lectures and develop tool for exploratory analysis using R language. Finally, we discuss the findings and open questions.

Keywords: behavioral game theory, guessing game, k-beauty contest, active learning, R.

1 Introduction

A key element of strategic thinking is to include into consideration what other agents do. Agent here is a person, who can make decisions and his/her actions have influence on the outcome. Naturally, person cannot predict with 100% what will others do, so it is important to include into model beliefs about other person thinking and update them during the game. Also, if we can't know what other player think, we can understand what is his/her best course of action. This is the main research topic of game theory.

All this makes decision making very interesting problem to investigate. In this work we will apply game theory to analyze such problems. Game theory provides mathematical base for understanding strategic interaction of rational players. There is important note about rationality, we should make. As Robert J. Aumann formulate in his famous paper [1], game theory operates with "homo rational", ideal decision maker, who is able to define his/her utility as a function and capable of computing best strategy to maximize it. This is the main setup of game theory and one of major lines of criticism. In reality, of course, people are not purely rational in game theory sense. They often do not want to concentrate on a given situation to search for best decision or simply do not have enough time or capabilities for this. Sometimes they just copycat behavior of others or use some cultural codes to make strange decisions. Also (as we see from the experiments) it seems that sometimes homo sapiens make decisions with reasons, one can (with some liberty in formulation) label as “try and see what happens”, “make random move and save thinking energy” and even “make stupid move to spoil game for others”.

This is rich area of research, where theoretical constructions of game theory seems to fail to work and experimental data shows unusual patterns. However, these pattens are persistent and usually do not depend on age, education, country and other things. During last 25 years behavioral game theory in numerous studies examines bounded rationality (best close concept to rationality of game theory) and heuristics people use to reason in strategic situations. For example we can note surveys of Vincent P. Crawford, Miguel A. Costa-Gomes and Nagore Iriberri [2] and Felix Mauersberger and Rosemarie Nagel [3]. Also there is comprehensive description of the field of behavioral game theory by Colin F. Camerrer [4].

The guessing games are notable part of research because of their simplicity for players and easy analysis of rules from game theoretic prospective. In this paper we present results of games played during 2018-2019 years in series of scientific popular lectures. The audience of these lectures was quite heterogeneous, but we can distinguish three main groups:

- kids (strong mathematical schools, ordinary schools, alternative education schools);
- students (bachelor and master levels);
- mixed adults with almost any background.

We propose framework of four different games, each presenting one idea or concept of game theory. These games were introduced to people with no prior knowledge (at least in vast majority) about the theory. From the other hand, games have simple formulation and clear winning rules, which makes them intuitively understandable even for kids. This makes these games perfect choice to test ability of strategic thinking.

1.1 Game theory definitions

We will consider games in strategic or normal form in non-cooperative setup. A non-cooperativeness here does not imply that the players do not cooperate, but it means that any cooperation must be self-enforcing without any coordination among the players. Strict definition is as follows.

A non-cooperative game in strategic (or normal) form is a triplet $G = \{N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N}\}$, where:

- N is a finite set of players, i.e., $N = \{1,\dots,N\}$;
- S_i is the set of admissible strategies for player i;
- u_i: $S \rightarrow R$ is the utility (payoff) function for player i, with $S = S_1 \times \dots \times S_N$ (Cartesian product of the strategy sets).

A game is said to be static if the players take their actions only once, independently of each other. In some sense, a static game is a game without any notion of time, where no player has any knowledge of the decisions taken by the other players. Even though, in practice, the players may have made their strategic choices at different points in time, a game would still be considered static if no player has any information on the decisions of others. In contrast, a dynamic game is one where the players have some information about each others' choices and can act more than once. Summarizing, these are games where time has a central role in the decision-making. When dealing with dynamic games, the choices of each player are generally dependent on some available information. There is a difference between the notion of an action and a strategy. A strategy can be seen as a mapping from the information available to a player to the action set of this player.

Based on the assumption that all players are rational, the players try to maximize their payoffs when responding to other players' strategies. Generally speaking, final result is determined by non-cooperative maximization of integrated utility. In this regard, the most accepted solution concept for a non-cooperative game is that of a Nash equilibrium, introduced by John F. Nash. Loosely speaking, a Nash equilibrium is a state of a non-cooperative game where no player can improve its utility by changing its strategy, if the other players maintain their current strategies. Formally, when dealing with pure strategies, i.e., deterministic choices by the players, the Nash equilibrium is defined as follows:

A pure-strategy Nash equilibrium (NE) of a non-cooperative game $G = \{\mathrm{N}, \{S_i\}_{i\in\mathrm{N}}, \{u_i\}_{i\in\mathrm{N}}\}$ is a strategy profile $s^{'} \in S$ such that for all $i \in \mathrm{N}$ we have the following:

$$u_i\left(s_i^{'}, s_{-i}^{0}\right) \geq u_i(s_i, s_{-i}^{0}) \text{ for all } s_i \in S_i.$$

Here $s_{-i} = \left[s_j\right]_{j\in\mathrm{N}, i\neq j}$ denotes the vector of strategies of all players except *i*. In other words, a strategy profile is a pure-strategy Nash equilibrium if no player has an incentive to unilaterally deviate to another strategy, given that other players' strategies remain fixed.

1.2 Guessing games

In early 1990[th] Rosemary Nagel starts series of experiments of guessing games, summarized in [5]. She wasn't the first one to invent the games, it was used in lectures by different game theory researchers (for example Hervé Moulin [6]). But her experiments were first experimental try to investigate the hidden patterns in the guessing game. Later, Teck-Hua Ho, Colin Camerer and Keith Weigelt [7] gave the name "*p*-beauty contest" inspired by Keynes comparison of stock market instruments and newspaper beauty contests. This is interesting quote, so lets give it here:

"To change the metaphor slightly, professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees." (John Maynard Keynes (1936) "The General Theory of Employment, Interest and Money", Chapter 12.V).

The beauty contest game has become important tool to measure "depth of reasoning" of group of people using simple abstract rules. Now there are variety of rules and experiments presented in papers, so lets only mention some of them.

2 Experiments setup

The setup is closer to reality then to laboratory and this is the point of this research. All games were played under following conditions:

1. Game were played during the lecture about the game theory. Participants were asked not to comment or discuss their choice until they submit it. However, this rule wasn't enforced, so usually they have this possibility if wanted;
2. Participants were not rewarded for win. The winner was announced, but no more;
3. During some early games we use pieces of paper, later we switch to google spreadsheets. The last tool has possibility to make multiple submission (with different names), but total number of submissions allows to control that.

The aim of this setup was to free participants to explore the rules and give them flexibility to make decision in uncertain environment. We think it is closer to real life learning without immediate rewards then laboratory experiments. Naturally, this setup has strong and weak sides. Let's summarize both.

The strong sides are:

1. This setup allows to measure how people make decisions in "almost real" circumstances and understand the (possible) difference with laboratory experiments.
2. These games are part of integrated approach to active learning, when games are mixed with explanations about concepts of game theory (rationality, expected payoff, Nash equilibrium etc), and they allow participants to combine experience with theory;
3. Freedom and responsibility. The rules doesn't regulate manipulations with conditions. So, this setup allows (indirectly) to measure how players cheat with rules or spend their time to solve the task.

Weak sides are:

1. Some percentage of players make "garbage" decisions. For example, choose obviously worse choice just to spoil efforts for others;
2. Kids has (and often use) possibility to talk out decision with the neighbors;
3. Sometimes participants (especially kids) lost concentration and didn't think about the game but made random choice or just didn't make move at all;
4. Even for simplest rules, sometimes participants failed to understand the game first time. We suppose it is due to conditions of lecture with (usually) 30-40 persons around.

Probably these weaknesses are inevitable in realistic scenarios. We claim that even if experiments are not in “pure” laboratory conditions, they reveal interesting behavior and are worth to introduce in this paper.

2.1 Rules of games

All games have the same preamble: Participants are asked to guess integer number in range 1 – 100, margins included. Note, that many setups, investigated in references, use numbers starting with 0. But the difference is small.

To provide quick choice calculation we have used QR code with link to Google Forms, where participants input their number. All answers were anonymous (players indicate nicknames to announce the winners, but then all records were anonymized). The winning condition is specific for every game.

1. **p-beauty contest.** The winning number is the closest to 2/3 of average.
2. **Two equilibrium game.** The winning number is the furthest from the average.
3. **Coordination with assurance.** The winning number is the number, chosen by plurality. In case of tie lower number wins.
4. **No equlibrium game.** The winning number is the smallest unique.

All these games are well-known in game theory. Let’s briefly summarize them. First game is dominance-solvable game. Strategy “to name numbers bigger than 66” is dominated, since it is worse than any other. So rational player will not play it and everybody knows that. Then second step is to eliminate all numbers higher than 44 and so on. At the end rational players should play 1 and all win. In our setup we go further than just give players learn from observation. After first round we explain in detail what is Nash equilibrium and how it affects the strategies. After this explanation all participants actually **knew** that choosing 1 is the equilibrium option, when everyone wins. We supposed, that this should help to improve strategies in next round, but it is not.

Second game is about mixed strategies. Easy to show that if you want to choose number smaller then 50 – best way is to choose 1, since all other choices are dominated. And if you want to choose number bigger then 50 – best idea is to choose 100. Also, it is meaningful to choose 50 – it almost never wins. So, if many players will choose 1 – you should choose 100 and vice versa. In this game the best way to play is literally drop a coin and choose 1 or 100.

Third game has many equilibriums, basically every number can be winning. But to coordinate players must find some focal points (Thomas C. Schelling [8]). Natural focal point (but not only one!) is the smallest number since smaller number wins in case of tie. This slim formulation allows nevertheless make successful coordination in almost all experiments.

Finally, last game is in a dark water. As far as we know there is no equilibrium or rational strategy to play it. So sometimes very strange numbers are winners here.

3 Results and data analysis

3.1 First game

First game results are given in the Table 1.

Table Error! No sequence specified.**.** Results for p-beauty contest game.

Pool of players	Round	Mean	Winning number	Participants amount	Percent of irrational (> 66) choices
Alternative humanitarian school	1	66.69	66.69	13	61.5
	2	3.91	3.91	12	0
	3	3.07	2.05	13	0
Alternative mathematical school	1	42.82	28.54	17	11.7
	2	24.37	16.24	16	0
Adults	1	40.57	27.05	19	10.52
Alternative humanitarian school	1	52.54	35.02	11	27.2
	2	15.41	10.27	12	8.33
Adult (Facebook online)	1	22.98	15.32	102	4.9
Ordinary high school	1	43.41	28.94	51	23.5
	2	46.5	30.99	62	33.8
Mathematical school 1	1	43.41	28.94	51	23.5
Mathematical school 2	1	30.58	20.38	50	8
	2	14.26	9.5	57	5.26
Mathematical school 3	1	37.07	24.71	29	6.89
	2	26.2	17.47	29	6.89
Mathematical school 4	1	42	27.99	18	16.6
	2	23.1	15.39	20	0
Ordinary school	1	48.6	32.46	26	23
	2	19.78	13.18	23	0
Adults (conference on data science)	1	37.25	24.83	60	15
	2	21.44	14.29	57	12.28
IASA KPI (BSc)	1	42.4	28.27	27	22.2
NAUKMA (MSc)	1	27.37	18.24	8	12.5
	2	8.62	5.74	8	0

Almost all winning numbers are fall (roughly) in the experimental margins, obtained in Rosemary Nagel work [5]. With winning number no bigger than 36 and not smaller than 18 in first round. Two exceptions in our experiments were Facebook on-line test (15.32), when players can read information about the game in, for example, Wikipedia. And other is alternative humanitarian school (66.69), which seems didn't got the rules from the first time.

Using R statistical visualization tool, we can analyze in details how players from different types change their decisions between first and second round (Figure 1).

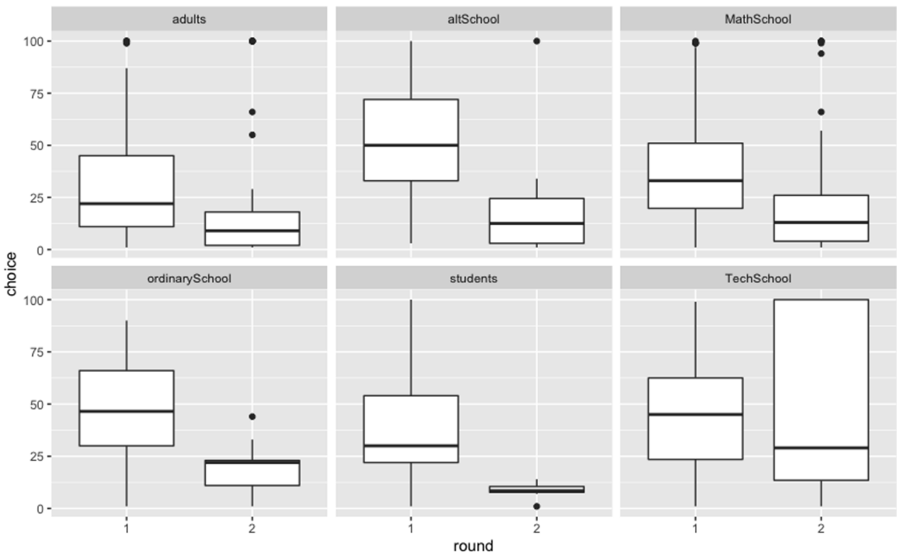

Fig. 1. Boxplot with two-round data about choice in first game.

Interesting metric is the percent of "irrational choices" – choices that can't win in (almost) any case. Let's explain, imagine that all players will choose 100. It is impossible from practice but not forbidden. In this case everybody wins, but if only one player will deviate to smaller number – he/her will win and others will lose. So, playing numbers bigger than 66 is not rational, unless you don't want to win. And here we come to important point, in all previous experiments this metric drops in second round and usually is very low (like less than 5%) [8]. But in our case, there are experiments where this metric become higher or changes very slightly. And initially values are much higher than expected. So here we should include factor of special behavior, we can call it "let's show this lector how we can cheat his test!". What is more interesting – this behavior more clear in case of adult then kids.

It is also interesting to see distribution of choices for different types of groups. We can summarize choices on the histograms (Figure 2).

Another metric [9] is how much winning choice in second round is smaller than in first. Due to concept of multi-level reasoning, every player in this game trying to its best to win but can't do all steps to winning idea. So, there are players, who just fight with nature. They think – let's assume all players chose numbers in random, then 50 will be average and 33 will be winning choice. It is first level reasoning. Then they try to estimate how many players chose 33, and calculate that best response for it will be 22 (of course we should also include those who failed to do first step and make some random choice) and so on. Based on result of first round and, in fact, explanation about the Nash equilibrium, players must know that it is better to choose much lower numbers. But graph shows that decrease is quite moderate. Only students show good performance in this matter. And tech school shows increase in winning number in second round! (Figure 3)

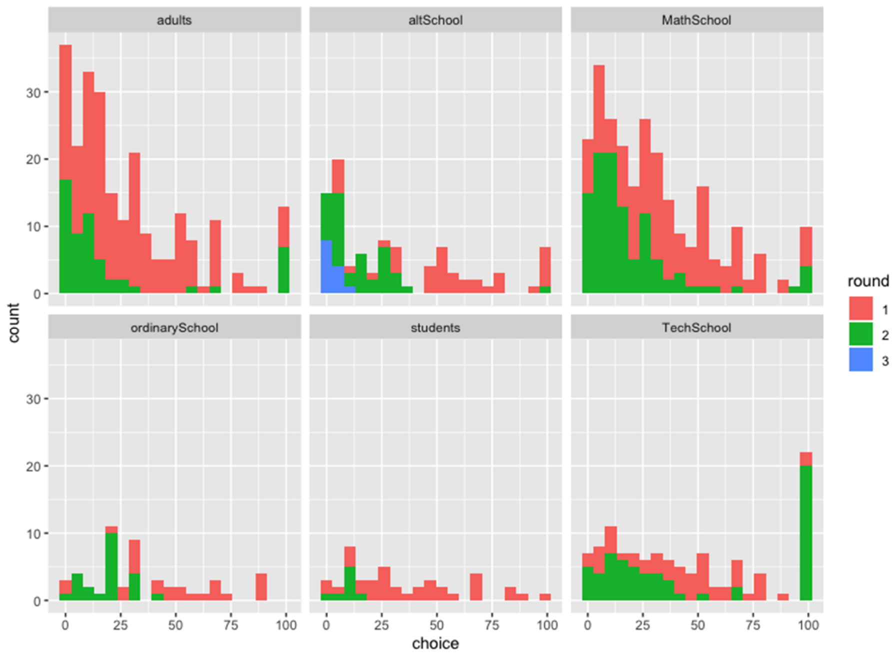

Fig. 2. Histogram of choices in first game for each type.

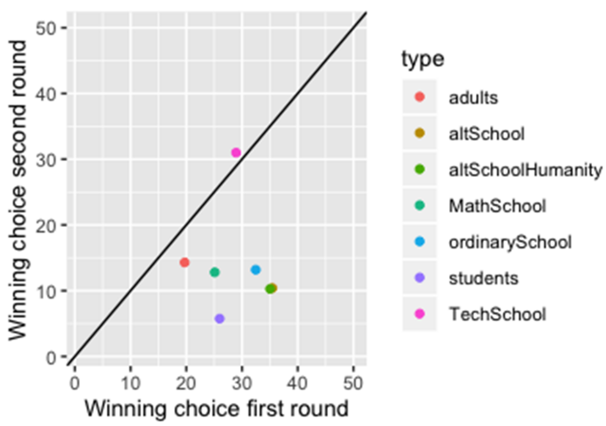

Fig. 3. Winning numbers in first and second round for each type.

3.2 Second game

In second game the key point is to understand that almost all strategies are dominated. The results are presented on Figure 4. and we can see that average can be bigger or smaller than 50, and accordingly winning choice will be 1 or 100. It is worth to note, that popular nature of these experiments and freedom to participate make the data gathering not easy. For example, many participants just didn't take any decision in second game. Results summarized in next table.

Table 2. Results for second game.

Pool of players	average	Percent choose 1	Percent choose 100	Number of participants
adults	43.64912	21.05263	26.31579	171
altSchool	43.77778	25.92593	14.81481	27
MathSchool	48.55844	22.72727	24.67532	154
ordinarySchool	51.21739	30.43478	30.43478	23

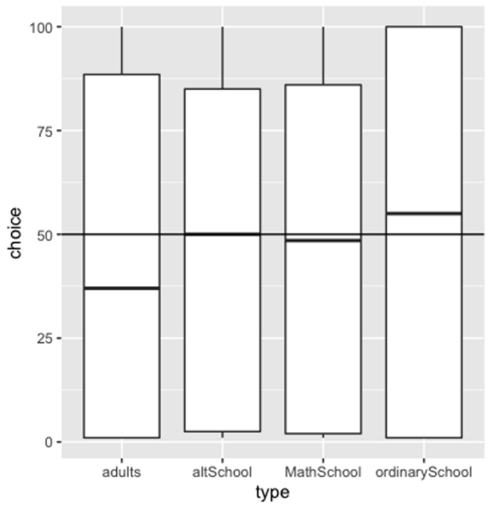

Fig. 4. Boxplot for choices in second game.

This is remarkable result, players without prior communications choose to almost perfect mixed equilibrium: almost the same percentage choose 1 and 100. This is even more striking taking into account no prior knowledge about mixed strategies and mixed equilibrium, kids play it intuitively and without any communication.

3.3 Third game

Third game is simpler than first two, it is coordination game where players should coordinate without a word. And, as predicted by Thomas Shelling, they usually do. Date presented on Figure 5 shows that 1 is natural coordination point, with one exception – Tech school (id = 1 here) decided that it would be funny to choose number 69 (it was made without singe word). Probably, it is the age (11th grade) here to blame. Also, we can note attempt to coordinate around 7, 50 and 100.

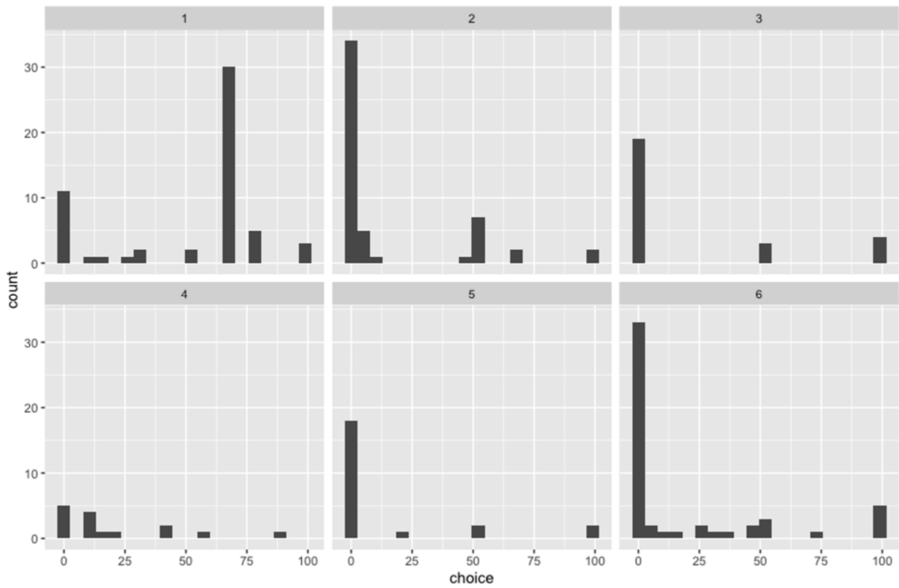

Fig. 5. Coordination game. Histograms of choices.

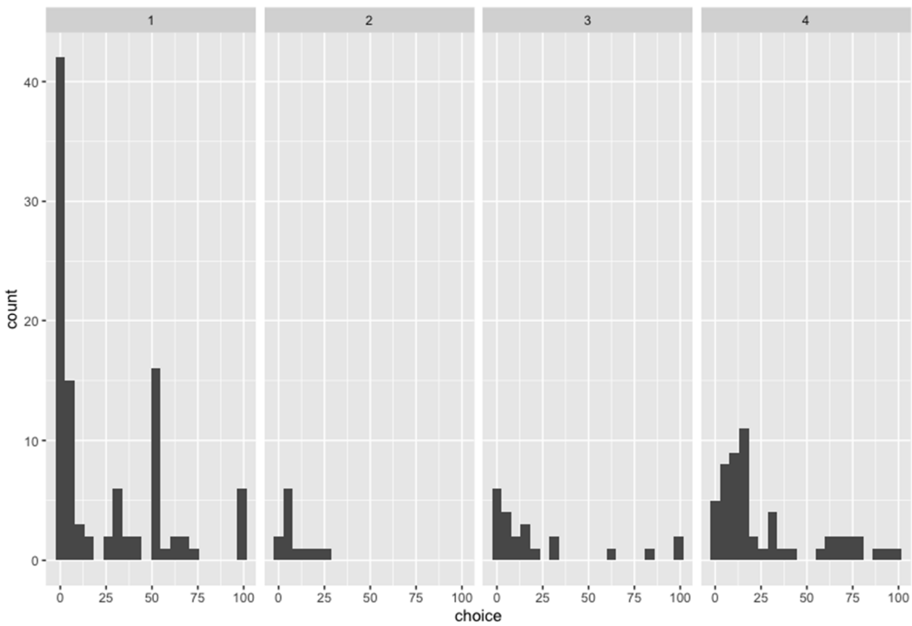

Fig. 6. Min unique game.

3.4 Fourth game

Here we just present the resulting histogram for each group and the winning numbers were: 12, 2, 4, 20 (Figure 6). Since no equilibrium here was theoretically found, we can only gather data at this stage and formulate hypothesis to found one.

All experimental data and R file for graphs can be accessed in open repository [10].

5 Conclusions

In this paper we have presented approach to make experimental game theory work for learning in educational process and be a research tool at the same time. Our result show classical pattern in decision making – actually every group behave in almost the same way dealing with unknown game. Some tried to deviate for unusual actions (like choosing 100 or choosing 69), and this is interesting point of difference with more "laboratory" setup of existing research. The main findings of the paper are following:

1. To learn the rules, you need to break them. Participants have chosen obviously not winning moves (> 66) partly because of new situation and trouble with understanding the rules. But high percent of such choices was present in second round also, when players knew exactly what is going on. This effect was especially notable in the cases of high school and adults and almost zero in case of special math schools and kids below 9th grade. We can formulate hypothesis that high school is the age of experimentation when children discover new things and do not afraid to do so.
2. If we considered winning number as decision of a group, we can see that group learning fast and steady. Even if some outliers choose 100, mean still declines with every round. It seems that there is unspoken competition between players that leads to improvement in aggregated decision even if no prize is on stake. Actually, it is plausible scenario when all participants choose higher numbers. But this didn't happen in any experiment. The closest case – Tech school, when bunch of pupils (possible coordinating) switch to 100 still only managed to keep mean on the same level.
3. In second game the surprising result is that players use mixed strategies very well. It is known (from experiments of Colin Camerer) that chimpanzee can find mixed equilibrium faster and better than humans. It seems that concept of mixed strategies is very intuitive and natural. But still in quite unfamiliar game players made almost equal number of 1 and 100, so each player unconsciously randomized his own choice.
4. In third game players coordinates to 1, as expected, because of condition that from numbers with equal choices – lesser wins. Also, we can note attempts of coordination around 7, 50 and 100. What is interesting is that in practice the condition was never applied – majority chooses 1 and that's it. If we decrease the numbers range to 1-10, other numbers have chance to win (5 or 7 for example). So, this is unexpected result – increasing of number of choices leads to bigger uncertainty when players trying to find slightest hint what to do, and this is condition of "lesser wins". When

players apply this condition to big area, they probably think – "1 is perfect choice, and other will think in that way also, this increase chances of winning".

The results have multiple applications:

- to provide kids with first-hand experience about strategic interactions and explain their decisions;
- to demonstrate how game theory experiments can be used in educational process;
- to understand difference in decision making among groups;
- to compare results with classical experiments and replicate them in modern Ukrainian education system.

References

1. Aumann, R.J. What Is Game Theory Trying to Accomplish? In: Arrow, K., Honkapohja, S. (eds.) Frontiers of Economics, pp. 5–46. Basil Blackwell, Oxford (1985)
2. Crawford, V.P., Costa-Gomes, M.A., Iriberri, N.: Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications. Journal of Economic Literature **51**(1), 5–62 (2013). doi:10.1257/jel.51.1.5
3. Mauersberger, F., Nagel, R.: Levels of Reasoning in Keynesian Beauty Contests: A Generative Framework. In: Handbook of computational economics, vol. 4, pp. 541–634. Elsevier (2018). doi:10.1016/bs.hescom.2018.05.002
4. Camerer, C.F.: Behavioral game theory: Experiments in strategic interaction. Princeton University Press, Princeton (2003).
5. Nagel, R.: Unraveling in Guessing Games: An Experimental Study. The American Economic Review **85**(5), 1313–1326 (1995)
6. Moulin, H.: Game Theory for the Social Sciences, 2nd edn. New York Univeristy Press, New York (1986)
7. Ho, T.-H., Camerer, C., Weigelt, K.: Iterated Dominance and Iterated Best Response in Experimental "*p*-Beauty Contests". The American Economic Review **88**(4), 947–969 (1998)
8. Schelling, T.C.: The Strategy of Conflict. Harvard University Press, Cambridge (1960)
9. Güth, W., Kocher, M., Sutter, M.: Experimental 'beauty contests' with homogeneous and heterogeneous players and with interior and boundary equilibria. Economics Letters **74**(2), 219–228 (2002)
10. Ignatenko, O.: ignatenko/GameTheoryExperimentData. https://github.com/ignatenko/GameTheoryExperimentData (2020). Accessed 10 June 2020