



# Using Bots in Strategizing Group Compositions to Improve Decision–Making Processes

Shai Neumann<sup>1</sup>(✉), Suraj Sood<sup>2</sup>, Markus Hollander<sup>3</sup>, Freda Wan<sup>4</sup>, Alexis-Walid Ahmed<sup>4</sup>, and Monte Hancock<sup>4</sup>

<sup>1</sup> Eastern Florida State College, Melbourne, FL, USA  
neumanns@easternflorida.edu

<sup>2</sup> Department of Psychology, University of West Georgia, Carrollton, GA, USA

<sup>3</sup> Center for Bioinformatics, Saarland University, Saarbrücken, Germany

<sup>4</sup> Sirius18, Reston, USA

**Abstract.** This paper explores utilization of bots created to play as participants in two games that have relevance in the field of economics. The first game is an iterated version of the prisoner’s dilemma that has relevance to decisions requiring trust while conducting business and involves a binary decision. The simulation examines one dimension of the Myers–Briggs personality type and its potential association with strategies that may be employed during an iterated version of the game, leading to different levels of performance. A web–based bot is used for such simulation. Results appear to support the possibility of exploiting personality type in this dimension in the context of similar circumstances.

The second game is an iterated power to take game that simulates interactions between taxing authorities (takers) and tax–payers (responders), and involves a continuous set of possible decisions for each stage of the game. A specific Myers–Briggs personality type of specified extremity is embedded into each player bot by randomly generating answers to a version of the Myers–Briggs Type Indicator. These answers are individually associated with preferences and strategies that allow the bots to react to the changing game state according to their personality type.

This paper explores the combination of, among other things, team sizes, initial conditions of income tax rates in some of the major economies on different continents, adjudication methods for group decisions, as well as personality type extremity. In a large number of simulations, various group compositions in terms of MBTI personality type are matched up against each other under these conditions, and their performance is ranked for each role.

Results appear to suggest that selective recruitment practices based on personality type can enhance performance of both takers and responders and may lead to improvements of certain aspects of economic conditions.

**Keywords:** Iterated prisoner’s dilemma · Bots  
Iterated power to take · Personality type

## 1 Introduction

Experimental economics is a branch of economics that deals with experiments in lab settings that are designed to simulate individual or group behavior in order to better understand the functioning of economic systems. In many cases, human subjects play games and their decisions during games are studied. Such settings have their limitations in terms of scope of participants, number of iterations of a game they can play, and validity outside the particular lab. Many experiments take place on university campuses and samples are drawn from a limited student population in a specific region of a specific country. The global economy involves a highly diverse population living under different cultural, political, and economic conditions.

Consideration of the increasing role bots play in social media, reaching high levels of sophistication that makes them sometimes difficult to distinguish from human behavior, raises the possibility that bots can be created to play games of experimental economics, with the advantage that settings will reflect realities in different countries, and different economic and political systems. Using psychological instruments, human diversity can be infused into bot behavioral preferences, groups sizes can vary a lot more, and number of iterations can be expanded significantly.

We explore two iterated games that may be associated with economic related decision making, namely the prisoner's dilemma and the power to take game. The first requires a binary decision and the second requires a decision in a continuous range of values in some interval. In both cases, the iterated nature suggests the possibility of bots producing insights that allow for enhancements of some human decision making processes. Additionally, using personality classification may improve predictability of particular strategies employed by individuals or groups.

As a personality model, we employed the widely used Myers–Briggs–Type–Indicator (MBTI) that divides the general population based on four dichotomous dimensions into 16 mutually exclusive subpopulations with distinct personalities and distinct expected preferences and behaviors under various circumstances [12]. The MBTI dimensions model how different personality types perceive information and how they form judgments about said information. The first dimension is “Introversion” (I) versus “Extroversion” (E) describes how different personalities relate to other people and the inner versus the outer world. The “Intuition” (N) versus “Sensing” dichotomy constitutes the second dimension that focuses on preferring abstract over concrete information. The third dimension, “Thinking” (T) versus “Feeling” (F), is about judging based on objective, logical judgments versus intra- or interpersonal values. “Perceiving” (P) versus “Judging” (J) is the last dimension that is about spontaneous adaptation versus adherence to plans.

## 2 Background

### 2.1 Prisoner’s Dilemma

The prisoner’s dilemma is a game in which two individuals are in jail and are separately interrogated for a suspected crime [14]. Each person may either “cooperate” with the other – saying nothing about any potential crime while being questioned – or “defect” against them by confirming their complicit involvement. Dawkins in [5] offered the following payoff matrix to describe the possible combinations of choices and corresponding outcomes for oneself and another suspect when the prisoner’s dilemma is “iterated”, i.e. repeated several times (Fig. 1).

		What you do	
		Cooperate	Defect
What I do	Cooperate	Fairly good <b>REWARD</b> for mutual cooperation <b>3 points</b>	Very bad <b>SUCKER’S PAYOFF</b> <b>0 points</b>
	Defect	Very good <b>TEMPTATION</b> to defect <b>5 points</b>	Fairly bad <b>PUNISHMENT</b> for mutual defection <b>1 point</b>

**Fig. 1.** Dawkins’ iterated prisoner’s dilemma payoff matrix.

The present study sought to add to the research on the iterated prisoner’s dilemma by adding personality factors as influencing variables. Specifically, two players were created with independently-determined Myers-Briggs personality types, one representing a human user and the other a bot. It was assumed that Myers-Briggs Thinking and Feeling types each have distinctly-preferred strategies while playing the prisoner’s dilemma. Of primary interest was how each type would perform relative to the other given this personality difference between them. Strategies were matched to Thinking and Feeling types based on the answers to questions such as:

1. Which personality preference is more likely to indicate one’s tendency to cooperate with others?
2. Which personality preference is more likely to indicate one’s tendency to act in their self-interest?
3. Which personality preference is more likely to indicate one’s tendency to make peace offerings following mutual conflict?
4. Which personality preference is more likely to indicate one’s tendency to be wary regarding the intentions of others?
5. Which personality preference, if any, is more likely to indicate a desire for all parties involved in a situation to benefit (i.e., to strive toward the “common good”)?

**Strategies.** “Nice” strategies were defined in Dawkins’ [5] sense as consisting of its user “never [being] the first to defect” (p. 12). Opposed to nice strategies are “nasty” strategies, the more regular use of which were further hypothesized to correlate negatively with their player’s score. Dawkins defined nasty strategies as being those with which a player would “sometimes defect, however rarely, when not provoked”.

Since users of nice strategies prefer to cooperate, it is arguable that they primarily act in both their and the other player’s interests. This does not hold true for users of nasty strategies, who are comparatively more likely to defect and thus act in their own interest at the expense of the other player’s.

**Hypothesis.** The hypothesis for this study was that a direct positive correlation would exist between the score of a player (whether automated or non-automated) and the extent to which they employ nice prisoner’s dilemma strategies. This hypothesis reflects Dawkins’ own conclusion, which was that “nice guys do well in this game” (p. 212).

## 2.2 Power to Take Game

The power to take is a game played by two opposing sides, with one side playing the role of a take authority that claims a percentage of the other sides’s income, and the responding side deciding what percentage of their own wealth to destroy based on the take rate. The game was used in a laboratory experiment by Bosman and van Winken in [1] with students as players to investigate the impact of emotions on behavior. Initially, it was played by individuals, however, Bosman et al. expanded the experiment in [2] to be played in teams in order to explore group decision making.

**Previous Findings in Power to Take Experiments.** In [1], Bosman and van Winken found that responders typically destroy nothing or everything. The latter typically happened when the responders experienced negative emotions in response to a take rate that was subjectively too high. In line with this finding, responders’ expectations regarding an acceptable take rate was a significant factor for the probability of destroying income.

In [2], Bosman et al. found that the individual decisions in a group follow the same trend observed in the game played by single players. Additionally, most takers thought a 50% take rate was fair, whereas responders tended to find a take rate of 0% as fair. However, only a small minority of the players were concerned with fairness in the group decision making process. Furthermore, both sides ignored the group decision making process of the other side and viewed the other team as a single entity. Lastly, they observed that group decisions were largely in line with a simple majority rule as well as the average of individual take inputs.

Overall, they came to the conclusion that emotional reactions, which they termed “emotional hazard”, result in decreased efficiency both in the two-player

and group power to take game. For group decision making processes it seems to be important to consider both the decision making rule as well as the impact of individual decisions. They pose the question of why there are large differences between individual players and how these differences relate to the group decision making process.

**Possible Impact of Personality Type on Group Decision Making.**

Bosman et al. suggested in [2] that expectations-based classification is an important explanatory variable for destruction rate as well as take rate. This fact may suggest that different personality types will have different expectations and will react differently to different realities. Differences in personality type might partially account for the different importance placed on fairness and other factors during the decision making process of both individuals and groups. The personality type composition of teams could have an impact on intra-group dynamics and could thus result in different group behavior.

**Relevance of the Power to Take Game for Economics.** The take authorities and take rate could be interpreted as the government imposing tax rates on the populations, whereby the destroy rate of the responders could represent the amount of effort people will make to reach a certain level of wealth given a tax rate. It is reasonable to assume that under extremely high tax rates a large segment of a highly talented workforce would prefer to reduce workload to avoid the unattractive prospect of seeing very little reward for maximum effort. When that is the case, the productivity of an economy suffers, and tax collection is reduced. If that does not happen in a certain society, then 100% tax rate would become a viable option and a taxing authority will likely be happy to impose it. However, in situations where the tax payers respond by reducing their productivity to nearly 0%, such an economy could be destroyed.

In an iterated version of the game, the players do not only respond to the immediate decisions of the opposing side, but also to the game history up to the current game round. This might be similar to tax authorities considering the effectiveness of prior tax rates when deciding to increase or decrease the tax rate. Similarly, tax payers might have expectations regarding an acceptable tax rate based on prior tax rates, as well as based on the reactions of the tax authority to signals of dissatisfaction.

**Research Questions.** Our interest in this context is to see what kind of personalities, individually, and in groups, would potentially perform either one of the two roles in the game particularly well, and which personalities would underperform in either role. A related question of interest is whether certain compositions of personality types can be linked to increased or decreased performance in the game, and, as an extension, which personalities or composition of personalities would be suitable as tax “collectors” and which would be successful as tax payers. Furthermore, we wish to investigate if there are take rates that overall result

in the best (or worst) performance regarding overall economy, tax collection or net earnings. Finally, it is of interest to examine the contribution of adjudication methods, team size, and prior, possibly cultural, history of tax rates to the group decision making process.

### 3 Methodology of the Prisoner's Dilemma

#### 3.1 Type Designation and Personality-Based Strategy Assignment

MBTI F types were assigned three nice strategies out of a total of five, meaning that they usually did not defect unless the other player did first. In contrast, T types were assigned three nasty strategies (out of five strategies, total). These decisions were made because “individuals who score the highest on scales of disagreeableness [meaning they are more likely to also score as T rather than F types [10]] appear to others as being conceited, egocentric, competitive, antagonistic, skeptical, overcritical, or distrustful toward rivals” ([11], p. 120). Those who are more likely to score as F types were thus assumed to be more cooperative and trustful toward other players.

Before the games started, the human user and bot (named “Sirius” after this project’s group) were typed independently. The human user was prespecified to always be an F, and Sirius a T. However, along each of the I/E, N/S, and J/P dimensions, players were typed randomly. After the program generated the user and Sirius’ types, they were assigned distinct strategies for the coming set of prisoner’s dilemma rounds based primarily on whether they were Thinking or Feeling types.

Some strategies (i.e., “Tit for Tat” and “Random”) were possible for both players regardless of their types, whereas others were only possible for either T or F types. For instance, “Always Defect” was only possible for the former and “Remorseful Prober” was only possible for the latter. F strategies included Tit for Tat, “Naive Peacemaker”, Remorseful Prober, “Always Cooperate”, and Random; T strategies included Tit for Tat, “Tit for Two Tats”, “Naive Prober”, Always Defect, and Random. Descriptions of each of these strategies were borrowed from existing literature and work surrounding the iterated prisoner’s dilemma [4, 7].

#### 3.2 Simulation

Following execution of the designation program, strategies were inputted into a minimally-adapted version of a Web-based prisoner’s dilemma simulator [4] (Fig. 2). Since multiple strategies were assigned uniquely to both T and F players and the comparative performance of these two types was of primary interest, 50 iterations per type and strategy designation were run.

Each simulation was followed by the appearance of a dialog box which showed the game’s final outcome.

The screenshot shows a game interface with a light green background. At the top, there are two score boxes: "Your Score: 3" and "Sirius' Score: 3". Below these is an "Outcome:" box containing the text: "Sirius also co-operated and gets 3 points. You get 3 points - a fairly good reward for mutual co-operation." Underneath the outcome box are two buttons: "Co-operate" and "Defect".

The lower section of the interface contains simulation controls. It starts with a "Run Simulation" button, followed by two dropdown menus: "Always Co-operate" and "Tit For Tat". To the right is a "Help" button. Below these is a link for "Strategy Definitions".

There are three rows of radio button options:
 

- "Random Intervention (%):" with options 1, 2, 5, 10, 20, 30 (selected), and 40.
- "Simulation Rate (secs / round):" with options 0.05, 0.2, 0.5, 1, 2, 5, 10 (selected), and 30.
- "Number of Simulation Rounds:" with options 50 (selected), 100, 200, 500, 1,000, and 10,000.

At the bottom, there are two rows of "Customised Strategy" settings, each with a table of values:
 

- Customised Strategy 1: T= 1.0, P= 1.0, R= 1.0, S= 0, B= 1.0
- Customised Strategy 2: T= 1.0, P= 1.0, R= 0, S= 0, B= 0

**Fig. 2.** First round of the iterated prisoner's dilemma (human strategy on the left; bot on the right).

## 4 Methodology of the Power to Take Game

In this study we simulated iterated power to take games with bots as players in order to examine factors and behaviors that might be relevant for group decision making processes and reactions to taxation circumstances over time. The investigated factors include personality type, team size, cultural background, adjudication method and team composition. Differences in behavior can be caused by, among other things, different personality traits that have been measured in models such as the Five Factor Model (FFM) [10]. For this study, a similar instrument, the Myers–Briggs Type Indicator (MBTI) [9], that nevertheless taps into personality traits described by the FFM [10] was employed to model automated bot behavior after human behavior in the power to take game.

### 4.1 Bias–Based Reasoning

In this study, we decided to implement a bias–based Knowledge Based Expert System (KBES) as described by Hancock [6]. The expert system allowed us to embed several heuristics for strategies into the player bots that might be employed by humans in the power to take game. Just as experts are biased by *a priori* beliefs, bias–based reasoning was used to determine the bots' *a priori* biases in making game–related decisions. This enabled the bots to react dynamically to changing game situations based on their personality–based preferences and disinclinations towards different factors and strategies.

A collection of preferences  $p_i \in [0, 1]$  for a certain factor or strategy and a collection of disinclinations  $d_i \in [0, 1]$  for that same factor or strategy are accumulated into the overall preference  $p$  and disinclination  $d$  as follows:

$$\begin{aligned} p &= p + p_i (1 - p) \\ d &= d + d_i (1 - d) \end{aligned} \tag{1}$$

with  $p := 0$  and  $d := 0$  at the beginning. This assumes values in  $\in [0, 1]$ .

This method was used to compute the strength of the bots' reactions to changing circumstances. Additionally, it was employed to determine their preference for certain personality-based factor we deemed relevant for behavior during the power to take game.

Given the preference  $p$  and disinclination  $d$  of a bot towards a certain factor or situation, the attitude  $\Delta$  of a bot was calculated as follows:

$$\Delta := p - d \tag{2}$$

$\Delta$  assumes values between  $-1$  and  $1$ , whereby values  $>0$  imply a preference for that factor in questions, whereas values  $<0$  imply a preference for the opposite. Since we were mainly interested in the positive preference towards a specific factor and not its opposite, we treated values  $<0$  as  $0$ .

## 4.2 Game Basics

In the first step, the game loads a configuration file that specifies: the number of rounds, the amount of money each round, adjudication method, prior take rate history, exact team composition in terms of MBTI type, the personality extremity, and the minimum variability that each bot exhibits. Based on these specifications, the bots are initialized for each team and play the game as described below.

**Bot Creation.** Given the MBTI four letter code, as well as the two parameters lower bound  $l \in [0, 1]$  and upper bound  $u \in [0, 1]$  that describe how extreme the personality is supposed to be:

1. For each dimension with  $n$  questions, randomly generate between  $l \cdot n$  and  $u \cdot n$  positive answers to the corresponding questions in the MBTI questionnaire. The random generation is uniform.  
For example, for the I-E-dimension with  $n = 10$  questions of an introvert,  $l = 0.8$  and  $u = 0.9$ , randomly generate 8 or 9 introvert answers plus, correspondingly, 2 or 1 extrovert answers.
2. Based on the bot's specific answers to the MBTI questionnaire, compute various aspects of personality that are relevant to determine employed strategies.



**Game Loop.** In each game round:

1. Each responder bot earns the specified amount of money.
2. Taker bot decision:
  - (a) The taker bots decide individually what percentage of the responder bots' income they want to claim.
  - (b) The game collects these individual taker decisions.
  - (c) Each taker bot decides if it wants to adjust its decision based on the opinions of its team members.
3. The game adjudicates the final take rate.
4. Responder bot decision:
  - (a) The responder bots decide individually how much of their own income they want to destroy.
  - (b) The game collects these individual responder decisions.
  - (c) Each responder bot decides if it wants to adjust its decision based on the opinions of its team members.
5. Remove the destroyed income from the responder bots.
6. Take the taker rate from the remaining responder bot income and add it to the collective money pool of the taker bots.

### 4.3 Relevant Personality Type-Based Factors

In similar studies involving MBTI and bots, strategies were assigned to the bots based on dimensions alone, sometimes taking into consideration the strength of the preference for that dimension [16]. However, after looking at a version of the MBTI test, different question in the same MBTI dimension appeared to imply different preferences for strategies in the power to take game. Thus, going by dimensions alone seemed insufficient for this context. On the other hand, neither did it seem to be a good idea to look at each test question individually in order to extrapolate entire strategies from that single question alone.

Therefore, we focused on several factors related to the information that was available to the bots during the game, or that were identified in previous studies as potentially relevant for the power to take game [1, 2]. We then looked at each question in the MBTI test and assigned values representing if a particular answer implied a preference or disinclination towards that factor, and how pronounced that preference or disinclination was. Bias-based reasoning was then used to compute the values of these factors for each bot during bot creation based on the specific answers it selected on the MBTI test. Some factors required additional computation that are explained below. An overview of the factors and their value ranges in extreme personalities can be found in Table 1.

**Consensus.** In group setting like our version of the power to take game, decisions are not made by individuals in a vacuum. Instead, they are exposed to the opinions of their team members. To what degree they are swayed by their team's opinions partly depends on their personality type. The Thinking versus Feeling MBTI dimensions seemed to be especially relevant for this factor.

**Table 1.** Factors relevant for the power to take game in MBTI types with 100% in each dimension. The value range of each factor is  $\in [0, 1]$ .

Type	Consensus	Experiment	Reactivity	Fairness	Greed	Information	Tradition
ENFJ	0.4391	0.5498	0.2166	0.5778	0.4222	0.0734	0
ENFP	0.6654	0.9813	0.4063	0.6322	0.3678	0	0
ENTJ	0	0.4151	0	0	1	0.8958	0
ENTP	0	0.7551	0	0	1	0.7189	0
ESFJ	0.5564	0	0.2192	0.2623	0.7377	0.0851	0.8673
ESFP	0.8351	0.4054	0.4085	0.3094	0.6906	0	0.1244
ESTJ	0	0	0	0	1	0.8967	0.5755
ESTP	0	0.2794	0	0	1	0.7206	0
INFJ	0.3656	0.3466	0.2003	0.5778	0.4222	0.0609	0
INFP	0.5801	0.6962	0.3815	0.6322	0.3678	0	0
INTJ	0	0.2478	0	0	1	0.8503	0
INTP	0	0.5304	0	0	1	0.6818	0
ISFJ	0.4731	0	0.2031	0.2623	0.7377	0.0733	0.8691
ISFP	0.7343	0.1096	0.3834	0.3094	0.6906	0	0.1283
ISTJ	0	0	0	0	1	0.8513	0.5792
ISTP	0	0.0172	0	0	1	0.6835	0
Min	0	0	0	0	0.3678	0	0
Mean	0.2906	0.3334	0.1512	0.2227	0.7773	0.4119	0.1965
Max	0.8351	0.9813	0.4085	0.6322	1	0.8967	0.8691

The consensus factor ranges in value from 0 to 1, whereby 0 implies that the bot is not influenced by its team’s opinions at all, whereas 1 implies that it wants to go entirely with the group consensus.

**Variability.** Human beings are not perfectly consistent in their behavior, and exhibit some degree of variability in their decisions. This is partly caused by factors unknown to the observer, but seems to be influenced by personality type as well. This factor seemed to be largely influenced by the MBTI Judging versus Perceiving dimension.

Given a base variability value  $b \in [0, 0.5]$  and the bot personality’s difference between preference and disinclination  $\Delta_{p,d} \in [0, 1]$  for variability, the variability  $v$  of a bot was computed as

$$v := (1 + \Delta_{p,d}) b \quad (3)$$

The variability factor thus ranges in value from  $b$  to  $2b \in [0, 1]$ , with higher values implying higher variability.

**Experiment.** This factor is similar to variability but taps into other aspects of personality as well. Namely, it describes how much a bot prefers to stick to



**Information.** The game is potentially played with a larger number of rounds. It seemed unlikely to us that human players would remember the entire game history perfectly, or care for it. Instead, we assumed that recall would be better for the last rounds and that some personality types would consider more information than others. This seemed to be mainly influenced by the Intuition versus Sensing, Perceiving versus Judging and Thinking versus Feeling MBTI dimension.

The information factor ranges in value from 0 to 1, where 0 implies that the bot only looks at immediate information, whereas 1 implies consideration of as much information as possible.

**Reactivity.** “Emotional hazard” was identified by Bosman and van Winken as a factor that decreased efficiency in power to take games with human players, since responders chose to destroy everything when experiencing negative emotions [1]. Therefore, we decided to introduce the reactivity factor that describes how strongly a bot is going to react if something goes against its wishes or expectations, or in other words, how volatile it is. The main relevant MBTI dimensions seemed to be Judging versus Perceiving and Thinking versus Feeling.

The reactivity factor ranges in value from 0 to 1, whereby 0 suggests no reaction in response to seemingly upsetting situations, whereas 1 suggests extreme reactions.

#### 4.4 Decision Making

**Individual Decision Making.** Each rule  $i \in \{1, \dots, n\}$  generates a proposed take or destroy rate  $x_i$  and corresponding weight  $w_i$ . The bot’s preliminary decision  $\bar{x}$  is the weighted mean

$$\bar{x} := \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \quad (5)$$

Since the weights depend on the personality factors, each proposed take or destroy rate only features as much into preliminary decision as it is of interest for the bot, and it is possible that the output of some rules is completely irrelevant for a bot.

**Taker Strategies.** A simple rule that fires each round is the “tradition rule” that casts a vote for the mean of the cultural background take rate. The importance of this rule depends entirely on how important the tradition factor is for the bot.

The “greed rule” is another simple rule that is about maximizing the collection by maximizing the take rate. The weight of this rule depends completely on how greedy the bot is.

The “fairness rule” is slightly more complex and attempts to propose a fair take rate that takes into consideration the bot’s own expectations as well as

the reactions of the responder team. The importance of this rule hinges on the fairness factor.

The “upset rule” is designed to add potentially irrational behavior in reaction to negative emotions caused by unmet expectations such as high destroy rates in response to culturally acceptable take rates, perceived fair take rates and exploitation of good will. Upset is accumulated using bias-based reasoning and depends on how important these factors are to the bot. How much the taker bot wants to increase the take rate depends on how upset it is, and the weight of this increase depends on the reactivity factor.

The most complex rule is the “strategy rule”. It contains common sense reasoning such as not increasing the take rate once the goal take rate is achieved, detecting and countering manipulation by the responder team as well as the employment of one of two different strategies that are selected based on the bot’s personality. Bots that are below average in the experiment factor try to introduce higher take rates with small, incremental increases, whereas bots above average in the experiment factor try to achieve their goal take rate by experimenting with larger, but targeted increases in take rate. The weight of these strategies depends on the greed, experiment and information factors as these seemed to be the factors most relevant for opportunistic and strategic thinking.

**Responder Strategies.** The responder bots possess a “greed rule” as well that attempts to maximize net earnings by minimizing the destruction. The degree by which the destroy rate is lowered and the weight of the proposed destroy rate depends on the greed factor.

The responders’ “upset rule” is similar to that of the taker bots and leads to high destroy rates when the bot exhibits a high reactivity and the take rate deviates too much from the cultural background, perceived fair take rate or when lowering of the destroy rate was exploited by the taker team.

The “strategy rule” implements more sophisticated strategic thinking that uses the destroy rate as a signal of what the responders consider to be an acceptable take rate. Additionally, it attempts to detect and circumvent taker strategies aiming to increase the take rate. The weight of this rule depends on the greed, experiment and information factors.

**Adjustment of Individual Decisions Based on Group Opinion.** After the bots made their individual decisions, they get the opportunity to adjust their opinion based on the mean of the individual decision in the entire team. In our model, the inclination of a bot to reconsider its opinion depended on the consensus factor as well as the (un-)certainty regarding its own opinion. The closer the take or destroy rates proposed by the different rules, the more certain the bot is that it made the right decision. Whereas a wide spread in possible rates leaves the bot conflicted and uncertain.

Given the proposed rates  $x_i$  and corresponding weights  $w_i$  of rules  $i \in \{1, \dots, n\}$ , as well as the number of non-zero weights  $m$  and the weighted mean  $\bar{x}$ , the uncertainty  $u$  is the weighted sample standard deviation:

$$u := \min \left( 0.5, \sqrt{\frac{\sum_{i=1}^n w_i \cdot (x_i - \bar{x})^2}{\frac{m-1}{m} \sum_{i=1}^n w_i}} \right) \quad (6)$$

With the consensus factor  $c$ , the adjustment rate  $a$  is computed as follows

$$a := \frac{1}{2}c + u \quad (7)$$

Given the mean of the individual decisions in the group  $g$  and the bot's own, preliminary decision  $\bar{x}$ , the final decision  $d$  is then

$$d := ag + (1 - a) \bar{x} \quad (8)$$

Lastly, the variability  $v$  is applied to the reconsidered decision by choosing a uniformly random number from the interval  $[\max(0, d - v), \min(1, d + v)]$ .

## 4.5 Experiment Design

In a first experiment, we investigated the performance of teams consisting of only one MBTI type with extreme personalities (0.8–0.9 in each dimension) playing other such teams, with different combinations of team size and cultural background.

In a second experiment, we introduced two additional team compositions, namely a team consisting of the real MBTI type distribution in the general U.S. population [13] and a team consisting of a random sample of that real distribution. We constructed all of these teams both with a personality extremity of 0.8–0.9 and a more realistic 0.55–1.0 as control, while keeping the team size constant.

In a third experiment, mixed teams of the best and worst takers and best and worst responders played against each other as well as against a sample of the real distribution. See Table 3 for the precise experiment design.

## 4.6 Implementation

Game configuration files were generated in YAML format. Configuration file generation, bot creation, the power to take game and experiment analyses were all implemented in Python 3 [15]. The simulations were run and tested on Windows 10.

# 5 Results

## 5.1 Prisoner's Dilemma

Mean score for Feeling types was 110.8, standard deviation 43.4, for the Thinking types 130.8, 40.6, respectively. A one-way t-test is significant at  $p = 0.037$ , suggesting improved expected performance of T type personality over the F type (Table 4).

**Table 3.** Experiment design for the power to take game.

	Experiment 1	Experiment 2	Experiment 3
# Rounds	100	100	100
Money/round	100	100	100
Adjudication	Mean	Mean, median	Mean, median
Countries	USA, Germany, UK, Japan, Russia	USA, Germany, UK, Japan, Russia	USA, Germany, UK, Japan, Russia
Variability	0.025	0.025, 0.05	0.01, 0.025, 0.05
Type extremity	0.8–0.9	0.8–0.9, 0.55–1.0	0.8–0.9, 0.55–1.0
Team sizes	1, 2, 3, 4, 100	100	2, 4, 10
Composition	Pure types	Pure types, real type distribution, random sample of real type distribution	Team of 2 best takers, team of 2 best responders, team of 2 worst takers, team of 2 worst responders, random sample of real type distribution
# Configurations	6,400	25,920	9,000

## 5.2 Power to Take

Examination of the factor value ranges of extreme personalities in Table 1 shows that all T types have no interest in consensus or fairness, are not reactive and prioritize maximizing their own gain. In addition, N types are not interested in tradition, which is also the case for STPs. SJs exhibited the highest interest in tradition and SFPs exhibited a small amount. It is notable that all types are greedy to some degree and no type is perfectly fair, with the NFPs being the most concerned with fairness followed by the SFPs.

**Experiment 1.** Examination of experiment 1 results identifies ENTPs and INTPs as the best taker personalities for highest tax collection rates and ISFJs and ESFJs are the worst personalities. Regarding net earnings, the best responder personalities are INTJs and ESTPs, and the worst personalities are ENFPs and INFPs. These trends hold across different team sizes and cultural backgrounds.

The best matchup for takers in terms of highest tax collection rate were ISTJs as responders, and an examination of matchups between ENTPs and ISTJs clearly demonstrates the ability of the ENTP taker bots to impose high take rates on the ISTJ responder bots and without triggering high destroy rates. In contrast, the worst matchup for taker bots are ENFPs as responders, since

**Table 4.** Results of 30 iterated Prisoner’s Dilemma simulations.

Round	User MBTI	User strategy	User points	Bot MBTI	Bot strategy	Bot points
1	ESFJ	Always Cooperate	150	INTJ	Tit for Tat	150
2	ENFJ	Random	117	INTP	Naïve Prober	112
3	ESFJ	Naïve Peacemaker	46	ESTP	Always Defect	66
4	ISFJ	Naïve Peacemaker	150	INTP	Tit for Two Tats	150
5	ESFJ	Always Cooperate	75	INTP	Random	200
6	ENFP	Always Cooperate	0	ENTJ	Always Defect	250
7	ENFP	Random	112	ISTP	Tit for Two Tats	112
8	INFP	Remorseful Prober	124	ISTJ	Naïve Prober	129
9	INFJ	Tit for Tat	110	INTP	Random	115
10	INFJ	Remorseful Prober	116	ISTJ	Random	111
11	ESFJ	Random	96	INTJ	Random	121
12	INFJ	Always Cooperate	147	ESTJ	Naïve Prober	152
13	INFP	Random	126	ENTJ	Tit for Two Tats	126
14	ESFP	Naïve Peacemaker	150	INTP	Tit for Tat	150
15	ENFJ	Always Cooperate	150	INTP	Tit for Two Tats	150
16	ENFP	Random	122	ISTJ	Naïve Prober	122
17	ESFJ	Random	116	ISTJ	Tit for Two Tats	111
18	ENFP	Random	23	ISTJ	Always Defect	158
19	ENFJ	Tit for Tat	49	ENTP	Tit for Two Tats	54
20	ESFP	Always Cooperate	150	ENTP	Tit for Two Tats	150
21	ENFP	Remorseful Prober	149	ENTJ	Tit for Tat	149
22	INFJ	Always Cooperate	147	ISTJ	Naïve Prober	152
23	ISFJ	Tit for Tat	125	ISTP	Random	125
24	ISFJ	Random	102	INTJ	Tit for Two Tats	122
25	INFP	Remorseful Prober	150	ESTP	Tit for Tat	150
26	ENFJ	Tit for Tat	49	INTJ	Always Defect	54
27	INFJ	Remorseful Prober	49	INTJ	Always Defect	54
28	ISFJ	Naïve Peacemaker	125	ENTP	Naïve Prober	130
29	ESFJ	Tit for Tat	150	INTP	Tit for Tat	150
30	ISFJ	Naïve Peacemaker	150	ENTP	Naïve Prober	150
Total points			3,325			3,925

they resort to high destroy rates that lead to lowering of the economy size, tax collection for the takers and also lowering of the net earnings of the ENFP responders; a clear loss for all parties involved. The best overall matchup for responders in terms of highest net earnings were ISFJs as takers.

**Experiment 2.** Matchups of extreme personality types (0.8–0.9 in each dimension) versus the real type distribution in the general, human population with actual type extremity (0.55–1.0 in each dimension) exhibited the same patterns observed in experiment 1 (Table 5). Again, the ENTPs and INTPs were the best tax collectors, while ESFJ and ISFJ taker bots achieved the highest economy rate at the expense of their tax collection. Similar to the observations of experiment 1, the ENTP taker bots managed to establish extremely high take rates, resulting in high tax collection rates, while maintaining a moderate level of resistance by the general population.



**Table 5.** Power to take, experiment 2: Overall ranking of takers and responders with type extremity of 0.8–0.9 versus the real type distribution with type extremity of 0.55–1.0. The rates are the mean of what the teams achieved compared to what was possible in the corresponding games. The highest rates are marked in light gray, whereas the lowest rates are marked in middle gray.

Team	Takers				Responders			
	Economy Rate		Collection Rate		Net Earnings Rate		Destruction Rate	
ENFJ	88.35%	17	76.11%	8	33.50%	12	32.45%	4
ENFP	88.28%	18	80.13%	5	17.37%	18	67.54%	1
ENTJ	88.46%	13	83.62%	3	36.90%	1	14.96%	8
ENTP	88.73%	11	84.31%	1	30.54%	15	35.56%	3
ESFJ	91.59%	2	37.13%	17	33.77%	11	4.89%	14
ESFP	89.58%	8	65.71%	10	36.45%	4	14.98%	7
ESTJ	91.19%	3	42.18%	15	32.45%	13	3.22%	17
ESTP	88.66%	12	80.72%	4	36.86%	2	9.70%	9
INFJ	89.24%	10	68.08%	9	36.21%	6	16.28%	6
INFP	88.45%	15	77.91%	7	27.14%	17	47.30%	2
INTJ	88.40%	16	78.17%	6	36.80%	3	9.05%	11
INTP	88.45%	14	83.71%	2	36.36%	5	19.35%	5
ISFJ	91.69%	1	37.09%	18	32.06%	14	3.64%	16
ISFP	91.03%	5	44.59%	14	35.34%	9	6.71%	13
ISTJ	91.12%	4	42.05%	16	29.91%	16	2.65%	18
ISTP	89.46%	9	62.78%	11	34.10%	10	4.56%	15
real	90.70%	6	50.49%	12	36.01%	7	9.50%	10
sample	90.68%	7	46.72%	13	35.93%	8	8.84%	12

ENFP and INFP responder bots destroyed the most income against a taking authority from the general population, whereas ISTJ and ESTJ responder bots destroyed the least income. Unlike in experiment 1, ENTJ and ESTP responder bots performed the best in terms of maximizing net earnings for themselves, while ENFP and INFP responder bots performed the worst, which is likely caused by resorting to high destroy rates. These trends held across different matchups and other conditions under investigation.

It is important to note that both the real distribution and sample real distribution responders ranked in the middle tertile in terms of net earnings and destruction rate. This was also the case for such takers regarding economy rate and collection rate.

Table 6 presents an interesting result concerning the relationship between take rate and economy rate. The significant Pearson correlation of  $r = 0.4199$  ( $p < 0.001$ ) indicates a general trend of higher tax rates being associated with larger economies. However, this trend does not extend to the full range of take rates. In fact, the worst take rates for the economy rate are the extremely low and extremely high tax rates, a fact that should both be expected and desired.

**Table 6.** Power to take, experiment 2: Ranking of take rates regarding economy rate, collection rate and net earning rate with the real type distribution with type extremity of 0.55–1.0 as responders. The economy, collection and net earning rates are the mean of the rates in reaction to the respective take rate. Given are the Pearson’s correlation coefficient  $r$  and the  $p$ -value of the correlation based on a two-sided  $t$ -test.

	Take rate	Economy rate	Take rate	Collection rate	Take rate	Net earnings rate
<i>Best</i>						
1	28%	91.80%	97%	87.26%	15%	77.21%
2	39%	91.75%	98%	87.05%	14%	76.97%
3	26%	91.59%	99%	85.60%	16%	76.48%
4	45%	91.46%	96%	85.01%	12%	76.11%
5	22%	91.39%	95%	84.92%	13%	75.87%
<i>Worst</i>						
1	12%	86.72%	12%	10.61%	95%	1.25%
2	99%	87.21%	13%	11.49%	96%	1.98%
3	13%	87.36%	14%	12.60%	97%	2.59%
4	90%	87.91%	15%	13.58%	98%	3.62%
5	89%	88.10%	16%	14.55%	99%	4.41%
r	0.419929		0.996563		-0.99833	
p-value	<0.001		<0.001		<0.001	

**Experiment 3.** When teams comprised of different personalities are matched against each other, we can observe a moderating impact of the adjudication method. This is largely because compositions of different personalities will tend to exhibit higher intra-group variability, so the adjudication method makes a difference as the group size increases.

## 6 Discussion

### 6.1 Prisoner’s Dilemma

Results of simulations appear to suggest an association between one dimension of MBTI and expected reward in an iterated prisoners dilemma situation. In particular, it appears that a T type personality would be expected to outperform a F type personality. If true, that would be inconsistent with Dawkins’ findings suggesting nice players win. As in many cases, conditioning on different assumptions and considering a different set of strategies for the PD game for the T type and F type, with somewhat different overlap, may lead to different results. So in reality, the practical implication of this part of the study is that such methodology of bot utilization for personality related binary decision-making

processes can be of value. Enhancements of this kind to binary decision-making processes may have limited power and may be realized only after a large number of iterations.

## 6.2 Power to Take

Results based on 41,320 configurations, each played for 100 rounds, appear to suggest that personality type matters for this game, and by extension, matters in the context of taxing authorities interacting with tax payers. This means that choices of particular personalities could lead to improved or reduced performance in any of the measures that one may be interested in, i.e. total tax collection, size of economy and net earnings, compared to other personality choices, or compared to the general population with its mixed personality type composition.

It may be worth noting that prior studies primarily involving students playing the game individually or in groups, obtained results with somewhat undesirable characteristics, namely destruction rates that tended to be either 0% or 100%, which lead to the responders decisions appearing to be essentially binary in those studies. This binary approach by a responder (tax payer) may be rational if the game is played once, since anytime the tax rate is under 100%, responders would maximize net earnings for the game by destroying 0%.

However, taxing circumstances repeat, and the reality is that tax collection as a function tax rates is a continuous function, so an iterated approach may be more appropriate in order to understand the evolution of the tax system with bot-based systems offering an advantage in the examined scope. In a large number of configurations played, the prevailing trends were of increasing tax rates over time, decreasing destruction rates over time, and decreasing net earnings over time. Such circumstances are not unrealistic in the real world, so rules used for bot behaviors, and the aggregate belief and disbelief method may be fairly effective for this problem.

Nevertheless, a potential problem is that we observed a deviation in net earning ranking from the average household income ranking in the general population [3]. This could be an artifact of using extreme personality types versus a real type distribution, but could also be an indication that additional or different factors and strategies might need to be considered, or that our interpretation of the relationship between MBTI test answers and bot behavior does not reflect reality.

In addition, the association of results of such games using our approach to economics may be limited in its validity. An economy is dynamic and iterating an identical game a large number of times may not always be a good model. There were elements of variability infused into behaviors and rules, but the essential rules of the games remained the same in terms of matrix of payoffs for the prisoners dilemma and per game wealth initially awarded to each responder each round in the power to take game.

## 7 Conclusion

This study suggests that personality may play a role in both binary and continuous decision-making processes as well as both single-player and group settings, and an iterated approach using bots playing humans may enhance our understanding of human decisions, potentially leading to better processes and better outcomes.

Future research may elect to examine different personality instruments, perhaps based on the Five Factor Model and infuse dynamics to the economy. Furthermore, it might be worth examining a variant of the iterated power to take game where the take authority has an income that is subject to the take rate, just like the responders. That way, the take authority would not just be concerned with maximizing the take rate but also take into consideration net earnings and the overall size of the economy, similar to taxing authorities in an economy.

## References

1. Bosman, R., van Winken, F.: Emotional hazard in a power-to-take experiment. *Econ. J.* **112**(476), 147–169 (2002)
2. Bosman, R., Hennig-Schmidt, H., van Winden, F.: Exploring group decision making in a power-to-take experiment. *Exp. Econ.* **9**(1), 35–51 (2006)
3. Career Assessment Site: MBTI Socioeconomic Infographic. <http://careerassessmentssite.com/mbti-personality-types-socioeconomic-infographic/>. Accessed 13 Jan 2018
4. Davis, W.: Iterated Prisoners Dilemma Online Game and Simulation (1997). <http://www.iterated-prisoners-dilemma.net/>. Accessed 21 Jan 2018
5. Dawkins, R.: *The Selfish Gene*, 30th Anniversary edn. Oxford University Press, Oxford (2006). (Original Work Published in 1967)
6. Hancock, M.: *Practical Data Mining*. CRC Press, Boca Raton (2012)
7. Knight, V.: Strategies Index (2015). [http://axelrod.readthedocs.io/en/stable/reference/all\\_strategies.html](http://axelrod.readthedocs.io/en/stable/reference/all_strategies.html). Accessed 31 Jan 2018
8. KPMG: Individual income tax rates table. <https://home.kpmg.com/xx/en/home/services/tax/tax-tools-and-resources/tax-rates-online/individual-income-tax-rates-table.html>. Accessed 21 Jan 2018
9. MBTI Personality Test. <http://www.maximusveritas.com/wp-content/uploads/2017/07/MBTI-Personality-Type-Test.pdf>. Accessed 11 Dec 2017
10. McCrae, R.R., Costa, P.T.: Reinterpreting the Myers-Briggs type indicator from the perspective of the five-factor model of personality. *J. Pers.* **57**(1), 17–40 (1989)
11. Mortensen, C.D.: *Human Conflict: Disagreement, Misunderstanding, and Problematic Talk*. Rowman & Littlefield Publishers, Lanham (2005)
12. Myers, I.B.: *The Myers-Briggs Type Indicator (Manual)*. Consulting Psychologists Press, Inc., Palo Alto (1962)
13. The Myers & Briggs Foundation: How Frequent is My Type. <http://www.myersbriggs.org/my-mbti-personality-type/my-mbti-results/how-frequent-is-my-type.htm?bhcp=1>. Accessed 17 Jan 2017
14. Poundstone, W.: *Prisoner's Dilemma: John Von Neumann, Game Theory and the Puzzle of the Bomb*. Doubleday, New York, USA (1992)

15. Python Software Foundation: Python Language Reference, Version 3.6. <http://www.python.org>
16. Salvit, J., Sklar, E.: Toward a myers-briggs type indicator model of agent behavior in multiagent teams. In: Bosse, T., Geller, A., Jonker, C.M. (eds.) MABS 2010. LNCS (LNAI), vol. 6532, pp. 28–43. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-18345-4\\_3](https://doi.org/10.1007/978-3-642-18345-4_3)