games

Behavioral Game Theory

Edited by

Russell Golman

Printed Edition of the Special Issue Published in *Games*

www.mdpi.com/journal/games

MDPI

Behavioral Game Theory

Behavioral Game Theory

Editor

Russell Golman

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editor
Russell Golman
Carnegie Mellon University
USA

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Games* (ISSN 2073-4336) (available at: https://www.mdpi.com/journal/games/special_issues/Behavioral_Game_Theory).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-03943-773-3 (Hbk)
ISBN 978-3-03943-774-0 (PDF)

Contents

About the Editor

Russell Golman (Associate Professor of Behavioral Economics and Decision Sciences) studies behavioral economics and behavioral decision making in the Department of Social & Decision Sciences at Carnegie Mellon University. His pioneering, interdisciplinary research has been published in a wide range of academic journals, including *Cognitive Psychology, Decision, Journal of Economic Literature, Journal of Economic Perspectives, Journal of Economic Theory, Psychological Review, RAND Journal of Economics*, and *Science Advances*. Professor Golman organized the Belief-Based Utility Conference at Carnegie Mellon in 2017 with generous funding from the Russell Sage Foundation and the Alfred P. Sloan Foundation.

games

MDPI

Editorial

New Directions in Behavioral Game Theory: Introduction to the Special Issue

Russell Golman

Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA; rgolman@andrew.cmu.edu

Received: 26 October 2020; Accepted: 28 October 2020; Published: 1 November 2020

Behavioral game theory accounts for how people actually make strategic decisions by incorporating social utility, limited iterated reasoning, and learning [1]. The papers in this Special Issue span this space of behavioral game theory research.

Seier [2] in this Special Issue explores whether fairness in strategic games tends to be driven by intuitive or deliberative responses. Many people are willing to incur selfish costs to uphold norms of fairness, from promoting efficiency or equity to punishing others who violate these norms [3]. Fair behavior could follow from deliberation, with self-control being used to do the right thing despite an intuitive inclination to be selfish, or it could be an intuitive response that is adaptive in naturalistic contexts, but that can be overcome deliberatively in the lab in artificial contexts in which selfishness does not have reputational costs. Seier [2] finds that people who give more intuitive answers on the cognitive reflection task tend to make more fair choices in strategic games: they give away more money in the dictator game, demand more money as receivers in the ultimatum game, and engage in more costly third-party punishment of norm violators in a multiplayer game. For many people, social utility is a fundamental element of their preferences.

Zhao [4] in this Special Issue studies how the extent of iterated reasoning performed in a strategic decision depends on constraints on the other player's ability (as well as one's own ability) to engage in iterated reasoning. Using two-player guessing games in which strategic choices map cleanly onto levels of reasoning in a level-k model [5–7], Zhao [4] finds that players engage in more steps of reasoning when their opponents have been placed under a condition of lighter (rather than heavier) cognitive load, and this effect is stronger when players themselves are under lighter cognitive load, and thus able to engage in more steps of reasoning in the first place. That is, players are capable of recognizing that cognitive load may inhibit the reasoning ability of their opponents, and they respond appropriately. The observed pattern of behavior reflects an adaptive response that transcends the level-k reasoning model. Other models, including logit quantal response equilibrium [8,9], noisy introspection [10], and the dual accumulator model [11], can also account for limited iterated reasoning in guessing games, and manipulating the precision of logit responses in these models can also affect the depth of reasoning that an individual exhibits. A behavioral insight affirmed here, and consistent with all of these models, is that while people are boundedly rational, in that they are not capable of unlimited iterated reasoning, they do respond sensibly to changes in their opponent's incentives or constraints.

Guilfoos and Pape [12] in this Special Issue study how strategic behavior changes as players play a game repeatedly (with new opponents) and get feedback. They econometrically estimate case-based learning [13], reinforcement learning [14], and self-tuning experience weighted attraction [15], applied to Selten and Chmura's [16] dataset of 864 subjects repeatedly playing one of twelve 2×2 games. Case-based learning fits the observed behavior best, and also best predicts out-of-sample choices for a held-out slice of the data. Comparing the models based on out-of-sample prediction ensures that the empirical support for case-based learning is not an artifact of model flexibility and overfitting.

Guisasola and Saari [17] in this Special Issue introduce a coordinate system for the full space of 2×2 games that distinguishes changes in payoffs that exclusively affect: (i) the selfish costs and benefits of one's own strategies averaged uniformly across the other player's strategies (the "individual preference component"); (ii) the dependence of these selfish costs and benefits on the choice of the other player's strategy (the "coordinative pressure component"); (iii) the externality imposed on the other player by the choice of one's own strategy (the "pure externality component"); (iv) a constant level shift of all payoffs (the "kernel component"). The coordinate system is useful for a number of applications. This paper focuses on applying it to 2×2 potential games, including coordination games and anti-coordination games. Predictions based on individual selfish costs and benefits, including Nash equilibrium, risk-dominance (equivalently, the global maximum of the potential function), level-k reasoning, quantal response equilibrium, noisy introspection, and the dual accumulator model, are invariant to changes in the pure externality component of a game. However, changes in the pure externality component of the game do affect social welfare. Thus, it is straightforward to design games that pose a tension between the strategy predicted by any model based on individual selfish costs and benefits and the strategy that maximizes social welfare. The empirical fact that people care about social welfare as well as other aspects of the interaction between the externality component of a game and the individual preference component [18] indicates that any model of the individual reasoning process needs to be augmented with a model of social preferences to more fully capture behavior. The decomposition of 2×2 games in the coordinate system presented in this paper could be useful for experimental research by making it easier to independently test models of individual reasoning and models of social preferences.

Jamison [19] in this Special Issue explores the role of pre-play cheap talk among players with common knowledge of rationality. Whereas cheap talk is often dismissed as not credible because it is easily imitated, it may actually be informative when players have partially aligned incentives [20,21] or social preferences [22], such that, conditional on a statement being interpreted correctly, an individual wants to make the statement in the first place. In the absence of pre-play communication, common knowledge of rationality implies that players will choose rationalizable strategies, but not necessarily successfully coordinate on a Nash equilibrium. Jamison [19] shows that cheap talk allows rational players to reach (only) efficient Nash equilibria. Understanding pre-play cheap talk among rational players gives us a benchmark for studying pre-play cheap talk in laboratory games and in the real world; a context in which players are boundedly rational, may have incomplete information, and may have uncertainty or biases about each other's social preferences [23]. These behavioral elements allow communication to be informative in new and interesting ways [24–29].

References

1. Camerer, C.F. *Behavioral Game Theory: Experiments in Strategic Interaction*; Princeton University Press: Princeton, NJ, USA, 2003.
2. Seier, M. The Intuition of Punishment: A Study of Fairness Preferences and Cognitive Ability. *Games* **2020**, *11*, 21. [CrossRef]
3. Camerer, C.F.; Fehr, E. Measuring social norms and preferences using experimental games: A guide for social scientists. In *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*; Oxford University Press: New York, NY, USA, 2004.
4. Zhao, W. Cost of Reasoning and Strategic Sophistication. *Games* **2020**, *11*, 40. [CrossRef]
5. Costa-Gomes, M.A.; Crawford, V.P. Cognition and behavior in two-person guessing games: An experimental study. *Am. Econ. Rev.* **2006**, *96*, 1737–1768. [CrossRef]
6. Georganas, S.; Healy, P.J.; Weber, R.A. On the persistence of strategic sophistication. *J. Econ. Theory* **2015**, *159*, 369–400. [CrossRef]
7. Nagel, R. Unraveling in guessing games: An experimental study. *Am. Econ. Rev.* **1995**, *85*, 1313–1326.
8. Capra, C.M.; Goeree, J.K.; Gomez, R.; Holt, C.A. Anomalous behavior in a traveler's dilemma? *Am. Econ. Rev.* **1999**, *89*, 678–690. [CrossRef]
9. McKelvey, R.D.; Palfrey, T.R. Quantal response equilibria for normal form games. *Games Econ. Behave.* **1995**, *10*, 6–38. [CrossRef]

10. Goeree, J.K.; Holt, C.A. Ten little treasures of game theory and ten intuitive contradictions. *Am. Econ. Rev.* **2001**, *91*, 1402–1422. [CrossRef]
11. Golman, R.; Bhatia, S.; Kane, P.B. The dual accumulator model of strategic deliberation and decision making. *Psychol. Rev.* **2020**, *127*, 477–504. [CrossRef]
12. Guilfoos, T.; Pape, A.D. Estimating Case-Based Learning. *Games* **2020**, *11*, 38. [CrossRef]
13. Gilboa, I.; Schmeidler, D. Case-based decision theory. *Q. J. Econ.* **1995**, *110*, 605–639. [CrossRef]
14. Erev, I.; Roth, A.E. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* **1998**, *88*, 848–881.
15. Ho, T.H.; Camerer, C.F.; Chong, J.K. Self-tuning experience weighted attraction learning in games. *J. Econ. Theory* **2007**, *133*, 177–198. [CrossRef]
16. Selten, R.; Chmura, T. Stationary concepts for experimental 2x2-games. *Am. Econ. Rev.* **2008**, *98*, 938–966. [CrossRef]
17. Guisasola, S.; Saari, D. With Potential Games, Which Outcome Is Better? *Games* **2020**, *11*, 33. [CrossRef]
18. Charness, G.; Rabin, M. Understanding social preferences with simple tests. *Q. J. Econ.* **2002**, *117*, 817–869. [CrossRef]
19. Jamison, J. Valuable Cheap Talk and Equilibrium Selection. *Games* **2020**, *11*, 34. [CrossRef]
20. Crawford, V.; Sobel, J. Strategic information transmission. *Econometrica* **1982**, *50*, 1431–1451. [CrossRef]
21. Farrell, J.; Rabin, M. Cheap talk. *J. Econ. Perspect.* **1996**, *10*, 103–118. [CrossRef]
22. Golman, R. Good manners: Signaling social preferences. *Theory Decis.* **2016**, *81*, 73–88. [CrossRef]
23. Di Tella, R.; Perez-Truglia, R.; Babino, A.; Sigman, M. Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism. *Am. Econ. Rev.* **2015**, *105*, 3416–3442. [CrossRef]
24. Charness, G. Self-serving cheap talk: A test of Aumann's conjecture. *Games Econ. Behav.* **2000**, *33*, 177–194. [CrossRef]
25. Chaudhry, S.J.; Loewenstein, G. Thanking, apologizing, bragging, and blaming: Responsibility exchange theory and the currency of communication. *Psychol. Rev.* **2019**, *126*, 313–344. [CrossRef]
26. Crawford, V. A survey of experiments on communication via cheap talk. *J. Econ. Theory* **1998**, *78*, 286–298. [CrossRef]
27. Ellingsen, T.; Östling, R. When does communication improve coordination? *Am. Econ. Rev.* **2010**, *100*, 1695–1724. [CrossRef]
28. Gneezy, U. Deception: The role of consequences. *Am. Econ. Rev.* **2005**, *95*, 384–394. [CrossRef]
29. Gurney, N.; Loewenstein, G. Filling in the Blanks: What Restaurant Patrons Assume About Missing Sanitation Inspection Grades. *J. Public Policy Mark.* **2020**, *39*, 266–283. [CrossRef]

games

MDPI

Article

The Intuition of Punishment: A Study of Fairness Preferences and Cognitive Ability

Markus Seier

Department of Economics and Business Economics, Aarhus University, 8210 Aarhus V, Denmark; msp1996@live.dk

Received: 24 March 2020; Accepted: 20 April 2020; Published: 7 May 2020

Abstract: Can differences in cognitive reflection explain other-regarding behavior? To test this, I use the three-item Cognitive Reflection Task to classify individuals as *intuitive* or *reflective* and correlate this measure with choices in three games that each subject participates in. The main sample consists of 236 individuals who completed the *dictator game, ultimatum game* and a *third-party punishment task.* Subjects afterwards completed the three-item Cognitive Reflection Test. Results showed that *intuitive* individuals acted more prosocially in all social dilemma tasks. These individuals were more likely to serve as a norm enforcer and *third-party punish* a selfish act in the *dictator game*. *Reflective* individuals were found more likely to act consistently in a self-interested manner across the three games.

Keywords: social preferences; third-party punishment; cognitive reflection ability; intuition; reflection; dictator game; ultimatum game

1. Introduction

Human societies depend on their members acting cooperatively. Social sanctioning is crucial for the maintenance of cooperative behavior when there exist material incentives to deviate from collectively desirable behavior, such as benefiting from a public good without bearing the cost of contributing. Sanctioning behavior can be explained by strong reciprocity, which is defined by a willingness to sacrifice resources to reward cooperative actions and to punish hostile actions even when this is costly and provides neither present nor future material rewards for the reciprocator [1,2]. Thus, individuals acting as norm enforcers enable cooperative behavior because of an understanding and expectation that a deviation will be sanctioned [3]. Social dilemma experiments reveal a great deal of strong reciprocity. For example, in [4], the majority of subjects were willing to engage in *third-party punishment*. That is, they punished a hostile action even though it did not affect their personal earnings.

Is sanctioning a norm violation an *intuitive* response, or does it take *deliberation* to sacrifice resources? To the best of my knowledge this question has not been investigated in the context of *third-party punishment*, where there is no indirect benefit from sanctioning through reputation-building or long-term material incentives from changing the behavior of people one interacts with in the future.

More generally, is cooperative behavior driven by an *intuitive* response or due to *deliberation*? Whether individuals rely on *intuition* or *reflection* in social dilemma experiments has been shown to generate differences in behavior. Applying cognitive reflection tests [5,6], subjects relying on *intuition* in decision-making are found to act more prosocially [7–11].

I contribute to this literature by examining whether behavior is consistent across three games and whether sanctioning the violation of a norm is an *intuitive* action. Applying a *third-party punishment task*, subjects are given the opportunity to, at a personal cost, sanction another subject who kept the entire endowment to herself in the *dictator game.*

Studying subjects' response time has as well been applied to access whether individuals rely on *intuition* in decision-making. Results in these studies are, however, not conclusive about whether a faster response time indicate more prosocial [12] or more egoistic [13] behavior. Identifying whether a choice is *intuitive* or *deliberate* from response time suffers from endogeneity issues as various cognitive processes contribute to response time. When controlling for strength-of-preference, there is no evidence that one type of choice is systematically faster than the other [14].

According to the Social Heuristic Hypothesis, *intuitive* individuals behave more prosocially in the lab because they internalize generally beneficial behavior from daily life that favors cooperative and fair behavior [15]. In light of this, the employed experimental design in this study investigates how strong these internalized fairness preferences are.

The purpose of this study is twofold. By having subjects complete the *dictator game*, the *ultimatum game* (both in the role of *proposer* and *recipient*) and finally the *third-party punishment task* the purpose is first to see if subjects display consistent behavior across games in line with the hypothesis that the "fair" outcome drives *instinctive* choices but that it takes *deliberation* to act selfishly. Secondly, this experiment investigates for the first time if the *instinctive* action is to engage in *third-party punishment* toward a *dictator* who kept the entire endowment to herself in the *dictator game*. The subjects' tendency to rely on *intuition* in decision-making is assessed by Frederick's three-item Cognitive Reflection Test (CRT) [5].

The sample consists of 295 students at Aarhus University, collected during spring 2019.

The results of this study confirmed, first of all, previous findings that *reflective* subjects act more selfishly and in accordance with the economic prediction in the *dictator*- and *ultimatum games*. They transferred less in *dictator game*, they offerred less as the *proposer* in the *ultimatum game*, and they were more likely to accept a low offer as *recipient*. Secondly, the experiment extended previous findings to *third-party punishment* by showing that the *intuitive* action was to sanction a norm-violator. Subjects relying on *intuition* in decision-making were found more likely to sacrifice resources to sanction a *dictator* who kept the entire endowment to herself. Taken together, the results of this experiment provide evidence that the *intuitive* action is to engage in "fair" behavior, or to sanction those not complying with the social norm of fair behavior.

In the following Section 2, I present the experimental design. The hypotheses are presented in Section 3. Section 4 presents the results of the experiment. Section 5 provides a general discussion of the findings. Section 6 discusses the limitations of this study. Section 7 concludes.

2. Experimental Design

2.1. Procedures

Subjects were recruited during four lectures in Psychology, Political Science, and Economics at Aarhus University. Three of these four lectures were for second semester students. The students were orally encouraged to participate during the break of the course and a link to the survey was distributed online.

Subjects were incentivized through a lottery scheme. In total, seven subjects were paid on average DKK 50 (≈$7.5) for completing the experiment and, in pairs, paid according to their choices in the task, for which they were drawn at random to receive payment. For each the *dictator game*, *ultimatum game*, and *third-party punishment task*, two subjects received payment. One subject was drawn to get paid for completing the CRT. For each correct answer on the CRT, one ticket was added to the bowl from where a subject was drawn. The subject received DKK 100 (≈$15) for completing the CRT.

2.2. Experimental Design

Subjects completed four social dilemma tasks: The *dictator game*, the *ultimatum game* with role uncertainty (i.e., subjects made choices in the role of both the *proposer* and the *recipient*) and decided whether to engage in *third-party punishment* by choosing if and how much to sacrifice to sanction a *dictator*, who kept the entire endowment to herself in the *dictator game*. After completing the four social dilemma tasks, subjects continued to the second part of the experiment to complete the three-item CRT.

Lastly, subjects were to state their gender, line of study and their email address in order to potentially get paid for participating in the experiment.

In the following, I will present each social dilemma task as well as the three-item CRT. The experimental instructions are reproduced in Appendix A.

2.2.1. Dictator Game

The first task was a standard *dictator game*. The subject acting in the role of the *dictator* was endowed with DKK 100 and had to decide on how much (in increments of DKK 10) to transfer to another subject acting as the *receiver*, with whom she was randomly matched. The *receiver* had no decision to make.

2.2.2. Ultimatum Game

For the second and third task, subjects were to make a decision first as *proposer* and later as *recipient* in the *ultimatum game*. The *proposer* is endowed with DKK 100 and chooses how much to offer (in increments of DKK 10) the *recipient*. The *recipient* indicates the minimum amount (acceptance threshold), she is willing to accept (in increments of DKK 10). If the offer is accepted, the proposed allocation is realized, and if the offer is rejected, both the *proposer* and the *recipient* receive nothing.

The strategy method [16] is employed to the *recipient's* decision because the sampling procedure allowed players to enter their choices at different time points. Even though applying the strategy method was necessary in this case, it is useful in the *ultimatum game*, since most offers are close to equal splits which means that there are few rejections, and thereby the actually relevant choices provide little information regarding the willingness to accept or reject low offers [17].

2.2.3. Third-Party Punishment Task

The fourth and final social dilemma task added a *third-party punishment* option to the *dictator game*. The subject is informed that she has been randomly matched to a pair of other subjects from the *dictator game*. One of the other subjects was assigned to the role of the *dictator* and chose to keep the entire endowment to herself[1]. The subject, who must decide on how much (if at all) to punish the *dictator* is endowed with DKK 50. For each DKK 1, the *third-party punisher* sacrifices, the *dictator* suffers a reduction in earnings of DKK 5. The *third-party punisher* must decide on how much to sacrifice between DKK 0 and DKK 20. By sacrificing DKK 20 of her own endowment, the *third-party punisher* can reduce the earnings of the *dictator* to DKK 0.

2.2.4. Three-Item Cognitive Reflection Test

After having completed the above-mentioned tasks, the subjects proceed to the three-item CRT [5]. The three-item CRT can be found in Appendix B.

The test is used to detect an individual's proclivity for applying two systems of decision-making: System 1 and System 2 processes [19]. System 1 is the intuitive "part" of the brain that relies on heuristics and automaticity. It possesses no computational capacity and is characterized as unconscious. It is fast, automatic and requires no effort. System 2 is the more analytical and rational system. It is deliberate and activated when facing complex calculations, different choices and requires the individual to be focused [20]. The performance on CRT indicates whether an individual is able to overcome the desire to go with the intuitive (incorrect) answer, reflect further upon the question and reach the, when explained to, relatively easy correct answer. For example the first question of the CRT: *A bat and*

[1] The experimental design applied actual matching on the subset of subjects who gave DKK 0 in the *dictator game*. Ex ante it could be expected that at least one subject would do so, based on previous *dictator game* experiments (In a meta study [18] found that 36.11% of all participants chose to give nothing).

a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost? ___ cents. Intuitive Answer: 10 / Correct Answer: 5.

Based on the answers to the CRT I divide subjects into three groups using the categorization used by [21]: Subjects who answered correctly two or more items on the CRT are categorized as *reflective*. Those opting for the intuitive, but wrong answer at least in two of the three items are *intuitive*. The subjects who are not categorized as either *reflective* or *intuitive*, form the residual group. For precise details of the categorization, see Appendix C.

3. Hypotheses

Looking to replicate previous findings of fair behavior by individuals relying on *intuition* in decision-making and that it takes *reflection* to pursue a self-interested objective gives three hypotheses in the *dictator-* and *ultimatum game* decisions.

Hypothesis 1. *Reflective individuals transfer less in the dictator game compared to intuitive individuals.*

Hypothesis 2. *Reflective individuals offer less as proposer in the ultimatum game compared to intuitive individuals.*

Hypothesis 3. *Reflective individuals require a smaller share to accept the offer as ultimatum game recipient compared to intuitive individuals.*

Including both the *proposer* decision in the *ultimatum game* and the transfer decision in the *dictator game*, it is possible to detect whether strategic considerations drive the *ultimatum game* offer. In the *dictator game,* such strategic considerations are absent, because it is a pure decision problem without strategic interaction. Expecting the *intuitive* action to be fair and *reflection* to lead to rational, self-interested decisions generates two hypotheses for *proposer* and *dictator* behavior.

Hypothesis 4a. *Reflective individuals offer more in the ultimatum game relative to their transfer in the dictator game.*

Hypothesis 4b. *Intuitive individuals do not offer more in the ultimatum game relative to their transfer in the dictator game.*

A main contribution of this study is the investigation of whether the *intuitive* action is to sanction those who violated the norm of fair behavior.

Hypothesis 5. *Intuitive individuals exhibit a greater willingness to punish a selfish dictator than reflective individuals.*

The other contribution to the existing literature is that this study investigates the behavior across four social dilemma decisions.

Hypothesis 6. *Reflective individuals act consistently more rational and self-interested in the four social dilemma decisions compared to intuitive individuals.*

4. Results

A total of 295 subjects completed the study. The main sample consists of 236 observations, for which all variables of interest are available. Of the 236 subjects in the main sample, 124 (52.5%) were male subjects (one subject did not state gender). 214 of the subjects were students at the faculty of Business and Social Sciences at Aarhus University, which leaves a minority from other faculties. This is not surprising, because the courses where the study was advertised are available in the faculty of Business and Social Sciences.

In each task, a few subjects chose the opposite extreme of strict self-interest (transferring DKK 100 in the *dictator game* and offering DKK 100 in the *ultimatum game* and accepting no less than DKK 100 in the *ultimatum game*). These "outliers" are included in the analysis. Excluding them does not alter the findings.

4.1. Cognitive Reflection Test Results

On average, the subjects answered 2.1 of the items on the CRT correctly. Of the 236 subjects, 48.7% answered all three items correctly, 24.2% answered two correctly, 14.4% answered one correctly and 12.7% did not answer any of the three items correctly. 9% of the subjects opted for the *intuitive* incorrect answer in all three items, 23.3% chose the *intuitive* answer in at least two items and 45.8% chose the *intuitive* incorrect answer at least once.

The *reflective* group consists of 172 subjects. The *intuitive* group consists of 56 subjects. The residual group consists of 8 subjects. As the residual group consists only of 8 subjects, these are grouped with the *intuitive* subjects throughout the statistical analysis. Therefore, the analyses mainly compares those *reflective* to those not *reflective.* The *non-reflective* group consists therefore of 64 subjects. Excluding the residual group, and thereby comparing the *reflective* to the *intuitive* subjects, does not change conclusions. (See Appendix D (Tables A1–A6, Figures A1–A5) for a summary of the findings excluding the residual group).

Men performed better in the CRT by answering an average of 2.3 items correctly compared to women with an average of 1.84 correct answers. This difference is statistically significant ($p = 0.003$, MWU[2]). The distribution of the answers can be found in Appendix E (Tables A7–A11).

In the following subsections, I will present the results for each of the tasks in the experiment. A graphical representation of the frequency of decisions consistent with rational, self-interested behavior by *non-reflective (reflective)* individuals can found in Figure 1. A more detailed presentation of decisions in each task can be found in Appendix F (Tables A12–A16, Figures A6–A9).

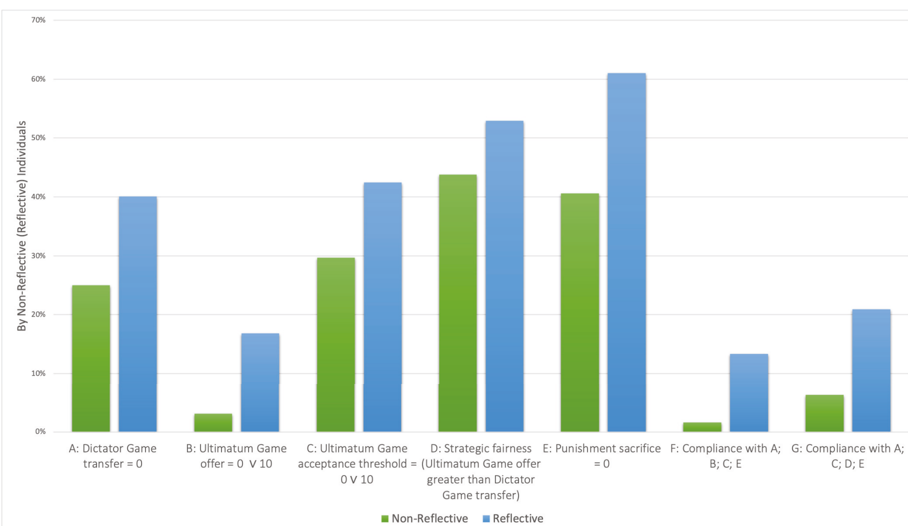


Figure 1. Frequency of Decision by Non-Reflective (Reflective) Individuals.

[2] Mann-Whitney-U: Note that the MWU is a test of differences in distribution.

4.2. Dictator Behavior

Result 1: Reflective subjects transfer less in the dictator game than intuitive subjects.

Reflective subjects transfer on average less than those not *reflective* (average transfer of DKK 28.2 and DKK 36.4, respectively). This difference is statistically significant at the 5% significance level ($p = 0.03$, MWU).

The average amount transferred to the *recipient* in the *dictator game* was DKK 30.4. The modal transfer was DKK 50, which 44.5% of the subjects chose, whereas 36% of the subjects chose to keep the entire endowment to themselves.

Transferring 0 DKK to the *receiver* and thus comply with the prediction from standard economic theory is more common for the *reflective* subjects (40.1% chose this versus 25% of the *non-reflective*). This difference is statistically significant at the 5% level ($p = 0.032$, $\chi^2 - test$). However, a part of the difference can be contributed to gender: Males are found significantly more likely to transfer DKK 0 to the *receiver* in the *dictator game* Thus, it appears that acting selfish in the *dictator game* is independent of being *reflective* when controlling for gender. Gender seems to be the significant factor that predicts behavioral differences (see Table 1).

4.3. Proposer Behavior in the Ultimatum Game

Result 2: Reflective subjects offer less in the ultimatum game than intuitive subjects.

Reflective subjects offer on average less than those not *reflective* (average offer of DKK 40.9 and DKK 50.5, respectively). This difference is statistically significant ($p = 0.0001$, MWU).

The average offer in the *ultimatum game* was DKK 43.5. The most frequently offered amount was DKK 50, which 68.6% of the subjects chose.

Of the *reflective* subjects, 15.7% offered DKK 10. Only one subject (1.8%) from the *intuitive* group offered DKK 10.

Distinguishing whether the *recipient* accepts or rejects an offer when indifferent, both offers of DKK 0 and DKK 10 can be considered consistent with rational and strictly self-interested behavior. 16.9% of the *reflective* subjects chose either of these offers as opposed to 3.1% of the *non-reflective*. This difference is statistically significant ($p = 0.005$, $\chi^2 - test$).

When controlling for gender, *reflective* subjects are estimated to be 12.6%-points more likely than *non-reflective* subjects to offer DKK 0 or DKK 10 in the *ultimatum game*. *Reflective* subjects are predicted to choose such an offer with a probability of 16.2% as opposed to a predicted probability of 3.6% for those *non-reflective* (see Table 1).

4.4. Recipient Behavior in the Ultimatum Game

Result 3: Reflective subjects are willing to accept lower offers in the ultimatum game than intuitive subjects.

Reflective subjects have on average a lower acceptance threshold relative to those not *reflective* (average threshold of DKK 27.8 and DKK 33.9, respectively). This difference is statistically significant at the 5% significance level ($p = 0.032$, MWU).

The average acceptance threshold was DKK 29.45. The modal acceptance threshold was DKK 10 and was chosen by 32.2% of the subjects whereas DKK 50 (requiring an equal split) was chosen by 29.7% of the subjects.

For the *reflective* subjects, the modal acceptance threshold was DKK 10, which was chosen by 36.6% in this category as opposed to 21.4% in the *intuitive* category. The modal acceptance threshold for the *intuitive* subjects was DKK 50, which was chosen by 37.5% in this category as opposed to 25% in the *reflective* category.

Both an acceptance threshold of DKK 0 or DKK 10 can be considered the rational, self-interested choice. 42.4% of the *reflective* subjects chose one of these thresholds as opposed to 29.7% of the *non-reflective* subjects. This difference is statistically significant at the 10% significance level ($p = 0.074$, $\chi^2 - test$). When controlling for gender, *reflective* subjects are estimated to be 11.4%-points more likely, compared to *non-reflective* subjects, to choose an acceptance threshold of DKK 0 or DKK 10 as *recipient* in the *ultimatum game*. *Reflective* subjects are predicted to choose such an acceptance threshold with a probability of 42.2% as opposed to a predicted probability of 30.8% for those *non-reflective* (see Table 1).

4.5. Dictator/Proposer Comparison

Result 4: Both reflective and intuitive subjects increase their offer in the ultimatum games relative to their transfer in the dictator game.

Across all subjects, the average transfer in the *dictator game* was DKK 30.4 and the average offer in the *ultimatum game* was DKK 43.5. Applying a Wilcoxon Sign Rank test, these means are significantly different ($p < 0.001$). Applying the test when distinguishing between *reflective* and *intuitive* subjects yields the same conclusion ($p's < 0.001$). Thus, both the *reflective* and *intuitive* subjects increase their offer in the *ultimatum game* relative to their transfer in the *dictator game*.

More than half of the subjects (50.4%) chose to increase their offer in the *ultimatum game* compared to their transfer in the *dictator game*—exhibiting strategic fairness. 52.9% of the *reflective* and 43.8% of the non-*reflective* subjects opted for this decision. This difference is not statistically significant ($p > 0.21$, $\chi^2 - test$).

When controlling for gender, *reflective* subjects are estimated to be 4%-points more likely to exhibit strategic fairness than *non-reflective* subjects. However, the effect is not statistically significant. *Reflective* subjects are predicted to exhibit strategic fairness with a probability of 51.3% as opposed to a predicted probability of 47.3% for those *non-reflective* (see Table 1).

4.6. Third-Party Punishment Behavior

Result 5: Intuitive subjects are more likely to punish a selfish dictator than reflective subjects.

Of the 236 subjects, 105 chose to punish the *dictator*, who kept the entire endowment to herself. The average amount sacrificed was DKK 4.8 which implies that a selfish *dictator*, on average, had her income reduced by DKK 24. The modal amount sacrificed was DKK 0, which 55.5% of the subjects chose. 10.2% of the subjects chose to reduce the earnings of the selfish *dictator* to DKK 0 by sacrificing DKK 20 of their endowment. 15.3% of the subjects chose to reduce the *dictator's* earnings by DKK 50 leaving the *dictator* with half of her initial endowment.

57.1% of the *intuitive* subjects chose to punish as opposed to 39% of the *reflective* subjects. This difference is statistically significant ($p = 0.017$, $\chi^2 - test$). The *reflective* subjects sacrificed, on average, DKK 3.97 as opposed to DKK 6.69 sacrificed by *intuitive* subjects. This difference is statistically significant ($p < 0.01$, MWU). Comparing the *reflective* subjects to those not *reflective* yields the same conclusion.

Considering only the subjects who opted for the opportunity to *punish* the selfish *dictator*, the *intuitive* subjects sacrificed, on average, DKK 11.7 as opposed to DKK 10.2 by the *reflective* subjects. This difference is not statistically significant ($p > 0.32$, MWU).

When controlling for gender, *reflective* subjects are estimated to be 20.1%-points more likely to not punish the dictator than *non-reflective* subjects. *Reflective* subjects are predicted to not engage in *third-party punishment* with a probability of 61.2% as opposed to a predicted probability of 41.1% for those *non-reflective* (see Table 1).

4.7. Consistency in Choices

Result 6: Reflective subjects are more likely to act consistently and in line with rational, self-interested behavior across all social dilemma tasks compared to intuitive subjects.

A rather clear prediction for rational, self-interested behavior exists for the *dictator game*, *recipient's* acceptance threshold in the *ultimatum game,* and the *third-party punishment task*. However, the decision as *proposer* in the *ultimatum game* is rather difficult to classify as expectations for the decision of the *recipient* matter. Thus, any offer can be considered rational, self-interested if that is the lowest amount the *proposer* expects to be accepted.

Due to the ambiguity in what constitutes rational, self-interested behavior in the *ultimatum game proposer* decision, I will consider offering DKK 0 or DKK 10 and strategic fairness separately.

First I consider whether *reflective* subjects are more likely to transfer DKK 0 in *dictator game*, offer DKK 0 or DKK 10 as *proposer* in the *ultimatum game*, acceptance threshold of DKK 0 or DKK 10 as *recipient* in the *ultimatum game* and not opting for the *punishment* opportunity in the *third-party punishment task*.

13.4% of the *reflective* subjects complied with the above-mentioned as opposed to 1.6% of those not *reflective*. This difference is statistically significant ($p = 0.008$, $\chi^2 - test$). When controlling for gender, *reflective* individuals are predicted to be 10.8%-points more likely than *non-reflective* subjects to choose as described in these tasks. *Reflective* subjects are predicted to choose as described with a probability of 12.7% as opposed to a predicted probability of 1.9% for those *non-reflective* (see Table 1).

A rational, self-interested individual could, as *proposer* in the *ultimatum game*, offer any share to the *recipient* if this is what the *proposer* believes to be the lowest amount to be accepted. However, in the *dictator game* there is no scope for such strategic considerations why a rational, self-interested individual would offer more as *proposer* in the *ultimatum game* relative to the transfer in *dictator game*. Considering whether *reflective* subjects are more likely to transfer DKK 0 in *dictator game*, have an acceptance threshold of DKK 0 or DKK 10 in the *ultimatum game*, exhibit strategic fairness as *proposer* in the *ultimatum game* and not opting for the *punishment* opportunity in the *third-party punishment task*, I find this to be the case. 20.9% of the *reflective* subjects complied with the above-mentioned as opposed to 6.3% of those not *reflective*. This difference is statistically significant ($p = 0.008$, $\chi^2 - test$). When controlling for gender, *reflective* subjects are predicted to be 12.2%-points more likely than *non-reflective* subjects to choose as described in these tasks. *Reflective* subjects are predicted to choose as described with a probability of 19.8% as opposed to a predicted probability of 7.6% for those *non-reflective* (see Table 1).

Table 1. Marginal effects from Logistic regressions.

VARIABLES	(1) A	(2) B	(3) C	(4) D	(5) E	(6) F	(7) G
1.Reflective	0.098	0.126 ***	0.114	0.040	0.201 ***	0.108 ***	0.122 **
	(0.069)	(0.037)	(0.071)	(0.074)	(0.074)	(0.032)	(0.048)
1.Male	0.233 ***	0.091 **	0.072	0.210 ***	0.029	0.089 **	0.165 ***
	(0.062)	(0.042)	(0.065)	(0.065)	(0.065)	(0.038)	(0.048)
Observations	235	235	235	235	235	235	235

Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. A: Dictator Game transfer = 0, B: Ultimatum Game offer = 0 or = 10, C: Ultimatum Game acceptance threshold = 0 or = 10, D: Strategic fairness; Ultimatum Game offer greater than Dictator Game transfer, E: Punishment sacrifice = 0, F: Compliance with A; B; C; E, G: Compliance with A; C; D; E.

5. Discussion

In line with several other studies, this study found more rational, self-interested behavior among more *reflective* individuals and more prosocial behavior among *intuitive* individuals. Further, this study found the more prosocial behavior among *intuitive* individuals to carry over to the *third-party punishment task*, where these individuals were found more likely to sanction a selfish act. A contribution of the present study was that subjects were to complete multiple social dilemma task, which allows to investigate the consistency across choices. In this aspect, *reflective* individuals were found more likely to act rationally in accordance with their self-interest across all four decisions.

Intuitive subjects give more in the *dictator game*, which is consistent with the findings of [7]. Transferring a positive amount to the *receiver* in the *dictator game* could be interpreted as altruistic preferences [2]. However, the findings of more rational and self-interested behavior by *reflective* subjects should be interpreted carefully, as gender seems to be the significant factor that drives differences in behavior in the *dictator game*. This is consistent with the findings of women giving more in a meta study on the experiments testing for gender differences [18].

In the *ultimatum game*, *reflective* subjects offered less than those not *reflective*. The decision of the *proposer* can be explained either by a "taste for fairness" or a "fear of rejection" (or a combination of these motives) [22]. Including the *dictator game* allows the inference with which motive matters for which group. However, the results indicate that both groups seem to act on a "fear of rejection". These findings contradict the findings of difference in transfer/offer being driven mostly by *reflective* individuals [10]. Even though "strategic fairness" appears to exist among both groups, the offers of the *intuitive* individuals are larger than those of the *reflective*. Thus, *intuitive* individuals appear to expect their offers in the *ultimatum game* to more likely be rejected. This is consistent with the consensus effect [23]. *Intuitive* individuals require a larger amount to accept an offer themselves.

Reflective subjects are more likely to accept offers in the *ultimatum game*, which confirms the findings of [8,9]. In those studies, the "strategy method" was not applied to the recipient's decision. Thus, *reflective* individuals exhibit a greater willingness to accept an unfair *ultimatum game* offer even when they are not directly faced with and possibly offended by the offer. Whether or not the strategic version of the *ultimatum game* induces lower acceptance thresholds is to some degree addressed in [24]. In this study, besides from playing the extensive form of the game, the subjects were required to state the minimum offer she would be willing accept. They found a significant negative correlation between the acceptance threshold and proposed offer which can be interpreted in light of *reflective* behavior. These individuals understand the bargaining position of the game as well as the risk of being rejected. Considering "negative reciprocity" as the motive for rejecting unfair offers in the *ultimatum game*, *reflective* individuals are more capable of overcoming their *intuitive* desire to *punish* the selfish act by the *proposer*. The willingness to accept an unfair offer is related to the ability to reflect further upon the decision and realize that accepting the offer is the better option.

Intuitive subjects are more likely to engage in *third-party punishment* and *reflective* subjects appear again more likely to act rational and self-interested. Thus, *intuitive* individuals are interpreted to be more likely to act reciprocally.

6. Limitations

Some factors related to the experimental design may have influenced how subjects behaved.

As the link to the survey were distributed at lectures encouraging students to participate, it is unknown when, where and possibly with whom the subjects completed the survey. Hence, there is concerns regarding their anonymity. Considering the relatively high share of correct answers in the CRT, one could expect subjects to have communicated with each other or have accessed the internet to look up the correct answer. Further, the chances of receiving payment for completing the CRT depended on the number of correct answers, which might have further incentivized subjects to look up the correct answer - at least incentivized them to think more carefully about the question, which was unintended. These limitations question whether the categorization of subjects is reliable. A reasonable explanation for the relatively high share of correct answers on the CRT in this study is the test's correlation with math abilities [5]. The vast majority of subjects were students of Economics, Political Science or Psychology. Especially students of Economics are expected to be relatively more capable of math. The survey questions did not elicit from which education the subjects were enrolled.

Only seven of the 295 subjects who completed the study received payment, providing only weak incentives. However, the observations here fit rather well the observations from other studies with stronger economic incentives. In a meta study, the average transfer was found to be 28% of the endowment [18], which is not far from 30.4% observed in this study. In a meta study on the ultimatum game, subjects were found to offer 40% of the endowment on average [17], which is comparable to the 43.5% observed here.

Further, around 20% of the subjects who started completing the survey opted out before the final question. Not being able to control the condition under which the survey was completed increases the probability of subjects sabotaging the experiment by choosing randomly or not reading through the instructions thoroughly. However, including or excluding the "outliers" of the present study did not change results.

7. Concluding Remarks

Reflective individuals are more likely to act rational and self-interested in social dilemma tasks and *intuitive* individuals are more likely to bring their internalized cooperative and fair behavior to the lab. Acknowledging that individuals differ in their cognitive reflection ability entails greater prediction and description of decision making. *Intuitive* individuals are more likely to act as a strong reciprocator and do not tolerate selfish deviations for material incentives. Explaining the *intuitive* decision in the lab by the Social Heuristic Hypothesis insights are gained regarding how society maintains the cooperative and fair behavior and could shed light on cultural differences. A topic for future research is to investigate whether the *intuitive* behavior is prosocial across cultures. Future research could differentiate the perspectives further to predict decision making with greater precision and understand the behavioral differences in more detail.

Funding: This research received no external funding.

Acknowledgments: I want to thank Alexander Koch for helpful comments and feedback during the process of designing the experiments as well as writing the article.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Survey Instructions

Q1: I would really appreciate your help in collecting data for my bachelor thesis. Completing this survey will only take a few minutes and you will have the chance to earn up to DKK 400 by answering seven survey questions. I will randomly draw 7 participants, who will get paid according to their choices. This will be explained in the survey. My name is Markus Seier and I am studying Economics. Your participation is voluntary. I will analyze the data in anonymous format. The email address that you can provide at the end of the survey will only be used to contact you in case you are among the participants drawn to receive a payment. Payments will be made by mobile pay. I will delete the email address as soon as payments are completed.

Q2: First, you complete four tasks regarding "division of money". Your decisions in these tasks determine your earnings if you are randomly drawn to be paid for answering this survey. If you are drawn to be paid for a particular question, you are paid according to your choices and the choices of the other participants with whom you are randomly matched. You can be drawn to be paid for multiple questions. After completing the four above-mentioned tasks, you proceed to the second part of this survey with three short questions. Lastly, you are to indicate your gender, at which faculty you study and provide your email address if you want to have a chance of getting paid up to DKK 400. Please continue to the next page where you are to complete four different tasks regarding division of money.

Q3: You are matched with another participant of this survey. You are given DKK 100 and must decide on how much to offer the other participant. You act as the "proposer". You earn DKK 100 subtracted what you have offered and the other participant earns what you have offered him/her. How much do you give to the other participant? Remember that you and the other participant will actually be paid according to your decisions if the computer draws your names.

- DKK 0 (That is: You get DKK 100. The other gets DKK 0.)
- DKK 10 (That is: You get DKK 90. The other gets DKK 10.)
- DKK 20 (That is: You get DKK 80. The other gets DKK 20.)
- DKK 30 (That is: You get DKK 70. The other gets DKK 30.)
- DKK 40 (That is: You get DKK 60. The other gets DKK 40.)
- DKK 50 (That is: You get DKK 50. The other gets DKK 50.)
- DKK 60 (That is: You get DKK 40. The other gets DKK 60.)

- DKK 70 (That is: You get DKK 30. The other gets DKK 70.)
- DKK 80 (That is: You get DKK 20. The other gets DKK 80.)
- DKK 90 (That is: You get DKK 10. The other gets DKK 90.)
- DKK 100 (That is: You get DKK 0. The other gets DKK 100.)

Q4: You are matched with another participant of this survey. You are given DKK 100 and must decide on how much to offer the other participant. If the other participant accepts your offer, you earn DKK 100 subtracted what you have offered and the other participant earns what you have offered him/her. If the other participant rejects your offer, you both earn DKK 0. How much do your offer the other participant? Remember that you and the other participant will actually be paid according to your decisions if the computer draws your names.

- DKK 0 (That is: You get DKK 100, The other gets DKK 0 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 10 (That is: You get DKK 90, The other gets DKK 10 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 20 (That is: You get DKK 80, The other gets DKK 20 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 30 (That is: You get DKK 70, The other gets DKK 30 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 40 (That is: You get DKK 60, The other gets DKK 40 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 50 (That is: You get DKK 50, The other gets DKK 50 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 60 (That is: You get DKK 40, The other gets DKK 60 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 70 (That is: You get DKK 30, The other gets DKK 70 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 80 (That is: You get DKK 10, The other gets DKK 80 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 90 (That is: You get DKK 10, The other gets DKK 90 if the offer is accepted. Otherwise, you both get DKK 0.)
- DKK 100 (That is: You get DKK 0, The other gets DKK 100 if the offer is accepted. Otherwise, you both get DKK 0.)

Q5: You must now decide whether to accept or reject an offer from another participant. The other participant is given DKK 100 and must decide on how much to offer you. If you accept, you earn what the other participant offered you and the other participant earns DKK 100 subtracted what he/she offered you. If you reject, you both earn DKK 0. What is the minimum offer, you are willing to accept? Remember that you and the other participant will actually be paid according to your decisions if the computer draws your names.

- DKK 0
- DKK 10
- DKK 20
- DKK 30
- DKK 40
- DKK 50
- DKK 60
- DKK 70
- DKK 80

- DKK 90
- DKK 100

Q6: I will randomly draw a pair of participants, from the first question, were the participant endowed with DKK 100 (the "proposer") chose to give DKK 0 and keep the DKK 100 for him/herself. You are given DKK 50 and can reduce the earnings of the proposer who chose to keep the DKK 100 for him/herself. You can reduce the earnings of the proposer by DKK 5 by giving up DKK 1 of your own earnings. That is, if you give up DKK X of your own earnings, you reduce the earnings of the proposer by DKK 5*X. How much of your own earnings are you willing to give up to reduce the earnings of the proposer? Remember that you and the other participant will actually be paid according to your decisions if the computer draws your names.

- DKK 0 (Reduce the earnings of the proposer by DKK 0)
- DKK 1 (Reduce the earnings of the proposer by DKK 5)
- DKK 2 (Reduce the earnings of the proposer by DKK 10)
- DKK 3 (Reduce the earnings of the proposer by DKK 15)
- DKK 4 (Reduce the earnings of the proposer by DKK 20)
- DKK 5 (Reduce the earnings of the proposer by DKK 25)
- DKK 6 (Reduce the earnings of the proposer by DKK 30)
- DKK 7 (Reduce the earnings of the proposer by DKK 35)
- DKK 8 (Reduce the earnings of the proposer by DKK 40)
- DKK 9 (Reduce the earnings of the proposer by DKK 45)
- DKK 10 (Reduce the earnings of the proposer by DKK 50)
- DKK 11 (Reduce the earnings of the proposer by DKK 55)
- DKK 12 (Reduce the earnings of the proposer by DKK 60)
- DKK 13 (Reduce the earnings of the proposer by DKK 65)
- DKK 14 (Reduce the earnings of the proposer by DKK 70)
- DKK 15 (Reduce the earnings of the proposer by DKK 75)
- DKK 16 (Reduce the earnings of the proposer by DKK 80)
- DKK 17 (Reduce the earnings of the proposer by DKK 85)
- DKK 18 (Reduce the earnings of the proposer by DKK 90)
- DKK 19 (Reduce the earnings of the proposer by DKK 95)
- DKK 20 (Reduce the earnings of the proposer by DKK 100)

You have now completed the first part of the survey. The next part consists of three questions, where you are to write your answer in the box below the question. Your chances of getting paid for this part depend on how many questions you answer correctly. For each correct answer, one lottery ticket with your name will be added to the pool from which the computer will draw one participant, who will be paid DKK 100.

Q7: A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost? (Write your answer in cents) Remember, a correct answer increases your chances of getting paid DKK 100.

Q8: If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (Write your answer in minutes) Remember, a correct answer increases your chances of getting paid DKK 100.

Q9: In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (Write your answer in days) Remember, a correct answer increases your chances of getting paid DKK 100.

Q10: Please indicate your gender.

- Male
- Female

Q11: At which faculty do you study?

- Arts
- Health
- Science & Technology
- BSS

Q12: Please write your email-address (studynumber@post.au.dk) The email address is to pay a participant who is drawn to receive his/her earnings in the survey. You are not required to provide your email address, but you cannot get paid if you do not.

Thank you for participating. You will be notified by email if you are drawn to be paid.

Appendix B. Cognitive Reflection Test

1. A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost? ___ cents. Intuitive Answer: 10 / Correct Answer: 5.
2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes. Intuitive Answer: 100 / Correct Answer: 5.
3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?_____ days. Intuitive Answer: 24 / Correct Answer: 47.

Appendix C. Cognitive Reflection Test Categorization

$$Intuitive \begin{cases} = 1 & if\ Q1 = 10\ \&\ Q2 = 100\ or\ Q1 = 10\ \&\ Q3 = 24\ or\ Q2 = 100\ \&\ Q3 = 24 \\ = 0 & Otherwise \end{cases}$$

$$Reflective \begin{cases} = 1 & if\ Q1 = 5\ \&\ Q2 = 5\ or\ Q1 = 5\ \&\ Q3 = 47\ or\ Q2 = 5\ \&\ Q3 = 47 \\ = 0 & Otherwise \end{cases}$$

$$Residual \begin{cases} = 1 & if\ Intuitive = 0\ \&\ Reflective = 0 \\ = 0 & Otherwise \end{cases}.$$

Appendix D. Results Excluding the Residual Group

Appendix D.1. Rational and Self-Interested Behavior

A graphical representation of decisions consistent with rational, self-interested behavior by *intuitive (reflective)* individuals can be found in Figure A1.

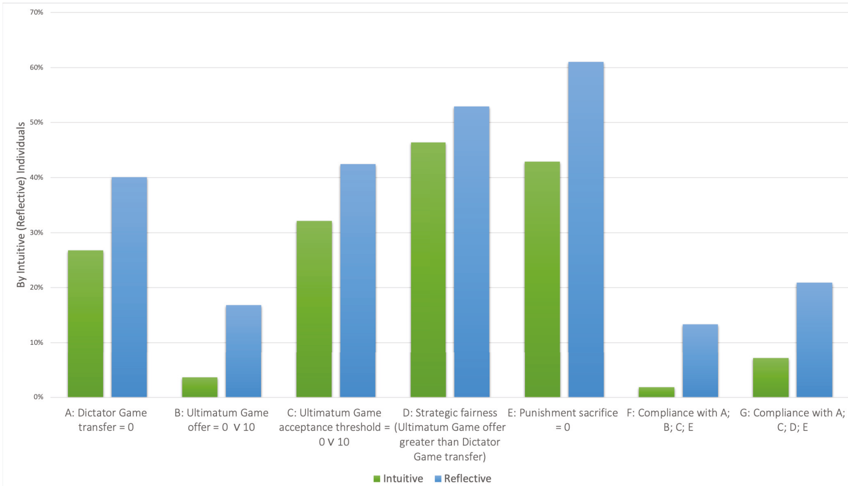


Figure A1. Frequency of Decision by Intuitive (Reflective) Individuals.

Appendix D.2. Transfer in the Dictator Game

The distribution of the *dictator game* transfer can be found in Table A1 and is illustrated in Figure A2.

Table A1. Frequency of Dictator Game Transfer by Intuitive (Reflective) Individuals.

Dictator Game Transfer	Intuitive	Reflective	Total
Transfer = 0	26.6%	40.12%	36.84%
Transfer = 10	1.8%	1.16%	1.32%
Transfer = 20	3.57%	4.65%	4.39%
Transfer = 30	3.57%	2.91%	3.07%
Transfer = 40	8.93%	5.81%	6.58%
Transfer = 50	48.21%	41.86%	43.42%
Transfer = 60	1.79%	1.16%	1.32%
Transfer = 70	0%	0%	0%
Transfer = 80	1.79%	0%	0.44%
Transfer = 90	0%	0%	0%
Transfer = 100	3.57%	2.33%	2.63%

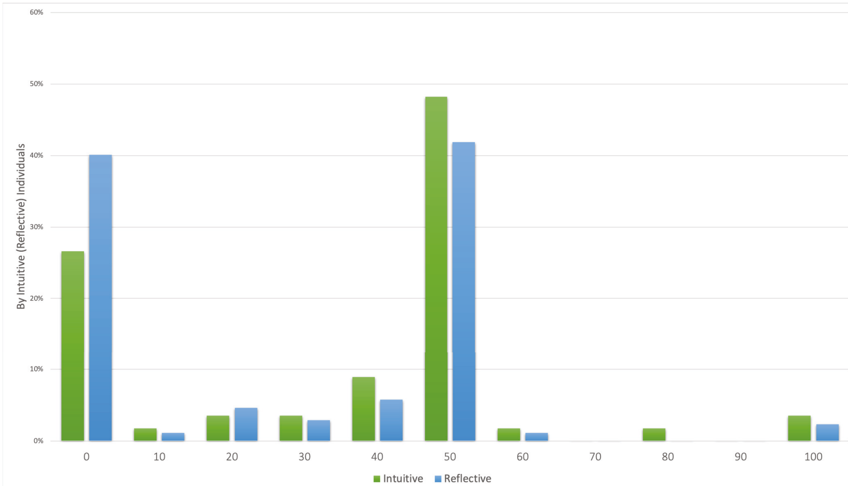


Figure A2. Frequency of Transfer by Intuitive (Reflective) Individuals.

Appendix D.3. Proposer Behavior in the Ultimatum Game

The distribution of the proposer decision in the *ultimatum game* can be found in Table A2 and is illustrated in Figure A3.

Table A2. Frequency of Ultimatum Game Offer by Intuitive (Reflective) Individuals.

Ultimatum Game Offer	Intuitive	Reflective	Total
Offer = 0	1.79%	1.16%	1.32%
Offer = 10	1.79%	15.70%	12.28%
Offer = 20	0%	4.07%	3.07%
Offer = 30	1.79%	4.65%	3.95%
Offer = 40	3.57%	7.56%	6.58%
Offer = 50	82.14%	63.37%	67.98%
Offer = 60	3.57%	2.33%	2.63%
Offer = 70	0%	0%	0%
Offer = 80	0%	0.58%	0.44%
Offer = 90	0%	0%	0%
Offer = 100	5.36%	0.58%	1.75%

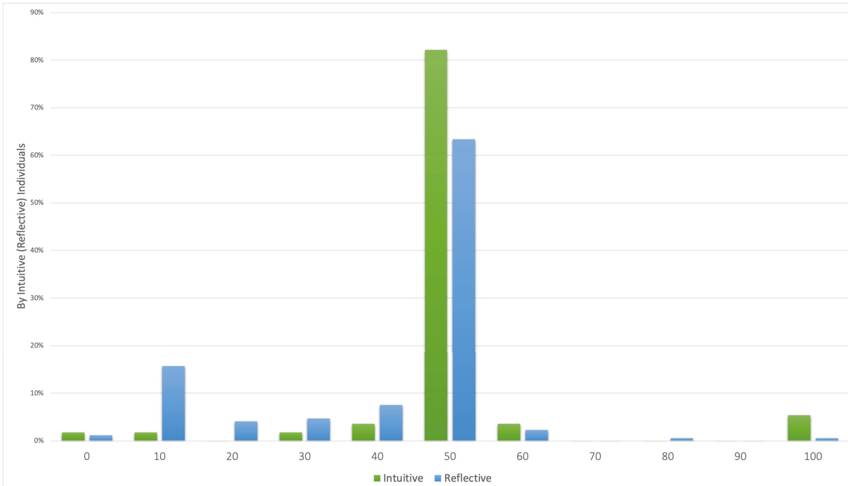


Figure A3. Frequency of Ultimatum Game Offer by Intuitive (Reflective) Individuals.

Appendix D.4. Recipient Behavior in the Ultimatum Game

The distribution of the recipient decision in the *ultimatum game* can be found in Table A3 and is illustrated in Figure A4.

Table A3. Frequency of Ultimatum Game Acceptance Threshold by Intuitive (Reflective) Individuals.

Ultimatum Game Acceptance Threshold	Intuitive	Reflective	Total
Threshold = 0	10.71%	5.81%	7.02%
Threshold = 10	21.42%	36.63%	32.89%
Threshold = 20	1.79%	5.23%	4.39%
Threshold = 30	10.71%	9.30%	9.65%
Threshold = 40	16.07%	16.86%	16.67%
Threshold = 50	37.50%	25.00%	28.07%
Threshold = 60	0%	0%	0%
Threshold = 70	0%	0%	0%
Threshold = 80	0%	0%	0%
Threshold = 90	0%	1.16%	0.88%
Threshold = 100	1.79%	0%	0.44%

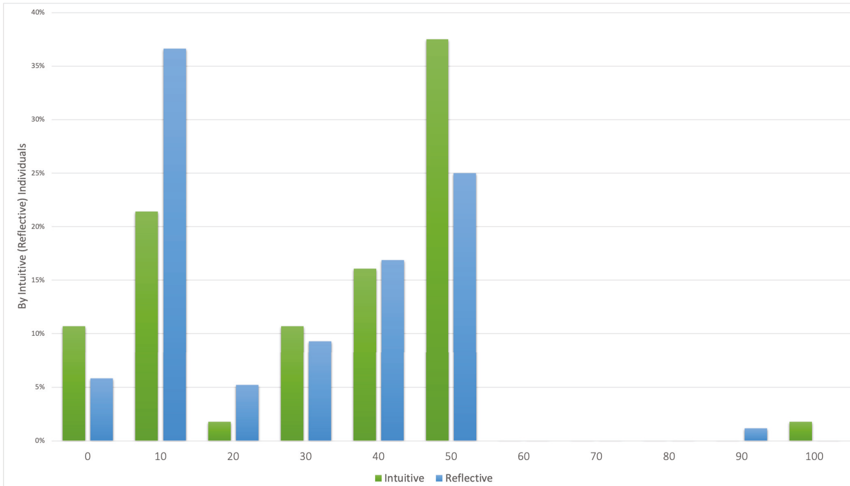


Figure A4. Frequency of Ultimatum Game Acceptance Threshold by Intuitive (Reflective) Individuals.

Appendix D.5. Third-Party Punishment Behavior

The distribution of the *third-party punishment* decision can be found in Table A4 and is illustrated in Figure A5.

Table A4. Frequency of Punishment Sacrifice by Intuitive (Reflective) Individuals.

Punishment Sacrifice	Intuitive	Reflective	Total
Sacrifice = 0	42.86%	61.05%	56.58%
Sacrifice = 1	0%	1.74%	1.32%
Sacrifice = 2	1.79%	0.58%	0.88%
Sacrifice = 3	1.79%	1.16%	1.32%
Sacrifice = 4	3.57%	1.16%	1.75%
Sacrifice = 5	8.93%	6.40%	7.02%
Sacrifice = 6	0%	1.16%	0.88%
Sacrifice = 7	0%	0.58%	0.44%
Sacrifice = 8	0%	1.16%	0.88%
Sacrifice = 9	0%	0%	0%
Sacrifice = 10	19.64%	13.95%	15.35%
Sacrifice = 11	0%	0.58%	0.44%
Sacrifice = 12	1.79%	1.16%	1.32%
Sacrifice = 13	0%	1.16%	0.88%
Sacrifice = 14	0%	0%	0%
Sacrifice = 15	1.79%	0.58%	0.88%
Sacrifice = 16	0%	0%	0%
Sacrifice = 17	0%	0.58%	0.44%
Sacrifice = 18	0%	0%	0%
Sacrifice = 19	0%	0%	0%
Sacrifice = 20	17.86%	6.98%	9.65%

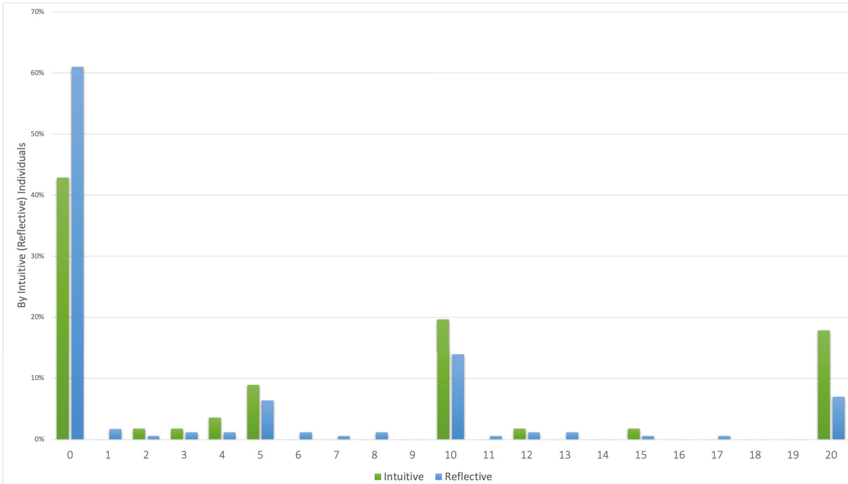


Figure A5. Frequency of Punishments Sacrifice by Intuitive (Reflective) Individuals.

Appendix D.6. Behavioral Differences between Intuitive and Reflective Individuals (Excluding Residual Group)

In Table A5 an overview of the results when excluding the residual group can be found. This include means of the different tasks as well as *p*-values from the statistical tests. A table with the marginal effects from logistic regressions can be found in Table A6. Excluding the residual group from the analyses and comparing those categorized as *reflective* only with those categorized as *intuitive* does not change much in the conclusions. Most notable differences are in terms of statistical significant in the MWU distribution tests and the contingency-table χ^2 tests where the *p*-values are greater for almost all of the tasks. The logistic regressions excluding the residual group reveal a very similar pattern in terms of statistical significant and interpretation of marginal effects.

Table A5. Results Excluding the Residual Group by Intuitive (Reflective) Individuals.

	Intuitive	**Reflective**	**Combined**	**MWU or χ^2 (*p*-Value)**
Dictator Game Transfer (mean)	35.7	28.2	30	0.074
A: Dictator Game Transfer = 0 (freq.)	26.8%	40.1%	36.8%	0.072
Ultimatum Game Offer (mean)	50.78	40.9	43.3	0.000
B: Ultimatum Game Offer = $0 \vee 10$ (freq.)	3.57%	16.9%	13.7%	0.012
Ultimatum Game Acceptance Threshold (mean)	32.7	27.8	29	0.127
C: Ultimatum Game Acceptance Threshold = $0 \vee 10$ (freq.)	32.1%	42.4%	39.9%	0.172
D: Strategic Fairness (Dictator Game Transfer > Ultimatum Game Offer) (freq.)	46.4%	52.9%	51.3%	0.400
Punishment Sacrifice (mean)	6.7	4	4.6	0.010
E: Punishment Sacrifice = 0 (freq.)	42.9%	61.1%	56.6%	0.017
Compliance with A, B, C & E	1.8%	13.4%	10.5%	0.014
Compliance with A, C, D & E	7.14%	20.9%	17.5%	0.018

Table A6. Marginal effects from Logistic regressions.

VARIABLES	(1) A	(2) B	(3) C	(4) D	(5) E	(6) F	(7) G
1.Reflective	0.081	0.123 ***	0.090	0.014	0.177 **	0.107 ***	0.114 **
	(0.072)	(0.040)	(0.075)	(0.076)	(0.078)	(0.033)	(0.051)
1.Male	0.242 ***	0.094 **	0.076	0.222 ***	0.036	0.092 **	0.169 ***
	(0.063)	(0.043)	(0.066)	(0.066)	(0.067)	(0.039)	(0.049)
Observations	227	227	227	227	227	227	227

Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. A: Dictator Game transfer = 0, B: Ultimatum Game offer = 0 or = 10, C: Ultimatum Game acceptance threshold = 0 or = 10, D: Strategic fairness; Ultimatum Game offer greater than Dictator Game transfer, E: Punishment sacrifice = 0, F: Compliance with A; B; C; E, G: Compliance with A; C; D; E.

Appendix E. Cognitive Reflection Test Results

The distribution of answers on the CRT for both men and women can be found in Table A7, for men alone in Table A8 and for women in Table A9.

Table A7. Distribution of Answers on the CRT for Both Men and Women.

Question/Answer	Correct	Intuitive	Other
1: Bat and Ball	58%	39%	3%
2: Widget	69%	25%	6%
3: Lily Pads	82%	15%	3%

Table A8. Distribution of Answers on the CRT for Men Alone.

Question/Answer	Correct	Intuitive	Other
1: Bat and Ball	65%	31%	3%
2: Widget	77%	20%	3%
3: Lily Pads	89%	10%	1%

Table A9. Distribution of Answers on the CRT for Women Alone.

Question/Answer	Correct	Intuitive	Other
1: Bat and Ball	50%	48%	2%
2: Widget	60%	30%	10%
3: Lily Pads	74%	21%	5%

The distribution of the number of correct answers on the CRT by gender and in total can be found in Table A10 and the distribution of the number of intuitive, wrong answers can be found in Table A11.

Table A10. Distribution of Number of Correct Answers on the CRT by Gender.

Gender/Number of Correct Answers	0 Correct Answers	1 Correct Answer	2 Correct Answers	3 Correct Answers
Men	5%	13%	27%	55%
Women	21%	16%	22%	41%
Men & Women	13%	14%	24%	49%

Table A11. Distribution of Number of Intuitive Answers on the CRT by Gender.

Gender/Number of Intuitive Answers	0 Intuitive Answers	1 Intuitive Answer	2 Intuitive Answers	3 Intuitive Answers
Men	60%	23%	12%	5%
Women	47%	22%	18%	13%
Men & Women	54%	22%	15%	9%

Appendix F. Additional Tables and Histograms of Choices by Non-Reflective (Reflective) Individuals

Appendix F.1. Transfer in the Dictator Game

The distribution of the *dictator game* transfer can be found in Table A12 and is illustrated in Figure A6.

Table A12. Frequency of Dictator Game Transfer by Non-Reflective (Reflective) Individuals.

Dictator Game Transfer	Non-Reflective	Reflective	Total
Transfer = 0	25%	40.12%	36.02%
Transfer = 10	1.56%	1.16%	1.27%
Transfer = 20	3.12%	4.65%	4.24%
Transfer = 30	4.69%	2.91%	3.39%
Transfer = 40	7.81%	5.81%	6.36%
Transfer = 50	51.56%	41.86%	44.49%
Transfer = 60	1.56%	1.16%	1.27%
Transfer = 70	0%	0%	0%
Transfer = 80	1.56%	0%	0.42%
Transfer = 90	0%	0%	0%
Transfer = 100	3.12%	2.33%	2.54%

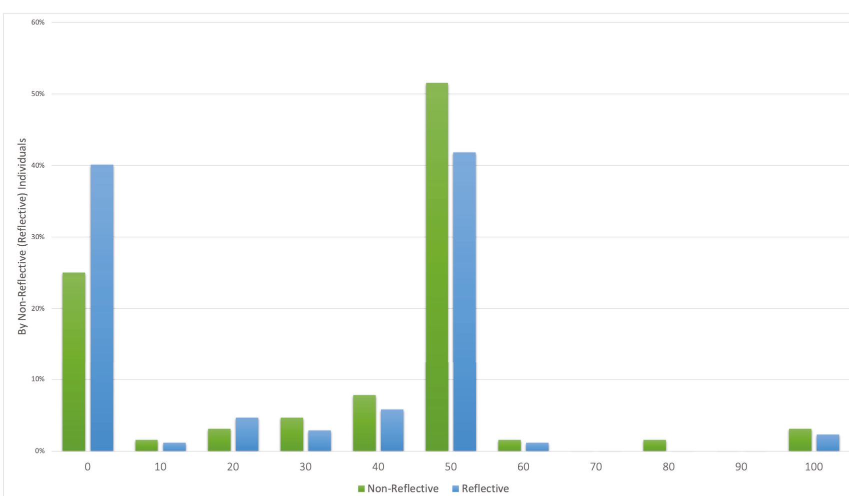


Figure A6. Frequency of Transfer by Non-Reflective (Reflective) Individuals.

Appendix F.2. Proposer Behavior in the Ultimatum Game

The distribution of the proposer decision in the *ultimatum game* can be found in Table A13 and is illustrated in Figure A7.

Table A13. Frequency of Ultimatum Game Offer by Non-Reflective (Reflective) Individuals.

Ultimatum Game Offer	Non-Reflective	Reflective	Total
Offer = 0	1.56%	1.16%	1.27%
Offer = 10	1.56%	15.70%	11.86%
Offer = 20	0%	4.07%	2.97%
Offer = 30	1.56%	4.65%	3.81%
Offer = 40	4.69%	7.56%	6.78%
Offer = 50	82.81%	63.37%	68.64%
Offer = 60	3.12%	2.33%	2.54%
Offer = 70	0%	0%	0%
Offer = 80	0%	0.58%	0.42%
Offer = 90	0%	0%	0%
Offer = 100	4.69%	0.58%	1.69%

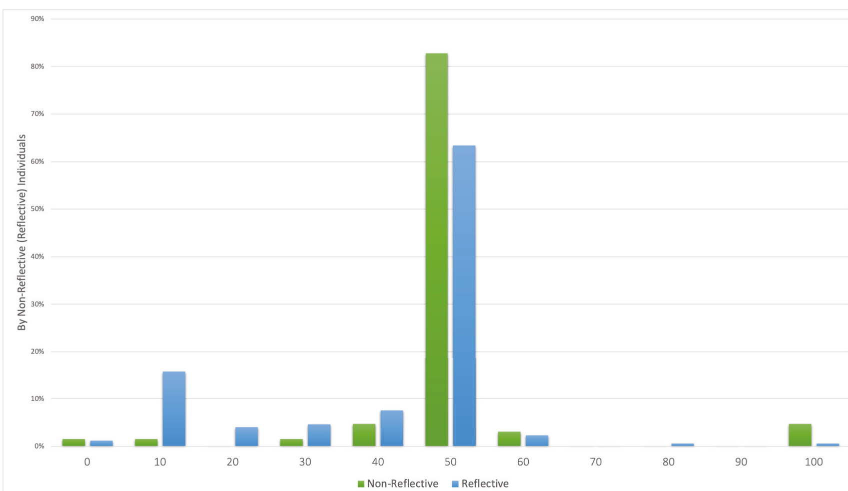


Figure A7. Frequency of Ultimatum Game Offer by Non-Reflective (Reflective) Individuals.

Appendix F.3. Recipient Behavior in the Ultimatum Game

The distribution of the recipient decision in the *ultimatum game* can be found in Table A14 and is illustrated in Figure A8.

Table A14. Frequency of Ultimatum Game Acceptance Threshold by Non-Reflective (Reflective) Individuals.

Ultimatum Game Acceptance Threshold	Non-Reflective	Reflective	Total
Threshold = 0	9.38%	5.81%	6.78%
Threshold= 10	20.31%	36.63%	32.20%
Threshold= 20	1.56%	5.23%	4.24%
Threshold= 30	10.94%	9.30%	9.75%
Threshold= 40	14.06%	16.86%	16.10%
Threshold= 50	42.19%	25.00%	29.66%
Threshold= 60	0%	0%	0%
Threshold= 70	0%	0%	0%
Threshold= 80	0%	0%	0%
Threshold= 90	0%	1.16%	0.85%
Threshold= 100	1.56%	0%	0.42%

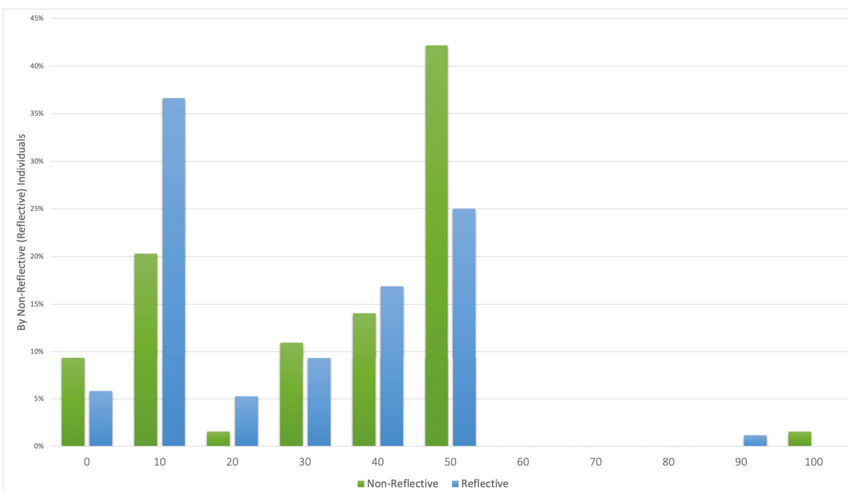


Figure A8. Frequency of Ultimatum Game Acceptance Threshold by Non-Reflective (Reflective) Individuals.

Appendix F.4. Third-Party Punishment Behavior

The distribution of the *third-party punishment* decision can be found in Table A15 and is illustrated in Figure A9.

Table A15. Frequency of Punishment Sacrifice by Non-Reflective (Reflective) Individuals.

Punishment Sacrifice	Non-Reflective	Reflective	Total
Sacrifice = 0	40.62%	61.05%	55.51%
Sacrifice = 1	0%	1.74%	1.27%
Sacrifice = 2	1.56%	0.58%	0.85%
Sacrifice = 3	1.56%	1.16%	1.27%
Sacrifice = 4	4.69%	1.16%	2.12%
Sacrifice = 5	7.81%	6.40%	6.78%
Sacrifice = 6	0%	1.16%	0.85%
Sacrifice = 7	0%	0.58%	0.42%
Sacrifice = 8	0%	1.16%	0.85%
Sacrifice = 9	0%	0%	0%
Sacrifice = 10	18.75%	13.95%	15.25%
Sacrifice = 11	1.56%	0.58%	0.85%
Sacrifice = 12	3.12%	1.16%	1.69%
Sacrifice = 13	0%	1.16%	0.85%
Sacrifice = 14	0%	0%	0%
Sacrifice = 15	1.56%	0.58%	0.85%
Sacrifice = 16	0%	0%	0%
Sacrifice = 17	0%	0.58%	0.42%
Sacrifice = 18	0%	0%	0%
Sacrifice = 19	0%	0%	0%
Sacrifice = 20	18.75%	6.98%	10.17%

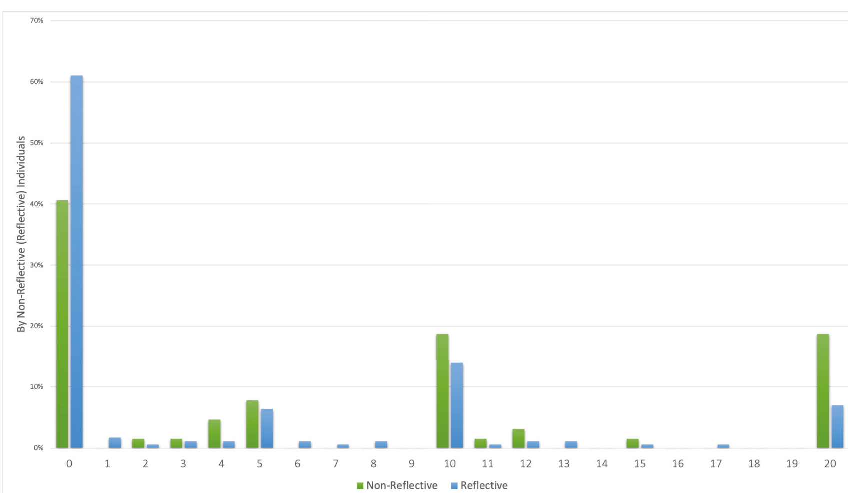


Figure A9. Frequency of Punishment Sacrifice by Non-Reflective (Reflective) Individuals

Appendix F.5. Behavioral Differences between Non-Reflective and Reflective Individuals

In Table A16 an overview of the results comparing *non-reflective* and *reflective* individuals can be found.

Table A16. Results by Non-Reflective (Reflective) Individuals.

	Non-Reflective	Reflective	Combined	MWU or χ^2 (*p*-Value)
Dictator Game Transfer (mean)	36.4	28.2	30.4	0.033
A: Dictator Game Transfer = 0 (freq.)	25%	40.1%	36%	0.032
Ultimatum Game Offer (mean)	50.5	40.9	43.5	0.000
B: Ultimatum Game Offer = 0 ∨ 10 (freq.)	3.1%	16.9%	13.1%	0.005
Ultimatum Game Acceptance Threshold (mean)	33.9	27.8	29.4	0.032
C: Ultimatum Game Acceptance Threshold = 0 ∨ 10 (freq.)	29.7%	42.4%	39%	0.074
D: Strategic Fairness (Dictator Game Transfer > Ultimatum Game Offer) (freq.)	43.8%	52.9%	50.4%	0.211
Punishment Sacrifice (mean)	7.1	4	4.8	0.002
E: Punishment Sacrifice = 0 (freq.)	40.6%	61.1%	55.5%	0.005
Compliance with A, B, C & E	1.6%	13.4%	10.2%	0.008
Compliance with A, C, D & E	6.3%	20.9%	17%	0.008

References

1. Camerer, C.F.; Fehr, E. Measuring social norms and preferences using experimental games: A guide for social scientists. *Found. Hum. Soc.* **2004**, *97*, 55–95.
2. Fehr, E.; Fischbacher, U.; Gächter, S. Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum. Nat.* **2002**, *13*, 1–25. [CrossRef] [PubMed]
3. Fehr, E.; Gächter, S. Reciprocity and economics: The economic implications of homo reciprocans. *Eur. Econ. Rev.* **1998**, *42*, 845–859. [CrossRef]
4. Fehr, E.; Fischbacher, U. Third-party punishment and social norms. *Evol. Hum. Behav.* **2004**, *25*, 63–87. [CrossRef]
5. Frederick, S. Cognitive reflection and decision making. *J. Econ. Perspect.* **2005**, *19*, 25–42. [CrossRef]
6. Toplak, M.E.; West, R.F.; Stanovich, K.E. Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Think. Reason.* **2014**, *20*, 147–168. [CrossRef]
7. Ponti, G.; Rodriguez-Lara, I. Social preferences and cognitive reflection: Evidence from a dictator game experiment. *Front. Behav. Neurosci.* **2015**, *9*, 146. [CrossRef]
8. De Neys, W.; Novitskiy, N.; Geeraerts, L.; Ramautar, J.; Wagemans, J. Cognitive control and individual differences in economic ultimatum decision-making. *PLoS ONE* **2011**, *6*, e27107. [CrossRef]
9. Calvillo, D.P.; Burgeno, J.N. Cognitive reflection predicts the acceptance of unfair ultimatum game offers. *Judgm. Decis. Mak.* **2015**, *10*, 332–342.
10. Inaba, M.; Inoue, Y.; Akutsu, S.; Takahashi, N.; Yamagishi, T. Preference and strategy in proposer's prosocial giving in the ultimatum game. *PLoS ONE* **2018**, *13*, e0193877. [CrossRef]
11. Nielsen, U.H.; Tyran, J.R.; Wengström, E. Second thoughts on free riding. *Econ. Lett.* **2014**, *122*, 136–139. [CrossRef]
12. Cappelen, A.W.; Nielsen, U.H.; Tungodden, B.; Tyran, J.R.; Wengström, E. Fairness is intuitive. *Exp. Econ.* **2016**, *19*, 727–740. [CrossRef]
13. Piovesan, M.; Wengström, E. Fast or fair? A study of response times. *Econ. Lett.* **2009**, *105*, 193–196. [CrossRef]
14. Krajbich, I.; Bartling, B.; Hare, T.; Fehr, E. Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nat. Commun.* **2015**, *6*, 1–9. [CrossRef]

15. Rand, D.G.; Peysakhovich, A.; Kraft-Todd, G.T.; Newman, G.E.; Wurzbacher, O.; Nowak, M.A.; Greene, J.D. Social heuristics shape intuitive cooperation. *Nat. Commun.* **2014**, *5*, 3677. [CrossRef]
16. Selten, R. Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperimentes. In *Beiträge zur Experimentellen Wirtschafts-Forschung*; JCB Mohr: Tübingen, Germany, 1967; pp. 136–168.
17. Oosterbeek, H.; Sloof, R.; Van De Kuilen, G. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Exp. Econ.* **2004**, *7*, 171–188. [CrossRef]
18. Engel, C. Dictator games: A meta study. *Exp. Econ.* **2011**, *14*, 583–610. [CrossRef]
19. Stanovich, K.E.; West, R.F. Individual differences in reasoning: Implications for the rationality debate? *Behav. Brain Sci.* **2000**, *23*, 645–665. [CrossRef]
20. Kahneman, D.; Egan, P. *Thinking, Fast and Slow*; Farrar, Straus and Giroux: New York, NY, USA, 2011; Volume 1.
21. Cueva, C.; Iturbe-Ormaetxe, I.; Mata-Pérez, E.; Ponti, G.; Sartarelli, M.; Yu, H.; Zhukova, V. Cognitive (ir) reflection: New experimental evidence. *J. Behav. Exp. Econ.* **2016**, *64*, 81–93. [CrossRef]
22. Thaler, R.H. Anomalies: The ultimatum game. *J. Econ. Perspect.* **1988**, *2*, 195–206. [CrossRef]
23. Blanco, M.; Engelmann, D.; Koch, A.K.; Normann, H.T. Preferences and beliefs in a sequential social dilemma: A within-subjects analysis. *Games Econ. Behav.* **2014**, *87*, 122–135. [CrossRef]
24. Markowska-Przybyła, U.; Ramsey, D.M. Norms of Egality and the Gap between Attitudes and Behaviour amongst Polish Students: An Experimental Study. *Econ. Sociol.* **2017**, *10*, 220–236. [CrossRef] [PubMed]

games

MDPI

Article

Cost of Reasoning and Strategic Sophistication

Wanqun Zhao

Rady School of Management, University of California, San Diego, CA 92093, USA; wanqun.zhao@rady.ucsd.edu

Received: 30 June 2020; Accepted: 14 September 2020; Published: 22 September 2020

Abstract: I designed an experiment to study the persistence of the prevailing levels of reasoning across games. Instead of directly comparing the k-level(s) of reasoning for each game, I used cognitive load to manipulate the strategic environment by imposing variations on the subject's cost of reasoning and their first- and second-order beliefs. Subjects have systematic changes in k-level(s) of reasoning across games. That finding suggests that subjects are responsive to changes in the strategic environment. Changes in k-level(s) of reasoning are mostly consistent with the endogenous depth of reasoning model when subjects are more cognitively capable or facing less cognitively capable opponents. Subjects have cognitive bounds, but often choose a lower-type action due to their beliefs about their opponents. Finally, cognitive ability plays a significant role in subjects making strategic adjustments when facing different strategic environments.

Keywords: level-k reasoning; guessing game; cognitive load; endogenous depth of reasoning; strategic thinking

1. Introduction

The use of the level-k model has prevailed in the literature for characterizing people's initial responses in laboratory strategic games [1,2]. The model characterizes the player's systematic deviations from the Nash equilibrium using a bounded rational-type explanation. The level-0 type's action is assumed to be uniformly distributed over all actions (or in some cases, level-0 type's action is the most prominent action available), whereas the level-1 type has the best response to the expected action of the level-0 type. The level-2 type has the best response to the expected action of the level-1 type. The iterations follow this pattern, as the level-k type always has the best response to the actions of level-$k-1$ type. Such patterns of off-equilibrium play have been evidenced in many laboratory experiments. In Nagel's p-beauty contest game, Nagel found spikes that correspond to the first and second rounds of iterative best responses [1]. Stahl and Wilson found similar evidence of level-1 and level-2 types with 10 matrix games [2]. Camerer et al. developed a cognitive hierarchy model [3]. Instead of holding a belief that all the other players are type k-1, level-k players in the cognitive hierarchy model assign a probability distribution over all the lower types. Many other studies used the level-k model to explain laboratory data (matrix game [4]; beauty contest game [5–8]; sequential game [9]; auction [10,11]; Crawford, Costa-Gomes and Iriberri also provide a comprehensive literature review [12]).

However, although the level-k model has proven its usefulness in characterizing initial responses for many laboratory games, its predictive power remains ambiguous because (1) it is often used posteriorly to classify a player's type given their actions and (2) the model lacks components related to individual characteristics that could help identify different types of players. It is important to understand how certain levels are reached for each individual, as it is a starting point for the discussion of the model's predictive power. Alaoui and Penta developed a framework called the endogenous depth of reasoning (EDR) model to explain what may happen in a player's head when they encounter a given strategic situation [13]. The EDR model captures individual characteristics by introducing cost of reasoning, which is determined both by the strategic environment and by a player's endogenous cognitive ability. The model includes game-specific characteristics by introducing the benefit of reasoning through payoffs of the games. Lastly, the model allows a clear separation of cognitive bounds and behavioral levels observed in games by introducing higher-order beliefs. Such separation makes room for individual adjustments of k-levels in different strategic environments. As a result, a level-1 action observed from a game does not necessarily classify the player as a level-1 player. Instead, such action can be a product of the player's cost and benefit analysis and his belief about his opponents.

The EDR model provides a plausible starting point to study the persistence of the level-k model. However, as individuals have heterogeneous costs of reasoning and belief systems in all kinds of strategic situations, it is hard to conduct direct comparisons across games to test whether the behavioral k-levels follow the EDR model's predictions. In this paper, I use Costa-Gomes and Crawford's two-person guessing games (henceforth CGC06) and cognitive load to create different strategic environments [14]. By controlling cognitive load, I create a standard for the cost of reasoning for all the subjects. Although individual cognitive ability may still have an effect, by using a within-subject experimental design, the individual effect will no longer impact the comparisons of strategic levels across games for the same subject. The revelation of information about the strategic environment is also carefully manipulated to clearly control the subject's belief space. The goal was to test whether the EDR model provides directional predictions about the changes on k-levels across games for any given subject. Alaoui and Penta tested the benefit part of their model using the 11–20 money request game with altered bonus rewards [13,15]. To the best of my knowledge, this was the first paper to provide experimental tests of the EDR model by introducing different strategic environments with controlled cost and belief space.

With the 18 two-person guessing games in the experiment, the results suggest that the subject's behavioral levels systematically vary across the games. Subjects are mostly responsive to the changes in the strategic environment. Their directional changes in behavioral levels can be predicted by the EDR model when they are more cognitively capable or their opponent is less cognitively capable. An inherent cognitive bound exists for the subjects in different strategic environments. When comparing a subject's behavioral levels across all the games while providing the same amount of cognitive resources, their behavioral levels rarely exceed their cognitive bound level for that strategic environment.

A few other papers also studied the correlation of individual k-levels with cognitive ability. Allred et al. investigated the effects of cognitive load on strategic sophistication [16]. In their experiments, they asked the subjects to perform a memorization task of either a three- or nine-digit binary number concurrently with strategic games such as beauty contest, 11–20, and 10 matrix games. They found that subjects with high loads (i.e., nine-digit number) were less capable of computing best responses, especially for the beauty contest game. They were also aware of their strategic disadvantages. The net result of cognitive load depended on the specific strategic context. Burnham et al. used a standard psychometric test to measure the cognitive abilities of their subjects, and correlated the test results with subjects' performances in a p-beauty contest game [17]. They found a negative correlation between cognitive test scores and entries in the beauty contest game, indicating that subjects with higher cognitive ability tend to be more strategically sophisticated in such games. Gill and Prowse used a 60-question non-verbal Raven test to assign subjects into high- and low-cognitive-ability groups [18]. They asked the subjects to play a p beauty contest game for 10 rounds, and found that subjects in the high-cognitive-ability group converged to equilibrium faster. These studies provided some evidence of the correlation of individual k-levels with cognitive ability or carefully controlled cognitive tasks. In my experiment, I used memorization tasks to manipulate the cost of reasoning for the subjects in the context of a two-person guessing game. According to Allred et al., higher cognitive load negatively affects a subject's ability to calculate the best responses in this type of guessing games [16]. To attain a higher level of strategic sophistication, players have to exert more effort to combat the effects of cognitive load; therefore, the cost of reasoning increases with cognitive load in this strategic situation. Every subject experienced both the low and high cognitive loads at some point during the experiment, so they were fully aware of the additional cost of reasoning that was added by these memorization tasks. As a result, their cost of reasoning and their belief about their opponent's cost of reasoning can be quantified by the cognitive load.

The stability of k-levels is an important aspect in the level-k model literature. Stahl and Wilson used twelve normal-form games to estimate the player's level [19]. They found that using a relatively low threshold, 35 out of 48 subjects could be classified as stable across games. Fragiadakis et al. asked the subjects to repeat their decisions in a series of two-person guessing games to subsequently best respond to their past actions [20]. They found that only 40% of the subjects who were able to replicate the decisions could be classified as a known behavioral type. A few works mentioned the predictive power of strategic sophistication. Arad and Rubinstein used a multidimensional Colonel Blotto game to observe subject's multidimensional iterative reasoning process [21]. They found that subjects with a higher level of reasoning in the 11–20 money request game also seem to have more rounds of iterative reasoning in this game.

Perhaps the most closely related work to this paper is Georganas, Healy, and Weber's 2015 paper [22]. They conducted an experiment to examine the cross-game stability of the k-levels. They used four matrix undercutting games and six two-person guessing games and compared them at the individual level. They found no correlation between the levels of reasoning across games. However, they found some evidence of cross-game stability within the class of undercutting game. I studied a similar question to the cross-game stability of the level-k model. Instead of introducing a second family of games, I used cognitive load to mimic different strategic environments, and restricted the subjects to fixed pairs while playing the games. The belief space was therefore carefully controlled, and the uncertainty from playing against a new random player for each round was completely eliminated. The data suggested that systematic level changes can be predicted by the EDR model under certain conditions. In Section 2, I provide a brief introduction to the EDR model to cover some necessary background and theoretical predictions. In Section 3, the experimental design is introduced in detail. Sections 4 and 5 cover the data analysis procedure and the discussion of the results, respectively. Section 6 provides the concluding remarks.

2. Theoretical Consideration

2.1. Model

I adopted Alaoui and Penta's EDR model for theoretical predictions [13]. In this model, players follow an endogenous reasoning process that determines the strategic bound in a particular context. With added structure on beliefs, the model is able to predict a player's actual level of play in any game that could use a k-level iterative best response reasoning process. The main benefit of using this model is that the structure of the model allowed me to conduct a comparative statics exercise on a player's reasoning process. One of the main goals of this study was to conduct a comparative static exercise on the cost side. Below, I provide more detailed descriptions of some key features of this model. These features are relevant to the experimental design and predictions for this paper.

A player's cognitive bound is a mapping from the incremental cost of reasoning ($c(k)$) and the incremental value of reasoning ($v(k)$) at each level to the intersection of the two terms.

$$\kappa(v,c) = min\{k \in \mathbb{N} \mid v(k) \geq c(k) \; and \; v(k+1) < c(k+1)\} \tag{1}$$

A player reaches their cognitive bound at the kth level by having a value of reasoning for that level exceeds cost of reasoning, but their cost–benefit analysis no longer supports the one-higher level (i.e., $k+1$) of reasoning. Further denote the cognitive bound of player i as $\bar{k_i}$, where:

$$\bar{k_i} = \kappa(v_i, c_i). \tag{2}$$

According to Alaoui and Penta, the value of reasoning is affected by the payoff of the game [13]. The cost of reasoning is an endogenous characteristic of an individual, which is largely related to their cognitive or reasoning ability. In this paper, I take their assumption on the value of reasoning and continue to assume that the payoff is the only incentive for players to apply logical reasoning in the games. I provide a further discussion on the cost of reasoning. Beyond an individual's endogenous ability, the strategic environment (such as cognitive load) provides many challenges for a person in applying strategic reasoning, which alters the cost of reasoning.

A player's belief is represented as a tuple. Since the game in my design is symmetric in payoffs, a player's belief can be restricted to the beliefs about the cost of reasoning. Therefore, the first element of the tuple, c_i, represents player i's own cost of reasoning. The second element is player i's beliefs of his opponent's (player j) cost of reasoning, denoted as c_j^i. The last element c_i^{ij} is player i's second-order belief, which is their belief about player j's belief of themselves. Any higher-order beliefs could be nested to the first- and second-order beliefs; therefore, a player's belief is represented as:

$$t_i = (c_i, c_j^i, c_i^{ij}). \tag{3}$$

2.2. Theoretical Predictions

I formulated the testable predictions following the EDR construction discussed in Section 2.1. For any game $G = \{X_i, u_i\}_{i=1,2}$, let $k_i(x_i)$ be the reflected behavioral level of player i by choosing action x_i, where X_i is the set of actions available for player i and u_i is the payoff function for player i.

1. **Changing the cost of reasoning**: For any c_j^i and c_i^{ij}, $k_i(x_i)$ (weakly) decreases with c_i. Fixing player i's first- and second-order beliefs, their cognitive bound weakly decreases with the cost of reasoning. The observed level of player i from the game will also weakly decrease. In my design, for the first 16 games holding cognitive load and information structure constant for the opponent, players will display lower strategic levels when the memorization task is a string of seven letters.
2. **Changing the opponent's cost of reasoning**: For any c_i and c_i^{ij}, $k_i(x_i)$ (weakly) decreases with c_j^i. If $c_i^{ij} = c_i$, then player i's cognitive bound is binding if they regard their opponent as more sophisticated.
   Player i reacts to the change in the cost of reasoning of their opponents. More specifically, if he observes his opponent's cost of reasoning increasing, he will adjust their strategy in the game to best respond to his opponent. That means they may choose to take an action that corresponds to a lower level of strategic sophistication. However, such adjustments of strategies are binding by the cognitive bound when the player believes their opponent has a lower cost of reasoning compared to their own cost. In the context of my experiment, a player should choose a weakly lower level of strategy if he observes his opponent's memorization task becoming more difficult (i.e., from a string of three letters to a string of seven letters).
3. **Changing the second-order belief**: For any c_i and c_j^i, $k_i(x_i)$ (weakly) decreases with c_i^{ij}. If $c_i \geq c_i^{ij}$, then player i's cognitive bound is binding. By fixing player i's own cost of reasoning and his opponent's cost, through only changing player i's second-order belief, player i should adjust their strategic actions. For example, when a player has a low cost of reasoning in the game, if they believe that their opponent has a wrong belief about themselves, namely, they believe that their opponent thinks the cost of reasoning for them is very high, then they can switch to an action that is associated with a lower level of reasoning. However, this adjustment of strategic actions according to the second-order belief is restricted by player i's own cognitive bound, meaning that they cannot make any adjustments that requires a higher level of reasoning than their cognitive bound. In the context of my experimental design, players should adjust their actions when the information structure shifts from full revelation of cognitive load to partial revelation.

4. **Cognitive bound**: Given c_i, for any c_j^i and c_i^{ij}, $k_i(x_i)$ never exceeds $\bar{k}_i$. When fixing player i's own cost of reasoning, their behavioral level should never exceed their cognitive bound. In the context of this experiment, on an individual level, actions observed in games 17 and 18 should correspond to the highest level of reasoning that one player can achieve under the respective cognitive load.

3. Experimental Design

In this section, I present the details of the experimental design. The experiment captured the process of level-k thinking through the two-person guessing game [14]. I provide a brief introduction to the game first, followed by the treatment design and the experimental timeline.

3.1. The Game

The two-person guessing game is an asymmetric, two-player game. Each player has a lower limit, $a_i > 0$, an upper limit, $b_i > 0$, and a target $p_i \in (0, 2)$. Players are required to input a guess that is within their lower and upper limit. However, their actual choice is not restricted by the limit. Denote player i's input by x_i. If a player guesses a number x_i that falls outside the limit interval, then their guess will be adjusted to the closest bound. For example, if $x_i < a_i$, then the adjusted guess y_i will be $y_i = a_i$. If $x_i > b_i$, then the adjusted guess y_i is $y_i = b_i$. However, any guess falling within the limit interval will not be adjusted; i.e., $y_i = x_i$.

The goal of the game is to make a guess that minimizes the difference between the player's own guess and the product of their target and his opponent's guess. Denote the difference by $e_i = \mid y_i - p_i \cdot y_j \mid$. The payoff is a quasi-concave function minimized at zero. Player i receives $u_i = \max\{0, 200 - e_i\} + \max\{0, 100 - \frac{e_i}{100}\}$. Since a player's guesses that have the same adjusted inputs will yield the same outcome for the subject, I use the adjusted guess y_i as a proxy of how players perform in the game.

In this game, the level-0 player is assumed to play randomly according to a uniform distribution over the action space. Denote the theoretical predicted guess made by a k-level player as x_i^k. Given the assumption imposed on the level-0 player's strategy, level-1 players will best respond to the expected value of level-0 player's guess, i.e., $x_i^1 = p_i \cdot \mathbb{E}\{y_j \mid y_j \in [a_j, b_j]\}$. The level-2 player's strategy will then be $x_i^2 = p_i \cdot \{\mathbb{1}(x_j^1 \in [a_j, b_j]) \cdot x_j^1 + \mathbb{1}(x_j^1 < a_j) \cdot a_j + \mathbb{1}(x_j^1 > b_j) \cdot b_j\}$. The reasoning process follows iterative best responses. It converges to the Nash equilibrium after finite rounds of iterations.

In this paper, I adopt 14 two-person guessing games used by CGC06 and 4 two-person guessing games used by Georganas et al. [14,22]. The parameters of each game are given in Table 1. All the players survive at least two rounds of iterative best responses before reaching the equilibrium (as stated in Table 1 "steps to eqm" column). Since in CGC06, only a few number of subjects reached level 3 in the reasoning process, the choice of parameters in this paper should be sufficient to identify a player's strategic levels in the game.

3.2. Cognitive Load

Before directed to the guessing game, subjects were required to memorize a string of letters and were told that they need to recall the given string after the guessing game. The string was composed of either three or seven random letters, for example, UMH or WIEZOFH. The subjects were given 15 s to memorize the string; then they were automatically directed to the guessing game.

I did not pay the subjects specifically for correct recalls. However, their payments on the guessing game were partially dependent on this memorization task. If the recall for the selected payment round was wrong, they were not paid for that round, and left the experiment with only the participation fee. Said payment scheme incentivized the subjects to memorize the cognitive load correctly, and therefore guaranteed the effects of different cognitive load treatments.

Table 1. The eighteen two-person guessing games.

Game #	P1's Limits & Target	P2's Limits & Target	Treatment	P1's Role	P2's Role	Steps to Eqm	Eqm at Boundary
1	[(100,900); 1.5]	[(300,500); 0.7]	[LL+]	role B	role A	5+	-
2	[(300,900); 1.3]	[(100,500); 0.7]	[HL-]	role B	role A	5+	lower
3	[(300,900); 1.3]	[(300,900); 1.3]	[HH+]	role B	role A	3	upper
4	[(300,900); 0.7]	[(100,900); 1.3]	[LH+]	role A	role B	5+	lower
5	[(100,500); 1.5]	[(100,500); 0.7]	[LH-]	role B	role A	5+	upper
6	[(100,500); 0.7]	[(100,900); 0.5]	[HL+]	role A	role B	5	lower
7	[(100,500); 0.7]	[(100,500); 1.5]	[LH-]	role A	role B	5+	-
8	[(300,500); 0.7]	[(100,900); 1.5]	[LL+]	role A	role B	5+	upper
9	[(100,500); 0.7]	[(300,900); 1.3]	[HL-]	role A	role B	5+	-
10	[(300,500); 0.7]	[(100,900); 0.5]	[HH-]	role B	role A	3	lower
11	[(100,500); 1.5]	[(100,900); 0.5]	[LL-]	role B	role A	5+	-
12	[(300,900); 1.3]	[(300,900); 1.3]	[HH+]	role A	role B	3	upper
13	[(100,900); 1.3]	[(300,900); 0.7]	[LH+]	role B	role A	5+	-
14	[(100,900); 0.5]	[(300,500); 0.7]	[HH-]	role A	role B	4	-
15	[(100,900); 0.5]	[(100,500); 0.7]	[HL+]	role B	role A	4	lower
16	[(100,500); 0.5]	[(100,500); 1.5]	[LL-]	role A	role B	5+	lower
17	[(100,900); 1.3]	[(100,500); 0.5]	L	-	-	5+	-
18	[(100,900); 1.5]	[(100,500); 0.7]	H	-	-	5+	-

3.3. Treatments

The experiment consisted of two blocks. In the first block, subjects were assigned into pairs. They played 16 two-person guessing games against each other within the fixed pairs. In the second block, they played two guessing games against the computer. There were a total of 18 two-person guessing games for them to complete for this experiment, and no feedback was given throughout the process.

3.3.1. Against Human

I adopted a $2 \times 2 \times 2$ design. For ease of explanation, I specify the two players in the guessing game as having role A and role B in this section. However, subjects were not aware of their role during the experiment. Each subject was given the role of A or B for each treatment exactly once. I used a within-subject design.

To examine the effects of changing the cost of thinking on a subject's level of reasoning, I varied the cognitive load for role A, holding role B's cognitive load constant. As mentioned in the previous section, role A needed to memorize a string of either three or seven random letters when playing the guessing game. To test the effects of changing the opponent's cost of thinking on a player's level of strategic sophistication revealed in the game, I also varied role B's cognitive load by two levels. Changing the cost of thinking of role B essentially tests the effects of changing the first-order belief for role A. Denote the cognitive load of three letters as low load (L) and seven letters as high load (H).

Lastly, I varied the disclosure of information on the cognitive load for role B. The exact cognitive load implemented on role A was either fully revealed to role B or partially revealed as a probability distribution. Denote full revelation as [+] and the counterpart as [-]. In the partial revelation treatment, role B was told that role A has a 0.5 probability of memorizing a string of three letters and a 0.5 probability of memorizing a string of seven letters. The full and partial revelations of the cognitive load information on role B were a method of measuring the effects of changing the second-order belief for role A. In the full revelation treatment, both roles A and B were aware that role A's memorization task is common knowledge. However, in the partial revelation treatment, role A knew their exact memorization task was hidden to role B; therefore, their second-order belief (i.e., their belief about role B's belief of their own cost of thinking) may not coincide with their actual cost of reasoning. A summary of treatments is provided in Table 2. In later sections, I used role A's label to identify the treatments, as I was essentially examining the treatment effects for role A only. The first letter in the label indicates role A's cognitive load (either L or H). The second letter indicates role B's cognitive load (opponent's cognitive load, either L or H), and the last element of the label indicates full or partial revelation (role A's second order belief, either [+] or [-]). Role B served as a supporting role to complete the information required for each treatment. The information presented to role B for each treatment is also presented in Table 2. However, when later discussing the experimental results, I only refer to each treatment using role A's label. Table 1 provides a summary of treatments and assignments of roles for each game. Each subject played as either role A or role B exactly once for each treatment. There are in total 16 games. For each treatment, the pair of games are symmetric in game parameters and cognitive load realizations. The games were played in two random orders (the first order was as game numbers listed in Table 1; the second order was: 2, 13, 14, 4, 3, 1, 16, 6, 11, 8, 12, 5, 10, 15, 7, 9, 18, 17. Since for each game, there were two players assigned with different cognitive loads, considering player 2's order of play, there were essentially four sequences. The number of subjects in each order was roughly balanced. After dropping subjects with missing data, there were 28 subjects playing the first order as player 1, 29 subjects playing the first order as player 2, and 27 subjects playing the second order as player 1 and player 2 respectively.). Before the start of each session, one of the two was randomly selected.

Table 2. The eight treatments.

	Role	Label	Cost	1st Order Belief	2nd Order Belief
1	Role A Role B	[LL+]	Low Low	Low Low	Low (full revelation) Low (full revelation)
2	Role A Role B	[HL+]	High Low	Low High	High (full revelation) Low (full revelation)
3	Role A Role B	[LH+]	Low High	High Low	Low (full revelation) High (full revelation)
4	Role A Role B	[HH+]	High High	High High	High (full revelation) High (full revelation)
5	Role A Role B	[LL-]	Low Low	Low 50% Low, 50% High	50% Low, 50% High (partial revelation) Low (full revelation)
6	Role A Role B	[HL-]	High Low	Low 50% Low, 50% High	50% Low, 50% High (partial revelation) Low (full revelation)
7	Role A Role B	[LH-]	Low High	High 50% Low, 50% High	50% Low, 50% High (partial revelation) High (full revelation)
8	Role A Role B	[HH-]	High High	High 50% Low, 50% High	50% Low, 50% High (partial revelation) High (full revelation)

3.3.2. Against Computer

Subjects played against the computer for the second block of the experiment. The computer always chooses a Nash equilibrium action. The concept of equilibrium was explained to the subjects. For example, subjects were told that "a combination of guesses, one for each person, such that each person's guess earns them as many points as possible, given the other person's guess, is called an equilibrium guess." A similar description of equilibrium guess is found in CGC06. Subjects were also given an example of an equilibrium guess following this description. However, they were not specifically taught how to derive an equilibrium guess. The reason for introducing the equilibrium concept was to encourage the subjects to perform as many rounds of iterative best responses as possible. The two guessing games in this part are labeled 17 and 18 in Table 1.

3.4. Experimental Timeline

A total of 111 subjects were recruited for this experiment. Sessions were conducted at the Incentive Lab at Rady School of Management, University of California—San Diego (San Diego, CA, USA). The experiment was programmed and conducted using z-tree [23]. The complete session lasted for 90 min. Subjects were given a 5 USD show-up fee for attending the experiment and an additional $5 if they passed the understanding test and completed the experiment. They earned an additional $8 on average depending on their decisions for the guessing games. For those who did not pass the understanding test, they spent about 30 min in this experiment and left with the $5 show-up fee.

Subjects were given instructions on the two-person guessing game first. After explaining the rules, I introduced four unincentivized practice rounds. During the practice rounds, subjects played against the computer and were told that the computer will always choose the mean of the target interval. After the subjects made a guess, feedback was provided for the subjects to reflect on the game rule and the payoff rule. An understanding test was then administered. The test was composed of six questions, similar to the understanding test in CGC06. Standard questions included calculatiosn of best responses and payoffs. Although subjects in the experiment were not restricted to following a level-k reasoning process, for the purpose of the experiment, I wanted to make sure the subjects were capable of calculating the best responses. A screenshot of the understanding test is shown in Figure 1. Subjects needed to answer four out of six questions correctly to proceed to the main part of the experiment.

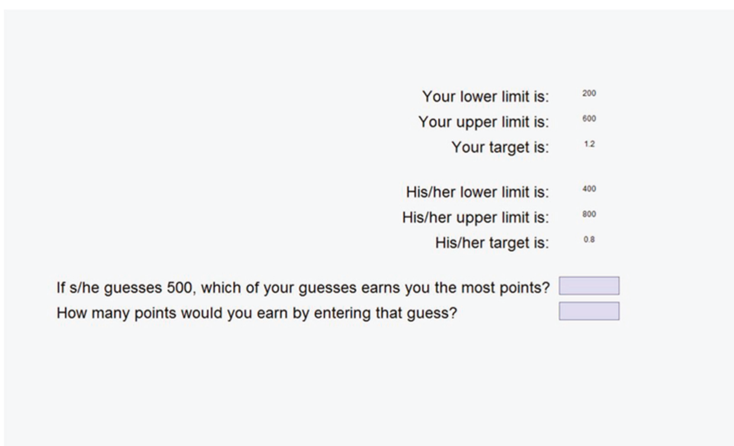


Figure 1. Screenshot of the understanding test.

Before playing the incentivized guessing games, subjects were introduced to the memorization task. They were given two unincentivized practice rounds for the low load and high load treatments. During the practice round, they had the standard 15 s to memorize the string of letters and were asked for immediate recall when the time was up. They, however, did no get to practice the guessing game with the cognitive load implemented.

The main experiment consisted of two parts, as discussed in Section 3.3. There were 18 two-person guessing games in total. For the first 16 games, subjects were randomly assigned into pairs and stayed within the same pair for all 16 decisions (one as player 1 and the other as player 2). For each game, subjects were given the same information set that consisted of the types of memorization task (either string of three or seven letters, or a probability distribution) for themselves and their opponents, whether their opponents knew about their exact memorization task, and the targets and limits for both players. An example of the actual decision screen is provided in Figure 2. Subjects were also asked to elicit their opponents' types of memorization task after they made their guesses and recalled the letters. This practice allowed me to check whether the subjects received and processed the correct information about their strategic environment. There was no feedback given in between the 18 guessing games. This prevented the subjects from learning anything about their opponents' past actions. Such practice also limited the subject's learning of the guessing game, as no payoff information was provided. (There was limited learning of the game. Upon checking a subject's levels with respect to the orders of the games they played, playing a later game was not associated with higher *k*-levels. The coefficient from the OLS regression was 0.005 and it was not statistically significant.)

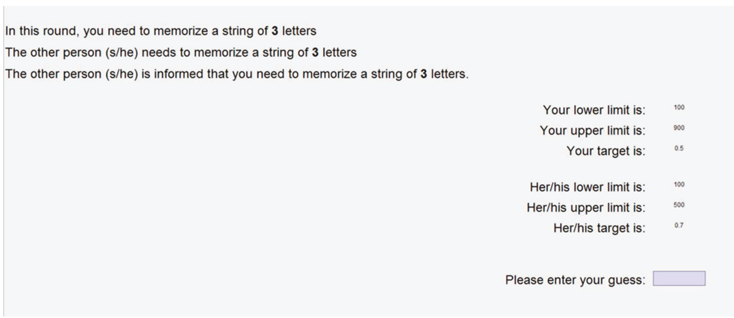


Figure 2. Screenshot of the incentivised two-person guessing game.

Subjects took a 10-question Mensa practice test at the end of the experiment. The test is used to measure the subject's analytical ability. Some questions ask the subject to identify the missing element that completes a sequence of patterns or numbers. Some questions are verbal math questions. A couple of studies in economics literature have used a similar test as a measure of cognitive ability [22]. I used this test in the experiment to measure whether there were any heterogeneous treatment effects on subjects with different exogenous cognitive abilities.

3.5. Discussion of the Experimental Design

First, I used letters to compose strings for the cognitive load treatment, unlike the conventional use of binary numbers [16]. This design restricts the subjects from using the cognitive load numbers as their inputs for the guessing game. It allows a clear separation of the two tasks, the memorization task and guessing game, and therefore increases the reliability of the treatment effects of cognitive load. I recognized subjects may be able to use other methods to memorize the string of letters, for example, using hand gestures. However, any such method also requires cognitive effort and therefore should not significantly lessen the effects of cognitive load for treatment purposes.

Subjects remained within the fixed pair for the first 16 incentivized guessing games. Since no feedback was given in between games, this design ensures the manipulation of cognitive load being the only source of changing beliefs for any subject. Subjects were different exogenously in terms of cognitive ability, so by staying in the same pair, they carried the same beliefs about their opponents' cognitive abilities throughout the whole session.

Lastly, for each of the 18 incentivized tasks, subjects were given 90 s to make a decision for the guessing game. According to Agranov et al., 90 s is enough for strategic players to make a decision in this type of guessing game [24]. To keep the effect of cognitive load constant across players, I only allowed the subjects to submit their guesses after the 90 s was up. Said practice avoids some subjects naïvely picking a guesses without strategic contemplation for the purpose of achieving correct recalls for the memorization task.

4. Data Analysis Procedure

All the subjects played 18 games in total, each against a fixed opponent during the experiment. There were 1998 observations of guesses. Grouping the guesses by games, I looked for level shifts observed with raw guesses. This exercise provided a general view of the effects of cognitive load on the games. I also used density plots of the guesses to visualize the treatment effects.

After the exploration of raw guesses, I estimated the level for each guess using the maximum likelihood method. Instead of assuming the subject's behaviors are determined by a single type across all the games, I assumed the subject's behavior in each game was determined by a single type and the types across games were allowed to be different. This was achievable with the design of my experiment with the variations on cognitive load.

Out of 1998 observations, 831 guesses correspond to a type's exact guesses. As about 40% of the observed guesses were a type's exact guesses, I followed the CGC06 approach in my estimation. Specifically, for each player i, game g, and level k, if player i was not making a type's exact guesses in game g, then I defined a likelihood function $L(y_{ig} \mid k, \lambda)$ for each level k for the player in that game, with beliefs $f_g^k(y)$ and sensitivity parameter λ, based on the assumption that they were trying to maximize their expected utility.

Formally, let x_{ig} be the raw guess observed for player i in game g. With the specification of lower limits a_{ig} and upper limits b_{ig}, the adjusted guess is then $y_{ig} = \min\{\max\{a_{ig}, x_{ig}), b_{ig}\}\}$. The density $f_g^k(z)$ represents a subject's belief about his opponent's action given their behavioral level being k. Although in the literature a subject's belief of the other player's level could follow a certain type of distribution, for example, Poisson distribution as in Camerer et al. (2004), in this study, I followed the standard approach that level-k player has point belief about his opponent, that his opponent is level-$(k-1)$ with probability 1. y_g^0 is defined as uniformly spread across the action space. The expected payoff of playing x_{ig} with behavioral level k's belief is then:

$$U_{ig}^k(y_{ig}) = \int_1^{1000} U_{ig}(y_{ig}, z) f_g(z) dz. \tag{4}$$

Let $U_{ig}^k = [\max(y_g^k - 0.5, a_{ig}), \min(y_g^k + 0.5, b_{ig})]$ be the interval of a type-k subject's exact adjusted guesses, allowing an error of 0.5. Any guess for game g, subject i, who is placed within U_{ig}^k, is then identified as an exact match for k-level. Conversely, define $U_{ig}^{k\ \complement} = [a_{ig}, b_{ig}] / U_{ig}^k$ as the complement of U_{ig}^k within the limit interval for subject i's game g. The likelihood function is then the following:

$$L(y_{ig}|k, \lambda) = \frac{exp[\lambda U_{ig}^k(y_{ig})]}{\int_{U_{ig}^{k\ \complement}} exp[\lambda U_{ig}^k(w)] dw}. \tag{5}$$

Since only one observation was used for the estimation, I took the sensitivity parameter (λ) as 1.33, which is the averaged estimated value of λ in CGC06 with only the subject's guesses. The maximum likelihood estimate of a subject's behavioral level in each game maximizes (5) over k, which is:

$$k_{ig} = \underset{k \in \{1,2,3,4,5,6\}}{\arg\max} \ L^*(y_{ig}|k). \tag{6}$$

To examine the treatment effects on behavioral levels, I pooled guesses into pairs for comparison. For example, to test the prediction on the changing cost of reasoning, I first identified games with the same first-order belief (either low or high cost of reasoning for opponent) and the same second-order belief (partial revelation), and then they were separated into comparison pairs by the subject's cognitive load tasks. The same selection was performed following the conditions listed in each testable prediction.

For each pair of games, I first conducted a binary comparison on their behavioral levels and I report the summary statistics. Since this is essentially a repeated measure of behavioral level from the same sample, I then conducted the Wilcoxon signed-rank test to check the distribution of behavioral levels. Lastly, I ran a GLS random effect regression to examine the treatment effects on behavioral levels. The regression was run by regressing the estimated level on the treatment variable. A subject's cognitive load was coded as 0 when it was in the low load treatment, and 1 when it was a high-load treatment. The same binary coding was also applied to the opponent's cognitive load treatment. The full revelation of information treatment was coded 0, whereas partial revelation was coded 1.

5. Results

5.1. General Examination of Raw Guesses

There were a total 1998 observations and 831 guesses corresponded to a specific level (levels 1 to 5, and equilibrium). When identifying levels, I assigned the lowest possible level to a guess that matched multiple types. For example, in game 3, equilibrium was reached after three rounds of iterative best responses, and the equilibrium was at the boundary of the target interval. In this case, although levels 3, 4, and 5, and the equilibrium all have corresponding guesses at 900, a subject's guess of 900 only assigned the subject to type level 3. This method of identification restricted over-assignments of the types.

Figure 3 shows the distribution of guesses that matched specific levels. Of the 831 guesses that matched a specific level, 43.92% were level 1 guesses, 31.41% were level 2 guesses, 14.20% were equilibrium guesses, and level 3 and higher corresponded to the remaining 10% of the guesses. To provide a clearer picture of the treatment effect, I used a Markov matrix for some treatments with these exactly matched guesses. Tables 3 and 4 present the level transitions between comparable games. For example, Table 3 consists of all the comparable pairs of changing a subject's own cost of reasoning, fixing the opponent with a high cognitive load (game 7 [LH-] and game 14 [HH-]). There were a total of 111 pairs of comparisons, 24 of which had both guesses that exactly matched a specific level. From games 7 to 14, 12 subjects reached level 1 in game 7 and 83.33% stayed at level 1 in game 14. Eight subjects reached level 2 in game 7, 87.5% of which stayed at level 2 and below in game 14. This result largely complies with the theory prediction that increasing cost of reasoning while fixing first- and second-order belief constant decreases the level of reasoning weakly. Likewise, Table 4 presents all the comparison pairs of changing the subject's first-order belief while fixing their own cost of reasoning and keeping their second-order belief constant (game 1 [LL+], game 8 [LL+], game 4 [LH+], and game 15 [LH+]). There were a total of 444 pairs of comparison, 99 of which had both guesses matched to a specific level. Forty pairs had level 1 guesses in the [LL+] treatment and 62.5% of them remained level 1 in the [LH+] treatment games. Similarly, 27 pairs had level two guesses in the [LL+] treatment. When changing the subject's first-order belief by increasing the cognitive load of their opponents, about 90% of these pairs had level 2 or lower guesses in the [LH+] treatment. These statistics largely coincided with the theoretical prediction—with increasing the cost of reasoning for the opponents, the subjects adjusted by weakly decreasing their behavioral levels of playing the game. Due to the limited number of exact matches, I was not able to conduct the same exercise for all the treatment pairs. However, complete discussion of the treatment effects is provided below with estimated behavioral levels.

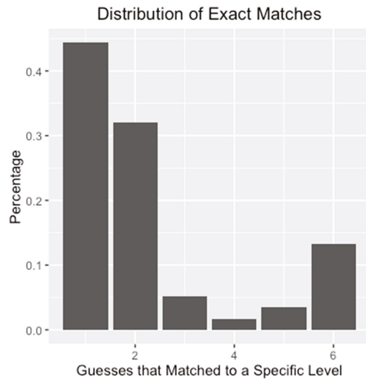


Figure 3. Distribution of exact matches.

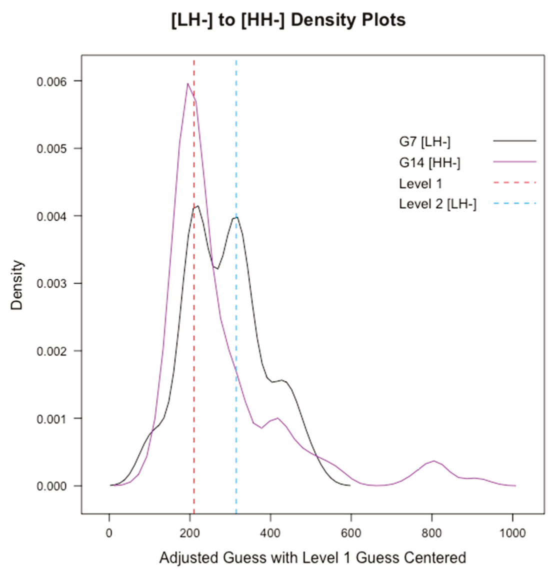
Table 3. Markov matrix of level transitions for increasing cost of reasoning, opponent with high load.

↓ to →	Level 1	Level 2	Level 3	Level 4	Level 5	Eqm	Num
Level 1	**83.33%**	0	16.67%	0	0	0	12
Level 2	**25%**	**62.5%**	12.5%	0	0	0	8
Level 3	0	0	**100%**	0	0	0	1
Level 4	0	0	0	0	0	0	0
Level 5	0	0	0	0	0	0	0
Eqm	0	**66.67%**	**33.33%**	0	0	0	3

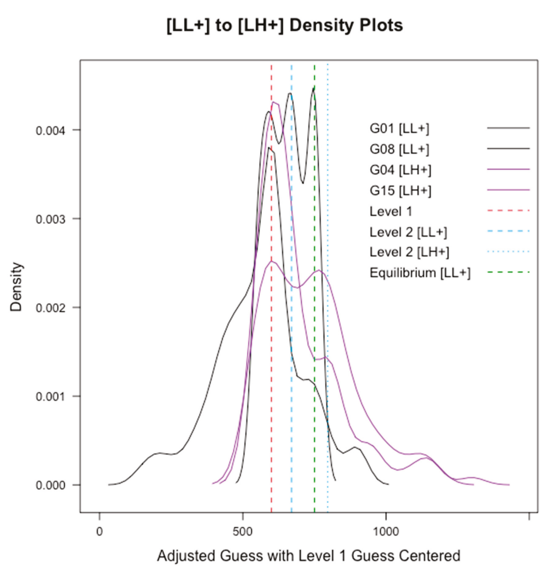
Table 4. Markov matrix of level transitions for changing first-order belief, subject with low load.

↓ to →	Level 1	Level 2	Level 3	Level 4	Level 5	Eqm	Num
Level 1	**62.5%**	7.5%	2.5%	0	15%	12.5%	40
Level 2	**3.7%**	**85.19%**	7.41%	0	0	3.7%	27
Level 3	0	0	0	0	0	0	0
Level 4	0	0	0	0	0	0	0
Level 5	0	0	0	0	0	0	0
Eqm	**31.25%**	**18.75%**	**28.13%**	0	**21.88%**	0	32

The pattern of subjects' adjustments to the changing strategic environment is also illustrated with density plots of each game. This time, all the raw guesses (after adjustments according to upper and lower limits) were used to plot the graphs. Figure 4 illustrates the treatment effects for the three theoretical predictions. To better compare across games, level 1 guesses were centered, and all the guesses were adjusted accordingly. The colored vertical lines illustrate the level-exact guesses. For example, in Figure 4a, the vertical red dashed line indicates level-1 guesses. Both density plots in the figure show peaks around the red vertical line, which indicate higher proportions of level-1 (or close to level 1) strategy used within the games across all the subjects. Notably, in the density plot for the [LH-] treatment (G7), there is another peak centered right at the level 2 guess for that game (indicated by blue dashed line). The density plot clearly shows that in the game where subjects have a lower cost of reasoning ([LH-]), guesses are congregated at both levels 1 and 2, whereas in the game where subjects have a higher cost of reasoning ([HH-]), only a peak at the level-1 guess is observed. Likewise, in Figure 4b, four games are plotted to illustrate the treatment effects of increasing cost of reasoning for the opponent. In Figure 4c, three games are used to demonstrate changing second-order beliefs. Note that both games 1 and 8 are relevant in both graphs, as the [LL+] treatment is relevant for both comparisons. As illustrated in the figure, in one of the games, the three peaks correspond to level 1, level 2, and equilibrium. When increasing the cost of reasoning for the opponent, the level 1 peak is still observable; however, only one game has a level-2 peak. Similarly, when changing the second-order belief from low load with probability 1 to (0.5, 0.5; L, H), only the level 1 peak remains, as then the subjects thought that their opponents thought there was a 50% probability that the subject was experiencing a high cognitive load. I omitted other vertical lines that indicated different levels due to the absence of peaks in the density plots.

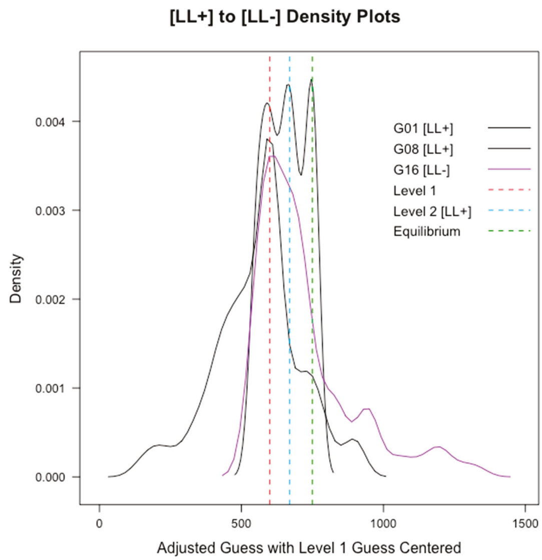


(**a**) Density plot of changing cost of reasoning (opponent high load).



(**b**) Density plot of changing cost of reasoning for opponent (subject low load).

Figure 4. *Cont.*



(**c**) Density plot of changing second-order belief (LL).

Figure 4. Density plot of raw guesses.

5.2. Distribution of Levels

From the preview of results from raw guesses in the previous subsection, changing the strategic environment appeared to lead to some structured changes in the depth of reasoning. However, only about half of the guesses were type-exact guesses. To better understand the treatment effects of the other half, I used maximum likelihood estimation to assign types, and then conducted analyses based on the estimated levels.

There were a total 1998 observations of guesses. As discussed in the previous section, I assigned a behavioral level for each observation. Surprisingly, a few guesses corresponded to exact level 4 and level 5 guesses in my data. Therefore, I included levels 1 to 5 and the Nash equilibrium type in my estimation. Of all the observations, 1167 guesses were estimated. The distribution of estimated levels for these guesses is shown in Table 5. The majority of the guesses were assigned to level 1 guesses. The level distribution for all the guesses is shown in Table 6. The game number is referred to the game number list in Table 1. Since all the subjects played each game exactly once, for each game listed, there were 111 observations.

Table 5. Summary of estimation results.

	L1	L2	L3	L4	L5	Nash
Exact	43.92%	31.41%	5.54%	1.56%	3.37%	14.20%
Estimated	71.89%	5.57%	12.94%	5.57%	3.43%	0.60%
Total	60.26%	16.32%	9.86%	3.90%	3.40%	6.26%

Table 6. The frequency of levels by game.

Game #	L1	L2	L3	L4	L5	Nash
All Guesses	**60.26%**	**16.32%**	**9.86%**	**3.90%**	**3.40%**	**6.26%**
1	72.97%	13.51%	0	2.70%	4.50%	6.31%
2	35.14%	16.22%	19.82%	15.32%	3.60%	9.91%
3	44.14%	22.52%	33.33%	0	0	0
4	63.06%	14.41%	1.80%	7.21%	13.51%	0
5	72.07%	12.61%	0	1.80%	0	13.51%
6	83.78%	4.50%	2.70%	9.01%	0	0
7	53.15%	12.61%	16.22%	5.41%	6.31%	6.31%
8	62.16%	12.61%	1.00%	0	0	24.32%
9	59.46%	6.31%	18.02%	1.80%	7.21%	7.21%
10	54.05%	45.95%	0	0	0	0
11	69.37%	10.81%	7.21%	1.80%	1.80%	9.01%
12	44.14%	23.42%	32.43%	0	0	0
13	70.27%	16.22%	1.00%	1.00%	5.41%	6.31%
14	73.87%	11.71%	14.41%	0	0	0
15	65.77%	15.32%	11.71%	7.21%	0	0
16	64.86%	16.22%	5.41%	2.70%	10.81%	0
17	34.23%	28.83%	3.60%	10.81%	2.70%	19.92%
18	62.16%	9.91%	9.91%	3.60%	5.41%	9.91%

The distributions of the levels were fairly similar to the results in CGC06, except that levels 4 and 5 were then included. Level 1 was the most prominent behavioral level. Of 1998 observations, 60.26% were level 1 guesses. In some games, level 1 was even more frequently observed. For example, in game 1, about 70% of the guesses were classified as level 1. A number of observations were levels 2 and 3 and Nash guesses. In my data, the occurrence of level 3 was more frequent in a few games. For example, in game 2 and game 3, more than 20% of observations were assigned to level 3. Although some observations corresponded to exact level 4 or level 5 guesses, the overall frequency of these two higher levels was much lower. In about one-third of the games, no guesses were classified into these two levels.

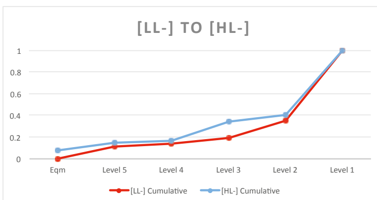
As shown in Table 6, there are a pair of games that have almost identical level distribution, game 3 and game 12. These two games have identical parameters and treatments (as shown in Table 1). Besides these two games, the frequency of levels in other games differed considerably. In some games, behavioral levels congregated toward levels 1 or 2, for example, games 1 and 6. In some games, such as games 2 and 9, behavioral levels spread out across the six categories. The variations in the distribution of levels across games could be due to the differences in the cognitive load tasks. The exact impact of the memorization tasks is discussed in detail in the following subsections.

5.3. Result 1: Increasing Cost of Reasoning

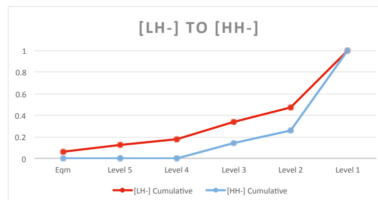
As mentioned in Section 2.2, the first testable prediction involved fixing the subject's first- and second-order beliefs and examining the effect of the changing cost of reasoning on the subject's behavioral levels. There were essentially two comparisons in this case: a comparison between treatment [LL-] and treatment [HL-], and between treatment [LH-] and [HH-]. Note that in both comparisons, the cost of reasoning for the subject varied from low to high; therefore, it was crucial to have partial revelation of the subject's (role A) memorization task. In the partial revelation treatment, role B (the opponent) only knew the probability distribution of the subject's memorization task (0.5, 0.5; L, H); therefore, even with the subject's own tasks varying between two treatments, the subject's second-order belief was controlled to be the same. There were 222 pairs of comparison in total. The summary statistics of the comparisons are presented in Table 7. The plotted distribution of behavioral levels is presented in Figure 5. To aid with the interpretation of the results, the behavioral levels in the figure are presented in a reverse order (i.e., higher level on the left and lower level on the right).

Table 7. The frequency of changing behavioral levels with increasing cost of reasoning.

Pair Name (From Game a to Game b)	# of Pairs	Treatment	Decreases	Constant	Increases
G16 to G9	111	LL- to HL-	20.72%	43.24%	36.03%
G7 to G14	111	LH- to HH-	39.64%	49.55%	10.81%
Combined	222	L?- to H?-	30.18%	46.40%	23.42%



(a)



(b)

Figure 5. Cumulative level distribution for increasing cost of reasoning.

When the opponent's cognitive load was controlled to be high and with partial revelation, subjects weakly decreased their behavioral levels 89.19% of the time (39.64% strict decrease). In Figure 5b, the [LH-] treatment is first-order stochastic dominant over the [HH-] treatment. The Wilcoxon test (Table 8) was significant at the 1% level for the comparison of the distributions of behavioral levels between these two strategic environments. When regressing the behavioral level on the treatment dummy, the result (Table 9) suggested that the coefficient for treatment dummy was 0.77, which was significant at the 1% level. This implied that the estimated behavioral level weakly decreased when the subject's own cognitive load increased when facing an opponent with high cognitive load. The finding is consistent with the EDR model. The relatively large proportion (49.55%) of constant levels may seem quite surprising at first look. One possible explanation is that these subjects may have had different cognitive bounds in the two treatments. In the [LH-] treatment, subjects may have adjusted their behavioral levels downward from their cognitive bound in that treatment due to some belief they formed when facing opponents with high cognitive loads. In the [HH-] treatment, subjects who had a lower cognitive bound (as they had a high cognitive load) may have displayed a lower behavioral level. When the two behavioral levels from two treatments coincided, I observed no changes in the behavioral levels in the treatment comparison.

The result for the comparison between [LL-] and [HL-] is less clear. As shown in Table 7, 63.97% of the comparisons had weakly decreasing behavioral levels (20.72% strict decrease), and a noticeable percentage (36.03%) of the comparisons had increasing levels. The Wilcoxon test statistic rejected the null hypothesis that the two strategic environments have the same distribution of behavioral levels at the 5% level. However, upon further checking using a one-tail Wilcoxon test, the distribution of behavioral levels shifted rightward when cognitive load changed from low to high when facing an opponent with a low cognitive load. When conducting the standard GLS random effect regression, the coefficient on the treatment dummy was positive and significant at the 10% level.

Table 8. Test results for equality of distribution.

Comparison Group (in Treatment)	Wilcoxon p-Values (Two-Tailed)	Wilcoxon p-Values (One-Tailed)
Changing Cost of Reasoning		
LL- to HL-	0.05 **	0.98
LH- to HH-	0.00 ***	0.00 ***
Changing Opponent's Cost of Reasoning		
LL+ to LH+	0.01 ***	0.00 ***
HL+ to HH+	0.00 ***	1
Changing Second Order Belief		
LL+ to LL-	0.11	0.05 **
LH+ to LH-	0.00 ***	0.99
HL- to HL+	0.00 ***	0.00 ***
HH- to HH+	0.00 ***	1
Against Computer (Nash)		
L to H	0.00 ***	0.00 ***

* indicates < 10% significance, ** indicates < 5% significance, and *** indicates < 1% significance.

Table 9. Regression results for treatment effects.

Comparison Group (in Treatment)	Relevant Dummy	Constant	Number of Obs.
Changing Cost of Reasoning			
LL- to HL-	0.34 * (0.18)	1.78 *** (0.14)	222
LH- to HH-	−0.77 *** (0.16)	2.18 *** (0.12)	222
Changing Opponent's Cost of Reasoning			
LL+ to LH+	−0.27 * (0.14)	2.03 *** (0.11)	444
HL+ to HH+	0.33 *** (0.10)	1.55 *** (0.07)	444
Changing Second Order Belief			
LL+ to LL-	−0.25 (0.18)	2.04 *** (0.12)	333
LH+ to LH-	0.41 *** (0.15)	1.77 *** (0.09)	333
HL+ to HL-	0.57 *** (0.15)	1.55 *** (0.10)	333
HH+ to HH-	−0.48 *** (0.09)	1.89 *** (0.06)	333

* indicates < 10% significance, ** indicates < 5% significance, and *** indicates < 1% significance. Standard errors in parenthesis.

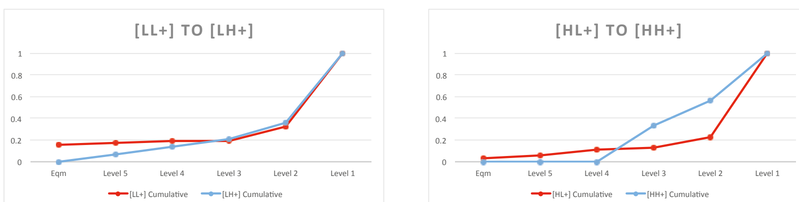
In this analysis, I treated equilibrium level as the highest level, since it requires the subjects to perform multiple steps of iterative best responses. However, since many games have equilibrium at the boundary of the limit interval (games 2, 4, 6, 10, 15, and 16 have the equilibrium at the lower limit; games 3, 5, 8, and 12 have the equilibrium at the upper limit), if the subject chooses an equilibrium action by naïvely playing at the boundary, then this behavioral level should not be considered as a higher level than any of the k levels. This was not the case for this comparison pair. Although game 9 ([HL-]) had 7.21% equilibrium guesses, those guesses were not at the boundary. However, upon further checking of games 16 ([LL-]) and 9 ([HL-]), I found that the level-5 type in game 16 had the same strategy as the equilibrium strategy of that game. Therefore, some of the equilibrium strategies in game 16 were pooled into level-5 type, which may be one possible explanation for the significant positive coefficient on the treatment dummy. Another explanation may be that the subjects felt more motivated to reason at higher strategic levels when they saw the opponents had easier strategic environments (memorizing three letters) as opposed to their own difficult strategic environments (memorizing seven letters). As a result, they displayed higher behavioral levels. This explanation suggests that other factors, such as motivation factor, may also play a role in determining a subject's behavioral levels.

5.4. Result 2: Increasing Cost of Reasoning for Opponent

To examine the effect of changing the first-order belief on a subject's behavioral level in games, I selected pairs of games with changing cognitive loads for the opponents. For example, a comparison of behavioral levels for games 1 and 4 served the purpose. In game 1 ([LL+]), player 1 has a low cognitive load when facing an opponent with a low cognitive load, and there is full revelation of each other's strategic environment. In game 4 ([LH+]), player 1 has a low cognitive load when facing an opponent with high cognitive load, and again, there is full revelation of the treatments. I found 444 pairs of comparison for the cases wherein the subjects had low cognitive loads, and another 444 pairs of comparison for the cases when they had high cognitive loads. The detailed comparison groups and summary statistics are shown in Table 10. The plotted distribution of behavioral levels is presented in Figure 6.

Table 10. The frequency of changing behavioral levels with increasing cost of reasoning for opponent.

Pair Name (From Game a to Game b)	# of Pairs	Treatment	Decreases	Constant	Increases
(G1, G8) to (G4, G15)	444	LL+ to LH+	23.87%	55.86%	20.27%
(G6, G13) to (G3, G12)	444	HL+ to HH+	15.54%	41.44%	43.02%
Combined	888	?L+ to ?H+	19.71%	48.65%	31.64%



(**a**) (**b**)

Figure 6. Level Distribution for increasing cost of reasoning for opponent.

The combined results were the opposite of the theory prediction, with a significant 31.64% of cases of increasing behavioral levels. However, upon further checking, the majority of the increasing cases occurred when subjects are having high cognitive load. When subjects had low cognitive load, 79.73% of the time, they weakly decreased their behavioral levels when their opponents' cognitive loads changed from low to high (23.87% strict decrease). Figure 6a illustrates that [LL+] games had more guesses at higher levels. This result is consistent with the EDR model. When a subject's cost and second-order belief was controlled across the two strategic environments, he was responsive to the changes in his opponent's cost of reasoning. However, some of these adjustments in behavioral levels were not strictly decreasing. If the subject believed that the increased opponent's cost of reasoning was not large enough to decrease the opponent's behavioral level by one, the subject's behavioral level remained the same across the two strategic environments. This partially explains the high percentage (55.86% and 41.44%) of constant behavioral levels in Table 10. When the subject had a high cognitive load and his opponent's cognitive load changed, the result did not comply with the EDR model. A total of 43.02% of the pairs showed increasing behavioral levels across the two strategic environments. The frequency of levels in Table 6 reveals that most subjects had level 1 guesses in games 6 (83.78%) and 13 (70.27%). This gave subjects much less room to adjust their behavioral levels downward compared to another strategic situation. Any behavioral level that was beyond level 1 in games 3 and 12 was considered as moving the behavioral level upward. This was one major limitation in observing the effects of changing the first-order belief when the subject had a high cost of reasoning (i.e., high cognitive load).

The Wilcoxon signed-rank test rejected the null hypothesis that the level distribution was the same for both treatment comparisons ([LL+] to [LH+] and [HL+] to [HH+]). However, the one-tail test suggested that when the subject had low cognitive load, increasing his opponent's cost of reasoning shifted the former's level to the left (to lower levels, significant at the 1% level). However, when the subject had high cognitive load, the level distribution shifted to the right. The regression coefficients suggested that increasing the opponent's cost of reasoning decreased the behavioral level when the subject had a low cognitive load (significant at the 10% level).

5.5. Result 3: Changing the Second-Order Belief

In the experiment, I used a (0.5, 0.5) probability distribution on the revelation of cognitive load treatments to control for the subject's second-order belief. In the full revelation treatment, role B knew the exact memorization task that was received by role A (the subject), either three (low load) or seven letters (high load) with a probability of one. Therefore, role A's (the subject) second-order belief was either ((1, 0); (L, H)) or ((0, 1); (L, H)). In the partial revelation treatment, role B knew that the probability of three or seven letters for role A was (0.5, 0.5), which made role A (the subject) have a second-order belief of ((0.5, 0.5); (L, H)). If comparing two games with different second-order beliefs for the subject, with everything else controlled as constant, then a second-order belief of low load with probability of one should be considered as more cognitively capable perceived by role B than a second-order belief of((0.5, 0.5); (L, H). The experiment, as shown in Table 11, supported that most subjects had a clear understanding of their opponent's cognitive load when the load was explicitly elicited, and they almost had uniform beliefs about their opponents' cognitive loads when they were in the partial revelation treatment as role B.

Table 11. Subject's belief about his opponent's cognitive load.

		Belief Elicitation			
		3 Letters	**7 Letters**	**Not Sure**	**Sum**
Treatments	3 Letters	591	51	24	666
	7 Letters	97	546	23	666
	(0.5 L, 0.5 H)	143	111	190	444

In the dataset, I found 888 pairs for comparison that allowed me to examine the effect of changing the second-order belief. I separated them into two groups: a comparison between the full revelation of low load to partial revelation, and a comparison between a partial revelation and a full revelation of high load. Both comparisons were performed in the direction of increasing second-order belief (i.e., c_i^{ij} increases). The detailed comparison pairs and summary statistics are listed in Table 12. The distribution of behavioral levels is plotted in Figure 7.

Table 12. The frequency of changing levels with changing second-order belief.

Pair Name (From Game a to Game b)	**# of Pairs**	**Treatment**	**Decreases**	**Constant**	**Increases**
Second order belief: Low to (0.5 Low, 0.5 High)					
(G1, G8) to G16	222	LL+ to LL-	25.68%	51.35%	22.97%
(G4, G15) to G7	222	LH+ to LH-	20.27%	43.69%	36.04%
Combined	444	L?+ to L?-	22.97%	47.52%	29.50%
Second order belief: (0.5 Low, 0.5 High) to High					
G9 to (G6, G13)	222	HL- to HL+	35.14%	52.70%	12.16%
G14 to (G3, G12)	222	HH- to HH+	14.86%	45.50%	39.64%
Combined	444	H?- to H?+	25.00%	49.10%	25.90%

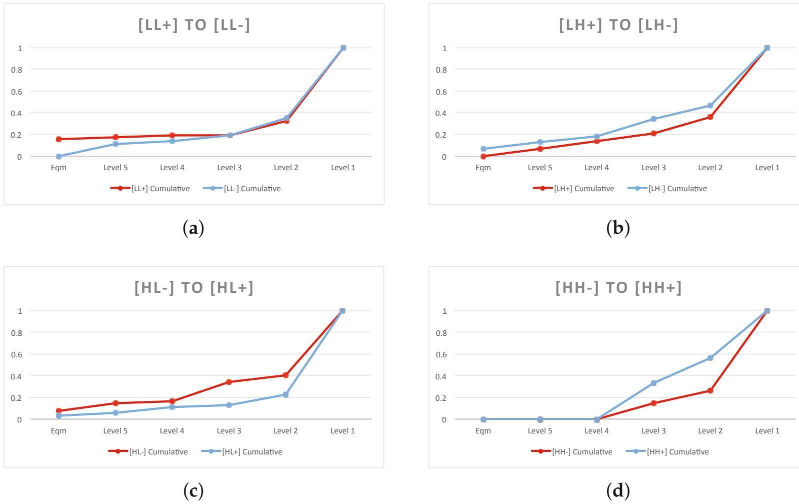


Figure 7. Level Distribution for changing second-order belief.

The effect of changing the second-order belief was generally weak, except for the cases where the subjects had high cognitive loads when facing opponents with low cognitive loads. For the treatment where both players had low cognitive loads, about 77.03% of the pairs had weakly decreasing behavioral levels when second-order belief changed from full to partial revelation. Among these comparisons, only 25.68% had strictly decreasing levels. This finding suggested that the changes in second-order belief may not have been strong enough for the subjects to adjust their behavioral level downward, even though both subjects had a low cognitive load and were relatively competent at contemplating over the strategic environment. To examine the effect of the second-order belief, it was first necessary to determine the effect of changing the first-order belief for the same group of subjects. In Table 10, the subject's behavioral responses to the changing opponent's cognitive load were limited when the subject had a high cognitive load. Now, consider the finding in the [LH+] to [LH-] comparison to the [HH-] to [HH+] comparison (Table 12); changing the second-order belief of the subject effectively changed his opponent's first-order belief. If the subject holds the belief about his opponent (who has a high cognitive load treatment) that the changes in his opponent's behavioral level are limited, then the subject should not decrease his behavioral level at all. This partially explains the low frequency of strictly decreasing behavioral levels for subjects who faced opponents with high cognitive loads.

The comparison between [HL-] and [HL+] is consistent with the EDR model. In Figure 7c, the [HL-] treatment is first-order stochastic dominant over the [HL+] treatment. Of the guesses, 87.84% had weakly decreasing behavioral levels, with 35.14% having strict decreases. Changing from partial revelation to full revelation of high cognitive load, the second-order belief decreased the subject's cognitive capability perceived by their opponent. Subjects were responsive to this change in the belief system, and adjusted their behavioral levels downward to best respond to their opponents. Testable prediction 3 suggests that if the subject's behavioral level is binding by their cognitive bound, then they are not able to make further adjustments according to their changing beliefs. The large percentage of constant levels for these comparisons supported this statement.

The Wilcoxon test results showed that the level distribution changed for changing second-order belief. When conducting a one-tailed test, the test result suggested that for [LH+ to LH-] and [HH- to HH+] treatments, the distribution of levels significantly (at the 1% level) shifted rightward (increasing behavioral levels). This may have occurred due to the subject's belief that their opponent with high cognitive load will engage in higher behavioral level. This result seems to comply with the results in Section 5.4, but the underlying reasons need further investigation.

The regression coefficient on the treatment dummy further supported the results. Since the treatment dummy was coded as zero with full revelation and one with partial revelation, the coefficient of 0.57 for [HL-, HL+] comparison suggested that the behavioral level decreased from partial to full revelation. It was significant at the 1% level. Again, the [LH+, LH-] and [HH-, HH+] comparisons were the opposite direction of model predictions, and they were also highly significant. In general, when the subjects faced opponents with high cognitive loads, they were responsive to changing second-order beliefs, but not in the direction that is predicted by the EDR model. However, when they faced more cognitively capable opponents, then they were mostly responsive to this change in the belief system because they thought their opponents were responsive to this information in their strategic environment. This finding is consistent with the EDR model when the opponent has a low cognitive load, which supports the opposite direction when the opponent is in a less cognitively capable situation.

5.6. Result 4: Cognitive Bound

In block 2 of the experiment, the subjects played against the computer. They were told that the computer was playing a Nash equilibrium strategy, and the equilibrium concept was explained. However, they were not taught the method to derive the equilibrium. The behavioral levels from the guesses in these two games should be considered as the highest levels they could achieve under each cognitive load treatment. I selected all the games with the same cognitive load treatment, either low cognitive load or high cognitive load, and pooled the results. A pairwise comparison between the pooled data and behavioral level obtained from games 17 and 18 allowed me to examine the existence of cognitive bounds. There were 888 pairs of comparison for each type of cognitive load, and the summary statistics are shown in Table 13.

Table 13. The frequency of changing behavioral levels comparing to cognitive bound.

Cognitive Bound	# of Pairs	Decreases	Constant	Increases
Low cognitive bound (G1, G2, G4, G7, G8, G11, G15, G16) to G17	888	18.36%	33.45%	48.20%
High cognitive bound (G3, G5, G6, G9, G10, G12, G13, G14) to G18	888	21.96%	48.20%	29.84%
Combined	1776	20.16%	40.82%	39.02%

The result for low cognitive load treatment was interesting: 48.20% of the guesses from block 1 games had behavioral levels lower than the subject's respective cognitive bound (level in game 17). Less than 20% of guesses had higher behavioral levels. This suggested that in many block 1 games, subjects purposely adjusted their behavioral levels downward due to different strategic situations, even though they had reached higher levels. For high cognitive load treatment, about 30% of behavioral levels increased from block 1 games to game 18. However, about 50% of the guesses had the same behavioral level across the two situations. Since high cognitive load inherently restricts the subject's cognitive ability, there may have been less room for downward adjustments for block 1 games. Due to the large percentage of weakly increasing levels from block 1 to block 2 games, I concluded that cognitive bound existed in most cases. In some situations, cognitive bound was strictly higher than the subject's behavioral levels in games. In some situations, cognitive bound was the same as the behavioral levels. Such cases were largely observed in the high cognitive load treatment.

To examine whether high cognitive load had a lower level distribution, I conducted a Wilcoxon signed-rank test on the estimated behavioral levels of games 17 (low load) and 18 (high load). Table 8 shows that the distributions of levels for the two treatments were significantly different at the 1% level. The one-tailed test indicated that the distribution of low load cognitive bound levels was to the right of the distribution of high load cognitive bound levels. This finding indicated that subjects had a higher cognitive bound when receiving low cognitive load treatment (memorizing a string of three letters) compared to receiving a high cognitive load treatment (memorizing a string of seven letters).

5.7. Robustness Check

During the guessing games, subjects needed to memorize a string of three or seven letters and recall the letters after they finished the guessing game. In this subsection, I present the results of this memorization task. Although the subjects were fully aware that if they failed to recall all the letters correctly, they would earn zero points for that round of the game, there were still some cases of wrong recalls due to reasons such as lack of attention or being too focused on the guessing game. I wanted to control the experimental results for such cases, as the subjects may have engaged in reasoning at higher levels when cognitive load did not fully apply. Table 14 shows the results of the memorization tasks. Most of the memorization tasks were perfectly performed. Not surprisingly, low cognitive load (three-letter memorization task) had more correct recalls, about 7% more than the high cognitive load task. The difference was significant at the 1% level.

Table 14. Results of the memorization task.

	3 Letters *** (Low Load)	7 Letters (High Load)	Total
Correct	97.30%	90.89%	94.09%
Wrong	2.70%	9.11%	5.91%
# of tasks	999	999	1998

∗ indicates < 10% significance, ∗∗ indicates < 5% significance, and *** indicates < 1% significance.

To check whether poor performance of the memorization task affected the treatment results, I excluded the data with wrong recalls and performed the analysis again. The comparison pair was dropped from the sample if either game of the pair had incorrect recalls. This was performed to ensure that the cognitive load was fully in effect, so that high cognitive load added difficulties to thinking through the guessing games at higher levels, and the cost of reasoning was higher.

Table 15 presents the treatment results after the robustness check. Treatments that involved high cognitive load had more data points dropped. For example, the [HH-] to [HH+] comparison had 444 pairs of comparison in the original sample. After robustness check, about 100 pairs were dropped. However, the results did not change much compared to the results presented in results 1 to 3 (Sections 5.3–5.5). The changes were mostly within 1%. I can therefore safely conclude that the original results were robust. The quality of the memorization task (i.e., whether the letters were correctly recalled) was almost independent of the treatment effects. Even in the cases of wrong recalls, the effect of cognitive load still applied to the subjects.

Table 15. Summary of the robust results with incorrect recalls dropped.

Pair Name (From Game a to Game b)	# of Pairs	Treatment	Decreases	Constant	Increases
Increasing Cost of Reasoning					
G16 to G9	107	LL- to HL-	20.56%	43.93%	35.51%
G7 to G14	77	LH- to HH-	40.26%	48.05%	11.69%
Combined	184	L?- to H?-	28.80%	45.65%	25.54%
Increasing Cost of Reasoning for Opponent					
(G1, G8) to (G4, G15)	426	LL+ to LH+	23.71%	56.10%	20.19%
(G6, G13) to (G3, G12)	392	HL+ to HH+	16.33%	42.09%	41.58%
Combined	818	?L+ to ?H+	20.17%	49.39%	30.44%
Second order belief: Low to (0.5 Low, 0.5 High)					
(G1, G8) to G16	214	LL+ to LL-	26.17%	50.93%	22.90%
(G4, G15) to G7	211	LH+ to LH-	21.33%	44.08%	34.60%
Combined	425	L?+ to L?-	23.76%	47.53%	28.71%
Second order belief: (0.5 Low, 0.5 High) to High					
G9 to (G6, G13)	202	HL- to HL+	35.64%	52.48%	11.88%
G14 to (G3, G12)	152	HH- to HH+	15.13%	45.39%	39.47%
Combined	354	H?- to H?+	26.84%	49.44%	23.73%

5.8. Cognitive Tests

In this subsection, I examine the results of the Mensa practice test. The test is composed of 10 questions and has a time limit of 10 min. Some subjects finished earlier, but they could never run overtime. Each correct answer is worth 1 point and all the unattempted questions are marked as 0 points. The score distribution of 104 subjects (seven missing) is presented in Table 16. There are a few very low points (2 or 3), and six subjects had scores of 10. Most subjects earned seven or eight points in this test.

Table 16. Summary statistics of cognitive test score and the counts of level changes following theory predictions.

	(1)	(2)	(3)	(4)	(5)	(6)
	Test Score	**Sum.Strict**	**Sum.Weak**	**Cost.Weak**	**1st.Weak**	**2nd.Weak**
Points possible	10	18	18	2	8	8
Max	10	13	18	2	8	8
Min	2	0	7	0	2	1
Median	7	4	13	2	6	6
Mean	7	4	13	1.5	5.5	6

To examine whether there are heterogeneous treatment effects in this experiment due to exogenous cognitive ability, I first determined a measure of the treatment effect. Out of all the results discussed in results 1 to 3 (Sections 5.3–5.5), there are in total 18 pairs of comparison. For each subject, I recorded one for the pair if the level change followed the theory prediction, and zero otherwise. As listed in Table 16, column Sum.Strict includes all the 18 comparisons, and only strict changes of levels are recognized. For example, if the pair game 16–game 9 had level 2 in both games, it is coded zero under Sum.Strict. However, column Sum.Weak allows weak changes; therefore, the above-mentioned scenario is coded as one under this column. The EDR model mostly discusses weak behavioral level changes because, in some cases, the changes in belief system or costs are not big enough to shift a behavioral level downwards by one level (evidenced by a large percentage of constant levels). Due to this reason, I considered "weak" changes, and decomposed them into columns (4) to (6), which cover the three main results. When limited to strict changes, a number of subjects had zero pairs following theory prediction (10 out of 111 subjects), and most subjects had only three or four pairs that had changes that could be predicted by the EDR model. However, when allowing weak changes, seven subjects had all the comparison pairs that were theory-predicted directional level changes, and most subjects had about 13 to 14 comparisons that could be predicted by the EDR model. The last three columns in Table 16 present results for each treatment separately.

To test whether cognitive ability had any correlation with the treatment effects, I ran a regression after dropping the subjects with missing test scores. The result is presented in Table 17. I used gender, class standing, and major as control variables. This information was collected at the end of the experiment. It appears that the cognitive test score and the female dummy variable were positively correlated with weak changes (at a 5% significance level), and the treatments changing the opponent's cost of reasoning (changing first-order belief) and changing second-order belief. The results showed some heterogeneous treatment effects in which the more cognitively capable subjects were more responsive to the treatments as predicted by the EDR model, especially in those requiring adjustments in response to the changing strategic environment of their opponents. When the strategic environment changed, these subjects were more likely to actively adjust their actions to gain possible strategic advantages.

Table 17. Regression results for cognitive test scores on correct directional changes of behavioral levels in block 1 games.

	Sum.Strict	Sum.Weak	Cost.Weak	First Order.Weak	Second Order.Weak
Test Score	0.03	0.36 **	−0.04	0.23 **	0.17 *
	(0.16)	(0.18)	(0.04)	(0.11)	(0.10)
Gender (F)	−0.73	1.43 **	0.03	0.83 **	0.57 *
	(0.57)	(0.60)	(0.13)	(0.37)	(0.34)
Class Standing	0.33	−0.18	−0.01	−0.13	−0.04
	(0.24)	(0.25)	(0.05)	(0.15)	(0.14)
Major	−0.02	0.00	−0.03	−0.03	0.01
	(0.12)	(0.13)	(0.03)	(0.08)	(0.07)
Constant	3.09 *	10.04 ***	1.76 ***	3.94 ***	4.34 ***
	(1.67)	(1.75)	(0.37)	(1.08)	(1.00)
# of Obs.	104	104	104	104	104

"Weak" includes constant levels and decreasing levels, while "strict" only includes strictly decreasing levels. * indicates < 10% significance, ** indicates < 5% significance, and *** indicates < 1% significance. Standard errors in parenthesis.

Since the result above suggested that more cognitively capable subjects' responses to changing strategic environment were more coherent with the EDR model, I separated the subjects into two groups according to cognitive test scores. Subjects with scores of eight or above were labeled as high cognitive subjects (high), and the remainder were labeled as low cognitive subjects (low). Table 18 presents results 1–3 again, separated by the cognitive test scores. As discussed in results 1 to 3 (Sections 5.3–5.5), I found significant asymmetries arising from the different strategic environments. Separating the subjects into two groups according to cognitive test scores allowed a closer examination of the source of the asymmetry. In Table 18, result 2 and result 3.1 highlight the relatively stable performance for the high cognitive subjects. As discussed in Section 5.4, subjects' responses to their opponents' changing cost of reasoning depended on their own cost of reasoning. In general, their adjustments in behavioral levels only followed the EDR model when they had a low cost of reasoning. This observation is untrue for the high cognitive subjects, who showed relatively stable performance regardless of their own strategic environment, with about 20% of the comparisons strictly following the EDR model. I observed a slight increase of 10% for those that did not follow the model; however, in general, the performance did not vary considerably. For the low cognitive subjects, the difference was huge. The 27.54% for comparison pairs that strictly followed the model decreased to 12.29%, and, more strikingly, the percentage of pairs that did not follow the model increased from 19.07% to 51.27%. This huge difference showed that the asymmetry found in the previous results was mostly due to these low cognitive subjects. There was a similar observation for result 3.1, where the high cognitive subjects had relatively stable performance regardless of their opponents' cognitive loads, whereas the low cognitive test score subjects were very sensitive to their opponents' strategic environments. Therefore, I concluded that the majority of asymmetric results found in results 2 and 3.1 were primarily driven by the low cognitive subjects. They were responsive to the treatments under the condition that they were in a more cognitively advanced situation. For results 1 and 3.2, both high and low cognitive subjects responded asymmetrically toward the treatment. However, as evidenced in Table 18, the changes from the low cognitive group were much greater than those of their counterparts.

Table 18. Results 1 to 3 separated by cognitive test scores.

			Changes in Behavioral Levels		
Treatment	**Cog Test**	**# of Pairs**	**Decreases**	**Constant**	**Increases**
Result 1: Increasing Cost of Reasoning					
LL- to HL-	High	45	22.22%	37.78%	40.00%
LL- to HL-	Low	59	16.95%	47.46%	35.59%
LH- to HH-	High	45	28.89%	55.56%	15.56%
LH- to HH-	Low	59	35.59%	57.63%	6.78%
Result 2: Increasing Cost of Reasoning for Opponent					
LL+ to LH+	High	180	19.44%	59.44%	21.11%
LL+ to LH+	Low	236	27.54%	53.39%	19.07%
HL+ to HH+	High	180	20.00%	47.22%	32.78%
HL+ to HH+	Low	236	12.29%	36.44%	51.27%
Result 3.1: Second order belief: Low to (0.5 Low, 0.5 High)					
LL+ to LL-	High	90	23.33%	53.33%	23.33%
LL+ to LL-	Low	118	28.81%	50.85%	20.34%
LH+ to LH-	High	90	20.00%	44.44%	35.56%
LH+ to LH-	Low	118	19.49%	42.37%	38.14%
Result 3.2: Second order belief: (0.5 Low, 0.5 High) to High					
HL- to HL+	High	90	35.56%	48.89%	15.56%
HL- to HL+	High	118	34.75%	55.08%	10.17%
HH- to HH+	High	90	21.11%	50.00%	28.89%
HH- to HH+	High	118	11.02%	40.68%	48.31%

The impact of cognitive ability on treatment effects was further evidenced by the regression results. In Table 19, the interaction term is significant for the comparison pairs that did not follow the EDR model ([HL+ to HH+], [LH+ to LH-], and [HH+ to HH-]). This implied that higher cognitive test scores skewed the effects of the treatment in the direction pointed by the EDR model. It seems that cognitive ability plays an important role for the subjects to display behavioral changes that can be predicted by the EDR model. The cognitive ability was captured endogenously by the treatment design in this experiment with two kinds of cognitive load. As discussed previously, the results differed systematically according to the amount of cognitive resources. Cognitive ability was also captured exogenously by the Mensa practice test, as discussed in this section. Within the asymmetric findings, subjects with higher cognitive test scores had more stable performance regardless of their own cognitive load, and were generally more predictable by the EDR model.

Table 19. Regression results for treatment effects and cognitive test scores on behavioral levels.

Comparison	Dummy	Score	Dummy ∗ Score	# Obs.
Changing Cost of Reasoning				
LL- to HL-	0.81 (0.79)	0.07 (0.07)	−0.06 (0.11)	208
LH- to HH-	−2.13 *** (0.73)	−0.11 (0.10)	0.19 * (0.10)	208
Changing Opponent's Cost of Reasoning				
LL+ to LH+	−1.15 (0.77)	−0.04 (0.09)	0.12 (0.10)	416
HL+ to HH+	1.67 *** (0.44)	0.08 (0.05)	−0.19 *** (0.06)	416
Changing Second Order Belief				
LL+ to LL-	−1.06 (0.75)	−0.04 (0.09)	0.11 (0.10)	312
LH+ to LH-	1.77 ** (0.80)	0.08 (0.05)	−0.19 * (0.11)	312
HL+ to HL-	1.08 (0.75)	0.08 (0.05)	−0.07 (0.10)	312
HH+ to HH-	−1.85 *** (0.37)	−0.11 *** (0.04)	0.19 *** (0.05)	312

* indicates < 10% significance, ** indicates < 5% significance, and *** indicates < 1% significance. Clustered individual standard errors in parenthesis.

6. Concluding Remarks

In this study, I designed a laboratory experiment to examine the consistency of players' strategic sophistication formulated by the level-k model. Following the endogenous depth of reasoning framework, I controlled the strategic environment by varying the cost of reasoning for the subjects, and their first- and second-order beliefs about their opponents.

My findings were consistent with the EDR model under some conditions. When the strategic environment was carefully controlled, subjects were very responsive towards the changes in the environment. Subjects who have more cognitive resources (in a low cognitive load treatment) or subjects who are facing opponents with less cognitive resources (in a high cognitive load treatment) change strategies systematically. This behavior can be predicted by the EDR model. Subjects in a strategically disadvantaged situation (high cognitive load treatment) have less room for strategic adjustments. In some of my findings, subjects appeared to try to achieve higher behavioral levels when they were under the high cognitive load treatment. The reason for this is still unclear. It may due to the awareness of the strategic disadvantage and the extra effort of the subjects under such situations, or some other behavioral factors existed that were not captured by the EDR model. The underlying reason needs further investigation. The effect of cognitive ability on the treatments was also captured by the cognitive test. Subjects with higher test scores were more predictable by the EDR model, regardless of the strategic environment. This finding is in line with the asymmetry observed in my results. As the source of asymmetry was mainly the amount of cognitive resources, it is not surprising that subjects with higher cognitive test scores adjusted better in these tasks.

A level of cognitive bound existed for subjects in different strategic situations. When playing games under the same amount of cognitive resources, subjects rarely had behavioral levels that exceeded their respective cognitive bounds for that strategic situation. Significant downward adjustments occurred from the cognitive bound in response to different strategic environments. Overall, when there is a strict control over the strategic environment, changes in k-levels across games are systematic. They can be explained by the EDR model to some extent, especially for subjects in a more cognitively advantaged situation. This study only discusses the directional changes in the levels. Further studies could examine the criteria and accuracy of such predictions.

Funding: This research received no external funding.

Acknowledgments: I would like to thank Charles Sprenger and Vincent Crawford for their valuable feedback and thoughtful suggestions. I am also grateful to the three referees and the editor for their helpful comments.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Nagel, R. Unraveling in Guessing Games: An Experimental Study. *Am. Econ. Rev.* **1995**, *85*, 1313–1326,
2. Stahl, D.O.; Wilson, P.W. Experimental evidence on players' models of other players. *J. Econ. Behav. Organ.* **1994**, *25*, 309–327. [CrossRef]
3. Camerer, C.; Ho, T.; Chong, J. A cognitive hierarchy model of games. *Q. J. Econ.* **2004**, *119*, 861–898. [CrossRef]
4. Costa-Gomes, M.; Crawford, V.P.; Broseta, B. Cognition and Behavior in Normal-Form Games: An Experimental Study. *Econometrica* **2001**, *69*, 1193–1235. [CrossRef]
5. Ho, T.H.; Cambrer, C.; Weigelt, K. Iterated Dominance and Iterated Best Response in Experimental "p-Beauty Contests". *Am. Econ. Rev.* **1998**, *88*, 947–969. [CrossRef]
6. Bosch-Domènech, A.; Montalvo, J.G.; Nagel, R.; Satorra, A. One, two, (three), infinity, ... : Newspaper and lab beauty-contest experiments. *Am. Econ. Rev.* **2002**, *92*, 1687–1701. [CrossRef]
7. Duffy, J.; Nagel, R. On the Robustness of Behaviour in Experimental 'Beauty Contest' Games. *Econ. J.* **1997**, *107*, 1684–1700. [CrossRef]
8. Grosskopf, B.; Nagel, R. The two-person beauty contest. *Games Econ. Behav.* **2008**, *62*, 93–99. [CrossRef]
9. Ho, T.H.; Su, X. A Dynamic Level-k Model in Sequential Games. *Manag. Sci.* **2013**, *59*, 452–469. [CrossRef]
10. Crawford, V.P.; Iriberri, N. Level-K Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions? *Econometrica* **2007**, *75*, 1721–1770,
11. Ivanov, A.; Levin, D.; Niederle, M.; Econometrica, S.; July, N. Can Relaxation of Beliefs Rationalize the Winner's Curse?: An Experimental Study. *Econometrica* **2010**, *78*, 1435–1452. [CrossRef]
12. Crawford, V.P.V.; Costa-Gomes, M.a.; Iriberri, N. Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications. *J. Econ. Lit.* **2013**, *51*, 5–62. [CrossRef]
13. Alaoui, L.; Penta, A. Endogenous Depth of Reasoning. *Rev. Econ. Stud.* **2016**, *83*, 1297–1333. [CrossRef]
14. Costa-Gomes, M.A.; Crawford, V.P. Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. *Am. Econ. Rev.* **2006**, *96*, 1737–1768. [CrossRef]
15. Arad, A.; Rubinstein, A. The 11–20 Money Request Game: A Level- k Reasoning Study. *Am. Econ. Rev.* **2012**, *102*, 3561–3573. [CrossRef]
16. Allred, S.; Duffy, S.; Smith, J. Cognitive load and strategic sophistication. *J. Econ. Behav. Organ.* **2016**, *125*, 162–178. [CrossRef]
17. Burnham, T.C.; Cesarini, D.; Johannesson, M.; Lichtenstein, P.; Wallace, B. Higher cognitive ability is associated with lower entries in a p-beauty contest. *J. Econ. Behav. Organ.* **2009**, *72*, 171–175. [CrossRef]
18. Gill, D.; Prowse, V. Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *J. Political Econ.* **2016**, *124*, 1619–1676. [CrossRef]
19. Stahl, D.O.; Wilson, P.W. On Players? Models of Other Players: Theory and Experimental Evidence. *Games Econ. Behav.* **1995**, *10*, 218–254. [CrossRef]
20. Fragiadakis, D.E.; Knoepfle, D.T.; Niederle, M. Who is Strategic? Technical Report. Unpublished work, 2016.
21. Arad, A.; Rubinstein, A. Multi-dimensional iterative reasoning in action: The case of the Colonel Blotto game. *J. Econ. Behav. Organ.* **2012**, *84*, 571–585. [CrossRef]
22. Georganas, S.; Healy, P.J.; Weber, R.A. On the persistence of strategic sophistication. *J. Econ. Theory* **2015**, *159*, 369–400. [CrossRef]
23. Fischbacher, U. Z-Tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* **2007**, *10*, 171–178. [CrossRef]
24. Agranov, M.; Caplin, A.; Tergiman, C. Naive play and the process of choice in guessing games. *J. Econ. Sci. Assoc.* **2015**, *1*, 146–157. [CrossRef]

games

MDPI

Article

Estimating Case-Based Learning

Todd Guilfoos [1,*] and Andreas Duus Pape [2]

1 Environmental and Natural Resource Economics, University of Rhode Island, Kingston, RI 02881, USA
2 Department of Economics, Binghamton University, Binghamton, NY 13902, USA; apape@binghamton.edu
* Correspondence: guilfoos@uri.edu

Received: 16 July 2020; Accepted: 7 September 2020; Published: 15 September 2020

Abstract: We propose a framework in order to econometrically estimate case-based learning and apply it to empirical data from twelve 2×2 mixed strategy equilibria experiments. Case-based learning allows agents to explicitly incorporate information available to the experimental subjects in a simple, compact, and arguably natural way. We compare the estimates of case-based learning to other learning models (reinforcement learning and self-tuned experience weighted attraction learning) while using in-sample and out-of-sample measures. We find evidence that case-based learning explains these data better than the other models based on both in-sample and out-of-sample measures. Additionally, the case-based specification estimates how factors determine the salience of past experiences for the agents. We find that, in constant sum games, opposing players' behavior is more important than recency and, in non-constant sum games, the reverse is true.

Keywords: learning; behavioral game theory; case-based decision theory

JEL Classification: D01; D83; C63; C72; C88

1. Introduction

Economists across the discipline—micro and macro, theory and empirics—study the impact of learning on individual and social behavior. Two questions are typical of this inquiry: first, whether and when learning leads to equilibrium behavior, and second, which model(s) of learning best explain the data. In this paper, we formulate a method to econometrically estimate Case-based Decision Theory (CBDT), introduced by Gilboa and Schmeidler [1], on individual choice data.

Like Expected Utility (EU), CBDT is a decision theory: that is, it shows that if an agent's choice behavior follows certain axioms, it can be rationalized with a particular mathematical representation of utility e.g., Von Neumann and Morgenstern [2], Savage [3]. The Expected Utility framework has states of the world, actions, and payoffs/outcomes. The CBDT framework retains actions and payoffs, but it replaces the set of states with a set of "problems", or circumstances; essentially, vectors of information that describe the choice setting the agent faces. CBDT postulates that when an agent is confronted with a new problem, she asks herself: how similar is today's problem to problems in memory? She then uses those similarity-weighted problems to construct a forecasted payoff for each action, and chooses an action with the highest forecasted payoff.

The primary motivation for our study is to estimate and measure the efficacy of CBDT to explain learning. Therefore, in this context, we refer to Case-based Learning or CBL. We develop a framework to estimate dynamic case-based decision theory econometrically and test it in a game-theoretic setting against other learning models. One significant difference between CBL and other learning models is the formulation of how information enters into decision-making. In CBL, information enters in how agents perceive past experiences to be salient to current choice. To do this, CBL incorporates psychological similarity.

An important part of this work is using a stochastic choice rule to estimate CBDT. CBDT is a deterministic theory of choice, but, in this study, we transform it into stochastic choice. The primary purpose of this transformation is estimating parameters of models on data, like much of the literature in learning algorithms that we compare CBDT against. However, it is worth noting that there is precedent in the literature to treat CBDT specifically as stochastic, e.g., Pape and Kurtz [4] and Guilfoos and Pape [5] use stochastic forgetfulness in their implementations to match human data. Moreover, there is a broader tradition in psychology of converting deterministic utility valuations into stochastic choice through the so-called Luce choice rule or Luce choice axiom [6] (see Section 3.6).

We test CBL and other learning models on data from a series of 2 $\times$ 2 experimental mixed strategy equilibria games. Erev and Roth [7] make an explicit case for the use of unique mixed strategy equilibrium games to investigate learning models, in part because the number of equilibria does not change with finite repetitions of the game and the equilibrium can be achieved in the stage game. Given the simplicity of the information available to subjects, these data provide a relatively conservative environment for a researcher to test CBL, as it restricts the degrees of freedom to the researcher. In an experiment, the information available to subjects is tightly controlled, so a well-defined experiment provides a natural definition of the problem vector for CBDT. We estimate parameters of the learning algorithms to understand how parameters change under different contexts, and because they provide information about the nature of choice. A benefit of estimating parameters of CBL is to compare how stable the parameters remain under different contexts. The data we use are well-studied by researchers investigating stationarity concepts and learning models [8,9].

We find that CBL explains these empirical data well. We show that CBL outperforms other learning algorithms on aggregate on in-sample and out-of-sample measures. Reinforcement learning and CBL perform similarly across individual games and they have similar predictions across games. This is also supported by our analysis of the overlap in RL and CBL in attraction dynamics when certain restrictions are made. When learning models outperform the known equilibria or stationary concepts (Nash Equilibrium, action-sampling equilibrium, payoff-sampling equilibrium, and impulse balance equilibrium) it prompts the question of which learning models characterize the data well and what insights are gained through learning models into decision making behavior.[1] For instance, it is known that some of the learning models in games do not converge to Nash Equilibrium and then we must consider what is it converging to, if anything, and how is it converging.

Our econometric framework for CBL provides estimates that measure the relative importance for each piece of information available to subjects and the joint significance of information in predicting individual choice; this can be interpreted as estimates of the salience of past experiences for the agents. We find that both recency and opposing players' behavior are jointly important in determining salience. We also find that in constant sum games, the behavior of opposing players is more important than recency, while, in non-constant sum games, recency is more important. The relative importance (as revealed by the relative weights) provides new insight into how subjects respond to stimuli in mixed strategy games, and provides a new piece of empirical data for future theory models to explain and understand. This points toward future work, in which more studies interact learning models with available information to identify how learning occurs in and across games.

We compare CBL to two learning models from the literature: Reinforcement Learning [7]; and self-tuning Experience Weighted Attraction [10]. Reinforcement Learning (RL) directly posits that individuals will exhibit behavior that in the past has garnered relatively high payoffs. Self-tuning Experience Weighted Attraction (self-tuning EWA) is a model that allows for the learners to incorporate aspects of reinforcement learning and belief learning. Both have achieved empirical success in explaining experimental game play; in particular, these two were the most successful

[1] The learning models from Chmura et al. [9] establish the fit of these stationary concepts and other learning models provide a worse fit of the data than the models considered here. We do replicate the findings for self-tuning EWA and find a better fit for reinforcement learning by estimating a greater number of free parameters.

learning models tested in Chmura et al. [9], whose data we analyze. We describe these models in greater detail in Section 3. We also formally investigate the relationship between CBL and RL; we show there is a mapping between RL and CBL when particular assumptions are imposed on both. Relaxing these assumptions is informative in understanding how the algorithms relate.

There is a small but persuasive literature evaluating the empirical success of CBDT. It has been used to explain human choice behavior in a variety of settings in and outside the lab. There are three classes of empirical studies. The first class uses a similarity function as a static model, which ignores dynamics and learning [11,12]. The second class is dynamic, but it utilizes simulations to show that case-based models match population dynamics rather than econometric techniques to find parameters [4,5]. The third class is experimental investigations of different aspects of case-based decision-making [13–18]. Our study is unique in that it proposes a stochastic choice framework to estimate a dynamic case-based decision process on game theoretic observations from the lab. Further, we relate this estimator to the learning and behavioral game theory literature and demonstrate the way in which case-based learning is different.

Neuroeconomic mechanisms also suggest that CBDT is consistent with how past cases are encoded and used in order to make connections between cases when a decision-makers faces a new situation [19]. Neuroeconomics is also in agreement with many other learning models. It is hypothesized by Gayer and Gilboa [20] that, in simple games, case-based reasoning is more likely to be discarded in favor of rule-based reasoning, but case-based reasoning is likely to remain in complex games. CBDT is related to the learning model of Bordalo et al. [21], which uses a similarity measure to determine which past experiences are recalled from memory. This is related to CBDT: in Bordalo et al. [21], experience recall is driven by similarity, while, in CBDT, how significant an experience weighs in utility is driven by similarity. Argenziano and Gilboa [22] develop a similarity-based Nash Equilibria, in which the selection of actions is based on actions that would have performed best had it been used in the past. While the similarity-based equilibria are closely related to this work, our case-based learning is not an equilibrium concept. Our work builds on the empirical design developed in the applied papers as well as those developing empirical and functional tools related to CBL e.g., [23,24].

2. Applying Case-Based Learning to Experiments

First, we compare the case-based approach to traditional expected utility. The expected utility framework requires that the set of possible states is known to the decision-maker and that the decision-maker has a belief distribution over this set of states. Case-based decision theory replaces the state space and its corresponding belief distribution with a "problem" space—a space of possible circumstances that the decision-maker might encounter—and a similarity function defined over pairs of problems (circumstances). One limitation of the expected utility approach is that it is not well-defined for the decision-maker to encounter a truly "new" state, which is, a state the decision-maker had never thought of before (it could be modeled as a state that occurs with probability zero, but then Bayesian updating would leave it at probability zero). The case-based approach overcomes this difficulty: the decision-maker can naturally encounter a "new" problem or circumstance, and need only be able to judge how similar that problem is to other problems the decision-maker has encountered: no ex-ante determination is required.[2] The problem space is also, arguably, more intuitive for many practical decision-making problems than the corresponding state space. For example, consider the problem of hiring a new assistant professor, where one's payoff includes the success of this candidate, fit with the current department, willingness of the candidate to stay, etc. Describing each candidate as a vector of characteristics that can be judged more or less similar is fairly intuitive, while constructing

[2] It is worth noting that there is a mapping between expected utility and case-based decision theory [25], which implies that in a formal sense replacing the state space with the problem space is not 'easier,' if one requires that the decision-maker must *ex-ante* judge the similarity between all possible pairs of problems.

a corresponding state space—possible maps of candidates to payoff-relevant variables—may not be. Reasoning by analogy, through similarity, can also make complex decisions more manageable. Moreover, the similarity between vectors of characteristics provides a specific means of extrapolating learning about one candidate to other candidates; the assumption of prior distributions, and updating those distributions, provides less guidance about how that extrapolation should be done.

In our setting multiple games are played in a laboratory in which players interact with each other to determine outcomes. The state space in this setting is large. The broadest interpretation of the appropriate state space is: the set of all possible maps from all possible histories of play with all opponents to all future play. While this is quite general, learning (that is, extrapolation from past events to future ones) requires the specification of a well-informed prior; if literally any path of play is possible given history, and if one had a diffuse (i.e., "uninformed") prior over that set, then any future path is equally likely after every possible path of play. Alternatively, the state space might assume a limited set of possible player types or strategies; in that case, the state space would be all possible mappings of player types/strategies to players. While this provides more structure to learning, it requires that the (correct) set of possible player types is known.

On the other hand, defining an information vector about history is less open-ended. There are natural things to include in such a vector: the identity (if known) of the player encountered; the past play of opponents, a time when each action/play occurred; and, perhaps other features, such as social distance [26] or even personality traits [27]. This implies a kind of learning/extrapolation in which the behavior of player A is considered to be more relevant to predictions of A's future behavior than is the behavior of some other player B; that if two players behave in a similar way in the past; that learning about one player is useful for predicting the play of another; and, that more recent events are more important than ones far in the past. These implications for learning naturally arise from a similarity function that considers vectors closer in Euclidean distance to be more similar (as we do here).[3] Interestingly, others have adopted the concept of similarity as a basis of choice in cognitive choice models [28].[4]

Note that this kind of extrapolation can be constructed in a setting with priors over a state space, under particular joint assumptions over the prior over the state space and the state space itself. Case-based decision theory can be thought of as a particular set of testable joint assumptions that may or may not be true for predicting human behavior.

3. Learning Algorithms

Learning models in economics have served dual purposes. First, learning algorithms can play a theoretical role as a model of dynamics which converge to equilibrium. This is the explicit goal of the "belief learning" model [29]. Second, learning algorithms can play an empirical role in explaining the observed dynamics of game play over time. This goal is explicit in the "reinforcement learning" model [7] which draws heavily on models from artificial intelligence and psychology.

Both purposes are incorporated in the Experience Weighted Attraction model, which, appropriately enough, explicitly incorporates the belief learning and reinforcement learning models [30–32]. EWA and its one parameter successor, self-tuning EWA, has proved to be a particularly successful account of human experimental game play. Here, we discuss these reinforcement and self-tuning EWA models and compare and contrast them to the case-based learning approach.

In the following repeated games, we assume the same following notation: there are a set of agents indexed by $i = 1, \ldots, n$, each with a strategy set S_i, which consists of m_i discrete choices,

[3] But see Section 7.2. In fact, Erev and Roth [7] discuss such a similarity between situations in which to define experimentation of a subject when choosing strategies.

[4] Similarity is used in a way that maps closely to how learning models work, in general, by repeating successful choices under certain conditions. Choices in Cerigioni [28] use similarity when automated through the dual decision processes familiar from psychology.

so that $S_i = \{s_i^1, \ldots, s_i^{m_i}\}$. Strategies are indexed with j (e.g., s_i^j). Let $s = (s_1, s_2, \ldots, s_n)$ be a strategy profile, one for each agent; in typical notation, s_{-i} denotes the strategy profile with agent i excluded, so $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$. Scalar payoffs for player i are denoted with the function $\pi_i(s_i, s_{-i})$. Finally, let $s_i(t)$ denote agent i's strategy choice at time t, so $s_{-i}(t)$ is the strategy choices of all other agents at time t.

Erev and Roth [7] argue that, empirically, behavior in experimental game theory appears to be probabilistic, not deterministic. Instead of recommending deterministic choices, these models offer what the EWA approach has come to call "attractions." An attraction of an agent i to strategy j at time t is a scalar which corresponds to the likelihood that agent will choose this strategy at this time relative to other strategies available to this agent. An attraction by agent i to strategy j at time t under an arbitrary learning model will be represented by $A_i^j(t)$. We compare these models by saying that different models provide different functions which generate these attractions, so we will have, e.g., $CBA_i^j(t)$ to represent the attraction that is generated by the case-based model.

Because a given attraction corresponds to a likelihood, a vector of attractions $\left\{A_i^j(t)\right\}_{j=1}^{m_i}$ corresponds to a probability distribution over available choices at time t and, therefore, fully describes how this agent will choose at time t.[5]

We consider case-based learning (CBL), reinforcement learning (RL), and self-tuning experience weighted attraction (EWA), in turn.

3.1. Case-Based Learning

We bring a formulation of case-based decision theory as introduced by Gilboa and Schmeidler [1] into the "attraction" notation discussed above, ultimately ending up with a case-based attraction $CBA_i^j(t)$ for each strategy s_j.

The primitives of Case-Based Decision Theory are: a finite set of actions $\mathcal{A}$ with typical element a, a finite set of problems $\mathcal{P}$ with typical element p, and a set of results $\mathcal{R}$ with typical element r. The set of acts is of course the same as the set of actions or strategies as one would find in a typical game theoretic set-up. The set of problems can be thought of a set of *circumstances* that the agent might face: or, more precisely, a vector of relevant information about the present circumstances surrounding the choice that the agent faces, such as current weather, time of day, or presence of others. The results are simply the prizes or outcomes that result from the choice.

A problem/action/result triplet (p, a, r) is called a *case* and can be thought of as a complete learning experience. The set of cases is $\mathcal{C} = \mathcal{P} \times \mathcal{A} \times \mathcal{R}$. Each agent is endowed with a set $\mathcal{M} \subseteq \mathcal{C}$, which is called the memory of that agent. Typically, the memory represents those cases that the agent has directly experienced (which is how it is used here) but the memory could be populated with cases from another source, such another agent or a public information source.

Each agent is also endowed with a similarity function $s : \mathcal{P} \times \mathcal{P} \to \mathbb{R}_+$, which represents how similar two problems are in the mind of the agent. The agent also has a utility function $u : \mathcal{R} \to \mathbb{R}$ and a reference level of utility H, which is called the aspiration value. An aspiration value is a kind of reference point. It is the desired level of utility for the agent; when the agent achieves her aspiration value of utility, she is satisfied with that choice and is not moved to seek alternatives.

When an agent is presented with problem p, the agent constructs the case-based utility for each available action $a \in \mathcal{A}$ and selects an action with the highest CBU.[6] CBU is constructed from memory in Equation (1).

---

5 We discuss the functional form of the probability distribution in Section 3.6.

6 As discussed in the introduction of this section, our implementation uses attractions, so choice is not deterministic, but rather stochastic with the probability of choosing an action increasing in the CBU.

$$CBU(p,a) = \sum_{(q,a,r)\in\mathcal{M}(a)} s(p,q)\,[u(r) - H] \quad (1)$$

where $\mathcal{M}(a)$ is defined as the subset of memory in which act a was chosen; that is $\mathcal{M}(a) = \{(q,a,r)|(q,a,r) \in \mathcal{M}\}$. (Following Gilboa and Schmeidler [1], if $\mathcal{M}(a)$ is the empty set—that is, if act a does not appear anywhere in memory—then $CBU(p,a)$ is assumed to equal zero.)

The interpretation of case-based utility is that, to form a forecast of the value of choosing act a, the agent recalls all those cases in which she chose action a. That typical corresponding case is called (q,a,r). The value associated with that case is the similarity $s(p,q)$ between that case's problem q and her current problem p, times the utility value of the result of that decision, minus the aspiration value H. Subsequently, her total forecast is the sum of those values across the entire available memory.

Now, let us bring the theory of case-based learning to an empirical strategy for estimating case-based learning in these experiments. Note that the experiments studied here—2×2 games with information about one's history—provide an environment for testing the theory of case-based learning, because the information vectors $\mathcal{P}$ presented to subjects is well-understood and controlled by the experimenter. (Outside the lab, more and stronger assumptions may be required to define $\mathcal{P}$).

3.1.1. Definition of Case-Based Attraction

CBL is defined by Equation (2). $CBA_i^j(t)$ is the case-based attraction of agent i to strategy j at time t; as discussed above, an attraction corresponds to the probability of selecting a strategy j. Here we present the equation and discuss each component in turn:

$$CBA_i^j(t) = A_0^j + \sum_{m=\max(t-M,0)}^{t-1} I(s_i^j, s_i(m)) \cdot S(x_t, x_m) \cdot [\pi(s_i(m)) - H] \quad (2)$$

The first term, A_0^j, is a taste parameter for strategy j. On the first instance of play, the second term is zero (we will explain below), so A_0^j also equals the initial attraction to strategy j. On the first instance of play, there are no prior cases to inform the experimenter of the subject's preferences, so it might be natural to assume that the agent ought to be indifferent among all actions, which would suggest that agents ought to choose all actions with equal probability in the first round. This does not appear to be the case in the data, hence the inclusion of this taste parameter (if initial actions are selected with equal probabilities, then these taste parameters will be estimated to be equal).

Now, let us consider the second term:

$$\sum_{m=\max(t-M,0)}^{t-1} I(s_i^j, s_i(m)) \cdot S(x_t, x_m) \cdot [\pi(s_i(m)) - H]$$

The variable M the (maximum) length of memory considered by the agent. The first case considered by the agent is listed as $m = \max(t-M, 0)$. This has a straight-forward interpretation: either considered memory begins at period 0, which is the beginning, or, if $t > M$ (and, therefore, $t - M > 0$), then only the last M periods are considered in memory. For example, if $M = 3$, then every utility calculation only considers the last three periods. If all experiences are included in memory then M is equal to ∞. We test the importance of the choice of M in the Section 6.

$I(s_i^j, s_i(m))$ is an indicator function that maps cases in memory to the appropriate attraction for the strategy chosen: that is, when the strategy chosen, $s_i(m)$, is equal to strategy s_i^j, then this function equals one and it contributes to the attraction for strategy s_i^j. Otherwise, this function is zero and it does not contribute.

$S(x_t, x_m)$ is the similarity function, which translates the elements of the problem into relevance: the greater the similarity value, the more relevant problem x_m is to problem x_t to the decision-maker.

$[\pi(s_i(m)) - H]$ is the payoff in memory, net the aspiration level, so results that exceed aspirations are positive and results that fall short of aspirations are negative.

3.1.2. The Functional Form of Similarity

We give a specific functional form to the similarity function in Equations (3) and (4), where x and y denote two different problem vectors. We choose an inverse exponential function that uses weighted Euclidean distance between the elements of the circumstances to measure the similarity of situations. This choice has support from the psychology literature [33]. Specifically, the information that individuals encounter in past experiences and can observe in the current case are compared through the similarity function. The more similar the current case to the past case, the greater weight the past case is given in the formulation of utility. (We explore other functional forms of similarity and distance between information vectors in Section 7.2.).

$$S(x,y) = \frac{1}{e^{d(x,y)}} \tag{3}$$

$$\text{where } d(x,y) = \sqrt{\sum_{i=1}^{\#Dims} w_i[(x_i - y_i)^2]} \tag{4}$$

$$\text{for some weights } w_i.$$

3.1.3. Comparision to RL And EWA

In RL and self-tuning EWA, attractions at time t are a function of attractions at time $t-1$, and attractions explicitly grow when the strategies they correspond to are valuable to the agent, a process called 'accumulation'. Note that CBL does not explicitly accumulate attractions in this way, and has no built-in depreciation or accumulation factor such as ϕ or $\frac{N(t)}{N(t-1)}$. However, closer inspection suggests that CBL implicitly accumulates attractions through how it handles cases in memory: as new cases enter memory, when payoffs exceed the aspiration level, they increase the attraction of the corresponding strategy. This appears to function as explicit accumulation does in the other theories. One important difference between the case-based approach to accumulation and the RL/EWA method is dynamic re-weighting: that is, when the current problem (information vector) changes, the *entire memory* is re-weighted by the corresponding similarity values. There is accumulation of a sort, but that accumulation is information-vector dependent. Accumulation through similarity allows for CBL to re-calibrate attractions to strategies that are based on information in the current and past problem sets.[7]

Depreciation can also be modeled in a natural way in CBL: if time is a characteristic in the information vector, then cases further in the past automatically play a diminishing role in current utility forecasts as they become more dissimilar to the present.

3.2. Reinforcement Learning

Reinforcement learning (RL) has origins in psychology and artificial intelligence and it is used in many fields, including neuroscience [34,35]. This is the formulation that we use here:

Consider a vector of attractions $\left\{RLA_i^j\right\}_{j=1}^{m_i}$. Suppose that strategy s^j is chosen, the payoff experienced by agent i is added to RLA_i^j. In this way, strategies that turn out well (have a high payoff) have their attraction increased, so they are played more likely in the future. After strategy profile $s(t)$ is chosen at time t and payoffs are awarded, the new vector of attractions is:

[7] Moreover, the similarity function can also be dynamic, which further allows for reconsideration of past events in a way RL/EWA accumulation does not.

$$RLA_i^j(t) = \phi RLA_i^j(t-1) + I\left(s_i^j(t), s_i^j\right) \cdot \pi_i(s_i, s_{-i}(t)) \qquad \forall j = 1, \ldots, m_i \tag{5}$$

This is the same basic model of accumulated attractions as proposed by Harley [36] and Roth and Erev [37]. The first term within the brackets, $\phi RLA_i^j(t-1)$, captures the waning influence of past attractions. For all attractions other than the one corresponding to the selected strategy $s(t)$ attractions tend toward zero, assuming that the single global factor is not too large. $I(s_i^j, s_i(t))$ will be used to denote the indicator function which equals 1 when $s_i^j = s_i(t)$ and 0 otherwise. The one countervailing force is the payoff π_i, which only plays a role in the selected strategy (as indicated by the indicator function). This version of reinforcement learning is a cumulative weighted RL, since payoffs accumulate in the attractions to chosen strategies. In a simplified setting where payoffs are weakly positive, this process can be thought of as a set of leaky buckets (with leak rate $(1-\phi)$), one corresponding to each strategy, in which more water is poured into buckets corresponding to the strategy chosen in proportion to the size of the payoff received.[8] Subsequently, a strategy is chosen with a probability that corresponds to the amount of water in its bucket.

There are simpler forms of RL that can be used, where ϕ is equal to 1, and not estimated. This simpler model is estimated in Chmura et al. [9] and performs worse than our modified model in explaining the data. We use Equation (5) to fit the data to make conservative comparisons to the CBL model.

3.3. Self-Tuning Experience Weighted Attraction

Self-tuning experience weighted attraction was developed by Ho et al. [10] to encompass experience weighted attraction [31] in a simple one parameter model. EWA incorporates both RL and belief learning, which relies on so-called "fictitious play", in which the payoffs of forgone strategies are weighted alongside realized payoffs. Self-tuning EWA is a compact and flexible way to incorporate different types of learning in one algorithm.

Equation (6) describes self-tuning EWA: a δ weight is placed on fictitious play and a $(1-\delta)$ weight is placed on realized outcomes. Self-tuning EWA has been successful at explaining game play in a number of different settings, including the data that we use in this paper.

$$EWA_i^j(t) = \frac{N(t-1)\phi(t)EWA_i^j(t-1) + \left[\delta + (1-\delta)\cdot I(s_i^j, s_i(t))\right]\cdot \pi(s_i^j, s_{-i}(t))}{N(t)} \tag{6}$$

In self-tuning EWA, the parameter N evolves by the rule $N(t) = \phi \cdot N(t-1) + 1$ and $N(0) = 1$. The $I(\cdot)$ function is an indicator function that takes a value of 1 when $s(t) = s_i$ and 0 otherwise. The parameter ϕ acts as a discount on past experiences, which represents either agents forgetfulness or incorporating a belief that conditions of the game may be changing. This parameter evolves, so that $\phi(t) = 1 - \frac{1}{2}Sp(t)$, where $Sp(t)$ is a surprise index. $Sp(t)$ measures the extent to which agent's partners deviate from previous play. More precisely, it is defined by the cumulative history of play $h_j^k(t)$ and a vector of the most recent play $r_j^k(t)$ for strategy j and opposing player k, as given in the Equations (7) and (8).

$$h_j^k(t) = \frac{\sum_{\tau=1}^{t} I(s_j^k, s^k(\tau))}{t} \tag{7}$$

$$r_j^k(t) = \sum_{j=1}^{2} \frac{\sum_{\tau=t-W+1}^{t} I(s_j^k, s^k(\tau))}{W} \tag{8}$$

[8] The bucket analogy is also apropos because Erev and Roth [7] describe a spillover effect, in which buckets can slosh over to neighboring buckets. We do not investigate the spillover effect in this paper, since with only two actions (in 2×2 games) the spillover effect washes out.

$W = 2$ because there are only two strategies available to all agents in these games. In the experiments used in this paper, the subjects are unable to identify opposing players and we treat all of the opposing players as a representative average player, following Chmura et al. [9], to define histories and the surprise index. Equation (9) defines the surprise index, which is the quadratic distance between cumulative and immediate histories.

$$Sp(t) = \sum_{j=1}^{2} (r_j^k(t) - h_j^k(t))^2 \tag{9}$$

The fictitious play coefficient δ shifts attention to the high payoff strategy. This function takes the value of $\delta = \frac{1}{W}$ if $\pi(s_i, s_{-i}(t)) > \pi(t)$ and 0 otherwise.

3.4. Relationship between RL and CBL

There is a strong connection between case-based learning and other learning algorithms, particularly reinforcement learning. One way to illustrate the connection between RL and CBL is by constraining both RL and CBL in particular ways, so that they become instances of each other. Subsequently, we can consider the implications of relaxing these constraints and allowing them to differ.

On CBL, we impose three restrictive assumptions: first, we constrain the information vector to include only time (so that the only aspect of situations/problems that the case-based learner uses to judge similarity is how close in time they occurred). Second, we set the aspiration level to zero, so that payoffs are reinforced equivalently in RL and CBL. Third, we assume the similarity function is of the form in Equation (10).

$$S(x_t, x_m) = \frac{1}{w^{|t-m|}}. \tag{10}$$

(Note, again, that x_t is a 'vector', which consists only of t).

Finally, on both, we impose the assumption that initial attractions to be zero for both CBL and RL, which means that, in both cases, choices are randomized in the initial period.

We can then derive the weight in similarity that leads to the same decay in attractions in both RL and CBL, as displayed in Equation (11).

$$\phi = S(x_t, x_{t-1}) = \frac{1}{w} \tag{11}$$

Under the assumptions on RL and CBL listed above, if one estimates the RL equation (Equation (5)) and then estimates the CBL equation (Equation (2)) on the same data, and then resulting estimators ϕ and w are necessarily related in the way described in Equation (11). We do not use these specialized forms to estimate against the data, but rather use them to demonstrate the simple similarities and differences in how CBL and RL are constructed. In Appendix A, we provide more details on the formal relationship between RL and CBL.

This is a base case, where RL and CBL are the same. Now, let us consider two complications relative to this base case and consider the implications for the different attractions.

First, let us allow for more variables in the information vector (in addition to time) and consider how this would change the CBL agent relative to the RL/base case agent. Adding more variables to the information vector can be thought of as a CBL agent being able to maintain multiple 'rates of decay', which could vary over time, where the CBL agent can choose which 'rate of decay' to use based on the current situation. For example, suppose opponent ID is included in the information vector. Subsequently, if the agent is playing a partner they encountered two periods ago, the CBL agent could choose to downweight the previous period's attraction and increase the weight given to the problem from two periods ago. In essence, this additional information, and combination of weights in the definition of distance, allows for the ϕ parameter to be 'recast' based on the memory of an agent and the current problem. The modification of reinforcement learning to include the recasting is an elegant way to incorporate the multiple dimensions of information agents use when playing games. It suggests that other empirical applications in discrete choice may also benefit in using CBL, because it contains core elements of reinforcement learning that have been successful in modeling behavior.

Second, let us consider an aspiration level that differs from zero. Suppose payoffs $\pi \geq 0$, as they are in the games that we consider here. Subsequently, under RL, and under CBL with $H = 0$, every experience acts as an attractor: that is, it adds probability weight to a particular action, the question is: how much probability weight does it add. However, when $H > 0$, then the change in attraction of an action is does not increase in π, but rather in $\pi - H$. This, importantly, changes the implications for attraction for payoffs that fall short of H. Under CBDT, such payoffs provide a "detractor" to that action, so they directly lower the attraction corresponding to this action.

3.5. Initial Attractions

We estimate initial attractions to strategies for all learning models. This adds two additional parameters to estimate for all models, for the row player the initial attraction to strategy Up and for the column player the initial attraction to strategy Left. This seems sensible, because, empirically, it does not appear that subjects choose strategies randomly in the first period of play, and, *a priori*, there is a systematic difference between payoffs when considering the expected play by the opposing player. We can compare the actions in the first round of the experimental data to the estimated initial attractions as a sensible test of the learning model. We fit all learning models using the stochastic choice rule and appropriate learning theory and then predict the choice for each period (round) and subject in the dataset.

3.6. Stochastic Choice Probabilities

As defined in Sections 3.1–3.3, each learning model generates a set of attractions for each strategy j: $RLA_i^j(t)$, $EWA_i^j(t)$, and $CBA_i^j(t)$. We use the same function to aggregate the attractions generated by these different models. That function is a logit response rule. Let $A_i^j(t)$ be any of these three attractions. Subsequently, Equation (12) gives the probabilities that the attractions yield:

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}} \tag{12}$$

Logit response has been expansively used in the learning literature of stochastic choice and, if the exponetial of the attractions are interpretes "choice intensities", this formulation is consistent with the Luce Choice Rule [6], as discussed in the introduction.[9] Equation (12) is used as the stochastic choice rule to fit data to the models to explain each individual choice j by each subject i in every time period t. The learning algorithm equations will be estimated using maximum likelihood in order to determine the fit of the each of the models and provide estimates of the specific learning parameters. This includes experimenting with various initial parameters and algorithms.[10]

In this logit rule, λ is the sensitivity of response to the attractions, where a low value of λ would suggest that choices are made randomly and a high value of λ would suggest that the choices determined by the attractions. This value will be estimated with the empirical data and could vary for a variety of reasons, such as the subject's motivation in the game or unobserved components of payoffs.

9 In addition to logit response, we also estimate a power logit function, but find that it does not change the conclusions or generally improve the fit of the learning models estimated here.

10 We use STATA to estimate the maximium likelihood functions using variations of Newton–Raphson and Davidon–Fletcher–Powell algorithms, depending on success in estimation. Code is available upon request.

4. Description of the Data

All of the games investigated are of the 2 × 2 form, as shown in Figure 1. The experiments from these games were collected by Selten and Chmura [8] and discussed in Chmura et al. [9]. Chmura et al. [9] investigate a series of learning models and determine which rules characterize individual and aggregate performance better. They find that self-tuning EWA fit the data best yet impulse-matching learning also fit the data well. The twelve 2 × 2 games include both constant sum and non-constant sum games.

The experiments were performed at the Bonn Lab with 54 sessions and 16 subjects in each session. 864 subjects participated in the experiment. The subjects in a given session were only exposed to one game. Games were 200 rounds long and subjects were randomly matched by round in groups of eight during the sessions. Knowledge of the game structure, payoffs, and matching protocols were public at the outset. The subject's role of row or column player are fixed during the experiment, so four subjects in the group of eight were assigned to be a row player and the others were assigned to be a column player. At the end of each round, the subjects were told their current round payout, the other player's choice, the round number, and their cumulative payout. The experiments lasted between 1.5 and 2 h and subjects received at 5 Euro show-up fee plus an average of 19 Euros in additional payouts.

Game 1	*L*	*R*
U	(10, 8)	(0, 18)
D	(9, 9)	(10, 8)

Game 7	*L*	*R*
U	(10, 12)	(4, 22)
D	(9, 9)	(14, 8)

Game 2	*L*	*R*
U	(9, 4)	(0, 13)
D	(6, 7)	(8, 5)

Game 8	*L*	*R*
U	(9, 7)	(3, 16)
D	(6, 7)	(11, 5)

Game 3	*L*	*R*
U	(8, 6)	(0, 14)
D	(7, 7)	(10, 4)

Game 9	*L*	*R*
U	(8, 9)	(3, 17)
D	(7, 7)	(13, 4)

Game 4	*L*	*R*
U	(7, 4)	(0, 11)
D	(5, 6)	(9, 2)

Game 10	*L*	*R*
U	(7, 6)	(2, 13)
D	(5, 6)	(11, 2)

Game 5	*L*	*R*
U	(7, 2)	(0, 9)
D	(4, 5)	(8, 1)

Game 11	*L*	*R*
U	(7, 4)	(2, 11)
D	(4, 5)	(10, 0)

Game 6	*L*	*R*
U	(7, 1)	(1, 7)
D	(3, 5)	(8, 0)

Game 12	*L*	*R*
U	(7, 3)	(3, 9)
D	(3, 5)	(10, 0)

Note: Payoffs for row (r) and column (c) players are given (r,c) in the matrix. Abbreviations of for Up, Down, Left, and Right are given as U,D,L, and R.

Figure 1. 2 × 2 Games.

5. Measuring Goodness of Fit

Following Chmura et al. [9], we use a quadratic scoring rule in order to assess the goodness of fit of each learning model.[11] This rule, as described in Equation (13), provides a measure of nearness from the predicted choice to the observed choice.[12]

$$q_i(t) = 2p_i(t) - p_i(t)^2 - (1 - p_i(t))^2 \quad (13)$$

The quadratic score, $q_i(t)$, is a function of the probability, $p_i(t)$, of the choice by action i in period t. p is the predicted probability that is derived from the parameters of the learning models. The score is equal to 1 minus the squared distance between the predicted probability and the actual choice.

The expected range of $q_i(t)$ is $[-1, 1]$. On one hand, if a learning model predicts the data perfectly, then $p_i(t) = 1$, which implies $q_i(t) = 1$. On the other hand, a completely uninformative learning model, in our setting, would be right half the time, so $p_i(t) = 0.5$, which implies $q_i(t) = 0.5$.

We employ the quadratic scoring rule in order to understand goodness of fit of each learning model in multiple tests. First, we calculate parameters on the entire playing history of all subjects and use the best-fitting parameters to estimate the predicted probabilities across playing history and calculate the mean quadratic score for each learning model. Next, we employ a rolling forward out-of-sample procedure. The out-of-sample process is chosen by fitting all models on the first X% of the data and using the fitted parameters of the model to predict the holdout sample of $(100 - X)$% of the data. We then calculate the mean quadratic score for the remaining out-of-sample observations. We repeat for different values of X; in particular, we use 40%, 50%, 60%, 70%, and 80% in-sample training data to predict choice on the remaining 60%, 50%, 40%, 30%, and 20% remaining data, respectively. The in-sample method is a standard way to judge goodness-of-fit, by simply looking at how much of the whole data the model can explain individual choice. The out-of-sample method guards against over-fitting the data, but, to be valid, it assumes stationarity of parameters across the in-sample and holdout data. For concerns of over-fitting the data with any learning model, this out-of-sample procedure is the preferred benchmark in choosing which model explain the data best.

In estimating CBL, we use information available to subjects to define the Problem set $\mathcal{P}$. In our main specification, we choose two elements in the information vector (i.e., problem vector). The first element is the round of the game (i.e., time). The round of the game plays the role of recency, or forgetting, in other learning models; cases that are distant in the past are less similar to present circumstances than cases that happened more recently. The second element is the opponents' play from the game. We account for other players actions by using a moving average of past play, treating all opponents as a representative player, just as we do for the surprise index in self-tuning EWA. We use a four period moving average. For example, a row player would use the moving average of how many times their opponents played Left as a component of similarity and, as opponents trend toward different frequencies of playing Left, the CBL would put less weight on those cases, $\mathcal{C}$. There are many possible choices on how to incorporate these information vectors and we explore them further in the Appendix B in Table A1. We find that these choices do not have a large effect on the performance of CBL.

We include cases as much as 15 periods in the past in memory (we explore the sensitivity of this assumption in Section 7.1).

---

[11] The quadratic scoring rule was introduced by Brier [38] to measure performance in weather forecasting. This scoring rule is also described in Selten [39].

[12] The use of other measures of goodness of fit generally provide the same qualitative measures, but ordering of preferred learning models can be reversed by employing Log-Likelihood when model fitness is relatively close. We prefer the quadratic scoring rule and use that throughout.

6. Results

To fit the learning models to data, we estimate Equation (2) for CBL, Equation (5) for RL, and Equation (6) for self-tuning EWA. All of the learning algorithms use the stochastic logit choice rule in Equation (12). In Figure 2, we report the mean quadratic score by the learning models discussed in the previous section across all 12 experimental games. We find when using in-sample measures between the learning models that CBL fits best, RL fits second best, and self-tuning EWA fits third best.[13] RL performs about as well as CBL across these experimental games. As expected, each learning model outperforms a baseline benchmark of random choice (i.e., a mean quadratic score of 0.5). Note that Chmura et al. [9] also find that self-tuning EWA and a selection of other simple learning models out-perform random choice, but they found self-tuning EWA was the best performing learning model in predicting individual choice with these data.[14]

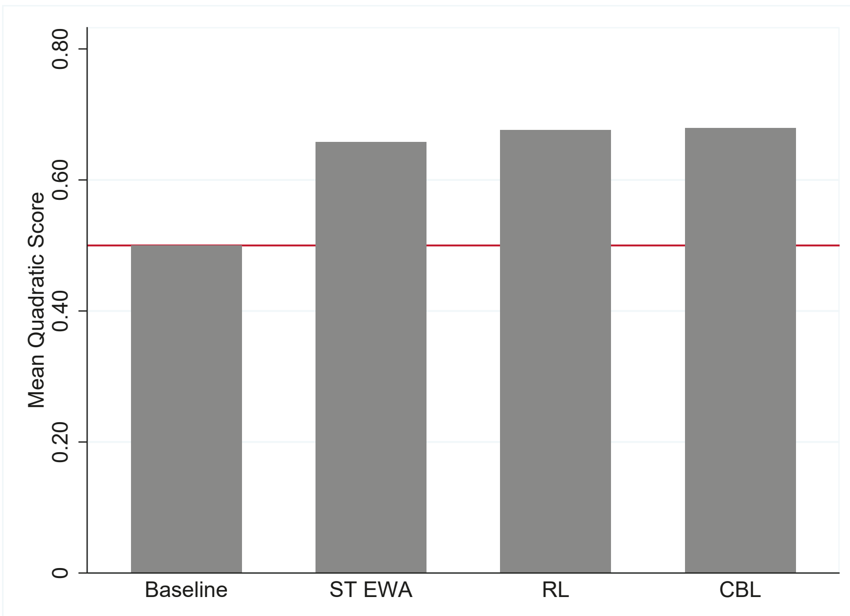


Figure 2. In-sample Fit of Learning Models. Note: The red line represents the quadratic score of the baseline model which is the predicted score of a learning model picking strategies at random. ST EWA refers to self-tuning EWA, RL refers to reinforcement learning, and CBL refers to case-based learning.

We use a non-nested model selection test proposed by Vuong [40], which provides a directional test of which model is favored in the data generating process. Testing the CBL model versus the RL model, the Vuong test statistic is 7.45, which is highly significant and favors the selection of the CBL model. In addition, we find that the CBL is selected over the ST EWA model with a Voung test statisitic of 37.91.

In Figure 3, we report the mean quadratic scores of the out-of-sample data using in-sample parameter estimates. We find similar conclusions as in the in-sample fit in Figure 2. CBL fits best,

[13] The mean squared error is 0.1618 for RL, 0.1715 for self-tuning EWA, and 0.1603 for CBL, where the ordering of selection of models is the same as the quadratic scoring rule.

[14] We estimate the initial attractions in our self-tuning EWA model while Chmura et al. [9] do not, which does not appear to make much of a difference in goodness of fit. They assume a random action initially for all learning models investigated. Chmura et al. [9] also estimates a one parameter RL model, which under performs self-tuning EWA.

followed by RL, and then by ST EWA. This leads us to presume that CBL may be better at explaining behavior across all these games, likely due to the inclusion of information about the moving average of opposing players' play during the game. It is important to note that the RL predicts almost as well as CBL with arguably a simpler learning model. The experiments we use, and have been traditionally used to assess learning models, are relatively information-poor environments for subjects compared to some other games. For example, many one-shot prisoner dilemma games or coordination games where information about partner's identity or their past play is public knowledge would be a comparably information-rich environment. This makes us optimistic that CBL may be even more convincing in information-rich environments. Because CBL makes use of the data about opposing players, CBL is an obvious candidate to accommodate this type of information in a systematic way that seems consistent with the psychology of decision making.

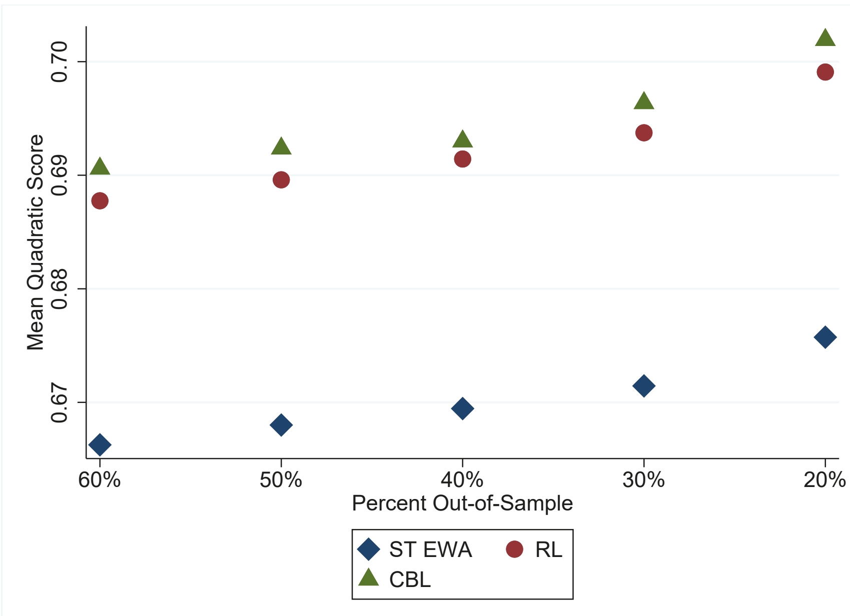


Figure 3. Out-of-sample Fit of Learning Models. Note: Each model is estimated using a portion of the data, while goodness of fit is measured on the remaining data. ST EWA refers to the self-tuning EWA, RL refers to reinforcement learning, and CBL refers to case-based learning.

7. Case-Based Parameters

In this section, we discuss the parameters of CBL estimation based on the full sample estimation. The parameters of CBL are λ, which measures the sensitivity of choice to CBU (see Section 3.6), A_0^L is the parameter measuring the initial attraction for Left for the column player, while A_0^U is the initial attraction for Up for the row player (see Section 3.1). These initial attractions are relative measures as the initial attractions for Down and Right are held at zero. W_i are weights in the similarity function on the different characteristics of the information vector (see Equation (4)). In particular, W_1 is the weight given to recency (here, round number), W_2 is the weight given to the moving average of actions of opposing players. These parameters are estimated to best fit the data using the logit rule in Equation (12).

We do not directly estimate the aspiration parameter, because it cannot be effectively empirically distinguished from the initial attraction parameters. If one considers Equation (2), one can see that the H parameter and the mean of the A_j parameters confound identification. We cannot distinguish between the average initial attractions to strategies due to priors and the aspiration value of the agent. Fortunately, we find that the fit of the CBL generally does not rely on the estimation of the aspiration level to achieve the same goodness-of-fit.

In Table 1, we report the estimated parameters using the full sample of observations in each treatment of each experiment. In all experimental treatments, we find a statistically significant value for λ, meaning that the learning algorithm estimated explains some choice. We find that the initial attraction parameters A_0^j are consistent with the frequency of choices in the first period. The relative weights of W_1 and W_2 are difficult to directly compare, as they are in different scales. We could normalize the data prior to estimation, but it is unclear what affect that might have on cumulative CBL over time. We explore ex-post normalization of the parameters in Appendix C and list results in Table A2. The empirical estimate of W_1 is positive and statistically significant. This indicates that, consistent with other learning models, recency is important to learning.

By comparing the coefficients W_1 in Table 1, we find that recency degrades similarity faster in non-constant sum games than in constant sum games. This difference suggests that in non-constant sum game, subjects 'forget' past experience faster when constructing expectations about the current problem and they put relatively more weight on the similarity of the moving average of opposing players.

The weight, W_2, on the moving average of past play of opposing players is positive and significant. A positive parameter gives greater weight to cases with similar average playing rates to the current problem. This parameter picks up adjustments to group actions over time.

Table 1. CBL Parameter Estimates. Note: ***, **, * denote statistical significance of 10%, 5%, and 1%. Clustered standard errors by subject are in parentheses. MQS is the mean quadratic score. † The standard error did not calculate using clustered standard errors and is instead calculated using the outer product of the gradient (OPG) vectors method.

			A: Combined Models				
	λ	A_0^L	A_0^U	W_1	W_2	N	MQS
Constant sum games	10.896 ***	−1.596 ***	1.775 ***	0.043 ***	15.917 ***	115,200	0.688
	(0.331)	(0.211)	(0.250)	(0.002)	(0.002)		
Non-constant sum games	37.831 ***	−1.024 ***	1.023 ***	1.136 ***	10.259 ***	57,600	0.653
	(1.748)	(0.148)	(0.143)	(0.000)	(1.006) †		
All games	10.179 ***	−1.946 ***	1.553 ***	0.049 ***	9.365 ***	172,800	0.679
	(0.186)	(0.195)	(0.204)	(0.000)	(0.364) †		
			B: Individual Game Models				
	λ	A_0^L	A_0^U	W_1	W_2	**N**	**MQS**
Game 1	15.543 ***	−1.397 ***	4.821 ***	0.205 ***	4.146 ***	19,200	0.809
	(0.823)	(0.437)	(1.021)	(0.016) †	(0.953) †		
Game 2	6.655 ***	−0.730 *	2.395 **	0.019 ***	5.316 ***	19,200	0.667
	(0.925)	(0.444)	(1.179)	(0.007)	(1.376)		
Game 3	15.175 ***	−1.753	2.465 ***	0.127	3.929 ***	19,200	0.768
	(0.584) †	(3.794)	(0.269) †	(0.134)	(0.398)		
Game 4	13.758 ***	−2.683 ***	2.139 ***	0.139 ***	3.167 ***	19,200	0.654
	(0.832)	(0.471)	(0.352)	(0.005)	(0.028)		
Game 5	110.704 ***	−0.416 ***	0.377 ***	3.204 ***	7.261 ***	19,200	0.630
	(8.723)	(0.071)	(0.078)	(0.000)	(2.352) †		
Game 6	65.058 ***	−0.406 ***	0.114	1.535 ***	3.246 **	19,200	0.594
	(5.276)	(0.125)	(0.090)	(0.003)	(1.507) †		
Game 7	33.815 ***	−0.431	2.135 ***	1.005 ***	8.853 ***	9600	0.741
	(3.764)	(0.316)	(0.618)	(0.104) †	(0.875)		
Game 8	5.300 *	−3.121 **	3.856 **	0.020	5.155 *	9600	0.637
	(3.092)	(1.475)	(1.941)	(0.041)	(2.783)		
Game 9	10.058 ***	−4.345 ***	2.512 ***	0.080 ***	2.784 ***	9600	0.743
	(0.724)	(1.155)	(0.695)	(0.000)	(0.004)		
Game 10	5.837 ***	−4.364 **	−0.563	0.020 ***	1.499 ***	9600	0.639
	(0.836)	(2.063)	(1.025)	(0.008)	(0.013)		
Game 11	20.525 ***	−1.588 ***	1.444 ***	0.394 ***	0.767 ***	9600	0.616
	(1.638)	(0.487)	(0.416)	(0.000)	(0.004)		
Game 12	7.184 ***	−1.864	0.604	0.012 **	8.471 ***	9600	0.593
	(1.078)	(1.162)	(0.890)	(0.005)	(2.254)		

7.1. Memory

We explore to what extent adding memory explains behavior in CBL. This is an important part of our depiction of learning, and we test the regularity to which it is important by varying the known memory of subjects in CBL. Figure 4 shows the improvement in the mean quadratic score as more memory is allowed in the CBL algorithm starting with three prior periods in memory ($M = 3$) and expanding to seventeen prior periods ($M = 17$). If we refer to three periods in memory, then subjects 'forget' periods that were further in the past than three periods (rounds) ago and do not consider them in comparing the current periods definition of a case. The figure demonstrates that increasing the length of short time-horizons provide an improvement in model fit, but most gains are exhausted by around nine periods. Because period number is included as an element of the problem definition $\mathcal{P}$, continuing to add more periods into the similarity function makes little difference past nine periods and degrades past fifteen periods. This provides the basis of our choice of fifteen periods for estimation.

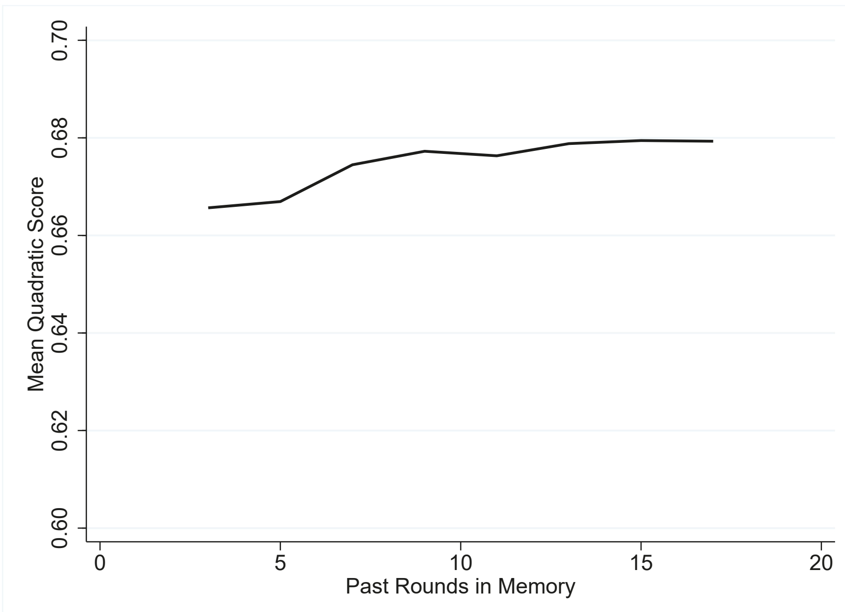


Figure 4. Length of Memory M.

7.2. Definition of Similarity

We estimate multiple similarity functional forms and measures of distance between the attributes used in the definition similarity between problems. We test how similarity as characterized by Equation (14) compares to in-sample fit of the data. The definitions of our similarity functions primarily differ in how similarity decays; above, we assume similarity decays exponentially and, in Equation (14), it decays according to the logarithm (Ones were added to avoid dividing-by-zero and log-of-zero problems).

$$S^2(x,y) = \frac{1}{ln(d(x,y)+1)+1} \tag{14}$$

In addition to the decay of similarity, we can also test a different definition of distance between elements of the problem. In our main specification, we use weighted Euclidean distance, as defined in Equation (4). Another popular definition in psychology for distance is the Manhattan distance given by Equation (15).

$$d^2(x,y) = \sum_{i=1}^{\#Dims} w_i |x_i - y_i| \tag{15}$$

Using these definitions, we report that the in-sample fit of the data, measured by the quadratic scoring rule, in Table 1 to be robust to the various definitions of similarity and distance. We indicate by column heading in Table 2 which equations were used corresponding to specific functional forms of similarity and distance. The functional form of distance between elements seems to be of minor importance to fit in the mixed strategy equilibria games explored here. Nevertheless, there is greater variation in the performance of the different similarity functions. The similarity function provided in Equation (14) performs better than the exponential. We also find the weight W_2 with Equation (14) is negative and statistically significant, which is unexpected. To avoid overfitting the data with parameters that do not make psychological sense, we use Equation (3) in our main specification. We conclude that CBL is robust to different definitions of similarity, and the inverse exponential function is a good fit with the experimental data at hand. This also corresponds to previous findings in psychology and economics [4,33].

Table 2. Similarity Definitions Measured by Mean Quadratic Score.

	(1)	(2)	(3)	(4)
	S, d	S^2, d	S, d^2	S^2, d^2
MQS	0.679	0.686	0.681	0.682

Note: S denotes the similarity function in Equation (3) and S^2 denotes the similarity function in Equation (4). d denotes the Euclidean distance function and d^2 denotes the Manhattan distance function.

8. Empirical Comparison of Learning Models

In this section, we investigate the dynamics of RL and CBL, the two best-fitting learning models, to more fully understand the results of these learning algorithms. Previously we discussed the potential overlap in RL and CBL, which in practice have similar fits to the data. CBL likely outperforms RL in aggregate due to its ability to incorporate important information in the the choice behavior of subjects. RL and CBL appear to converge on choices overtime. We illustrate convergence in prediction between CBL and RL in Figure 5. There is a possibility that RL and CBL are increasingly correct about different types of individual decisions and could not actually be converging to similar predictions of behavior. For example, say there are three types of decision makers (A, B, and C). CBL and RL predict players of type A well, but not B or C. As more information is added and the learning models improve goodness of fit, CBL predicts player type B better and RL predicts player type C better. Both of the models are doing better, but are doing it on different observations and therefore on not converging on the types of predictions they get correct. The convergence between CBL and EWA by round in Figure 5 demonstrates that the gains in accuracy are accompanied by a convergence in agreement between the two learning algorithms, although convergence is slight. The coefficient of the regression line in Figure 5 is −0.00006 with a clustered standard error by game type of 0.00002. This coefficient is statistically significant with a t-statistic of −2.98.

We also provide the model fits by individual games in Appendix D. Table A3 and A4 show the in-sample and out-of-sample model fits by individual game for all learning models.

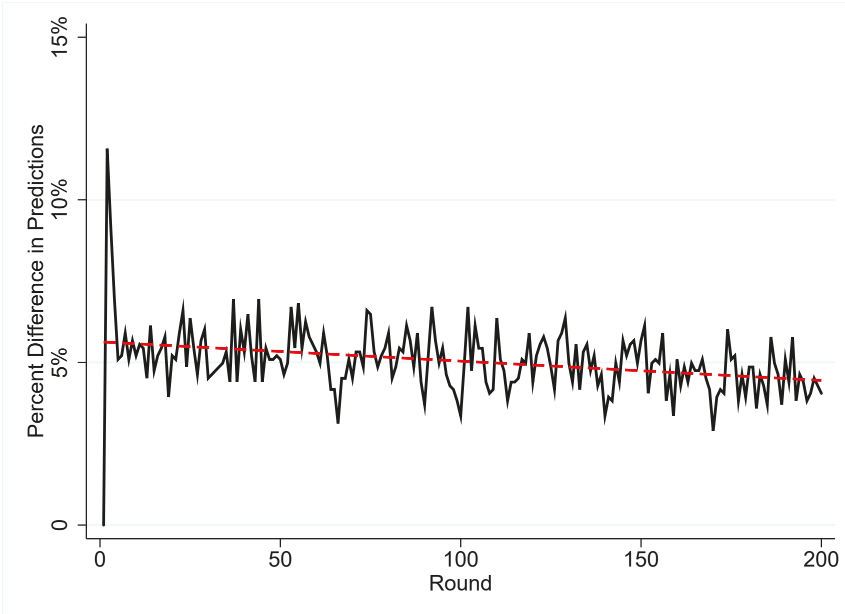


Figure 5. Convergence of RL and CBL. The red line denotes an OLS regression line of round on percent difference in predictions.

9. Discussion

The parameters of CBL are also related to other theories of learning. Aspiration levels can be incorporated into both the self-tuning EWA and CBL models of learning. They are somewhat inherent in the self-tuning EWA algorithm already because the attractions compare the average payoffs of previous attractions to the the current attraction and act as an endogenous version of aspiration level that incorporates foregone payoffs. Another similarity between these theories of behavior is recency, or the weighting of events that have more recently occurred in the past. In self-tuning EWA and RL, the parameter ϕ and cumulative attractions account for recency and, in CBL, a time indicator in the definition of the problem and definition of memory account for recency. In all the learning models, recency allows individuals to 'forget' old occurrences of a problem and adapt to new emergent behavior or payoffs.

In this study, we show the effectiveness of CBL in an environment used traditionally for learning models. CBL can easily be applied to other contexts with the same basic construction showed here. It would be simple to estimate the same algorithm on other games through the same procedures or choice data from outside the lab. The harder question is how to define the Problem, $\mathcal{P}$, in these different environments. In Appendix B, we discuss different definitions of the Problem for this context. We find the results aree robust to the definition of Problem in Table A1. In more information-rich contexts, it may be difficult to decide the number of characteristics, how information is presented in the similarity function, and whether fictitious cases are present in memory. One approach, if experiments are used, could be to track attention to particular pieces of information (e.g., through mouse clicks, eyetracking, or even asking subjects). Collecting secondary information on choices may be beneficial testing axioms of case-based decision theory. Further, as in Bleichrodt et al. [17], through experimental design, decision weights for information can be constructed to further understand the properties of CBL that can be non-parametrically estimated.

One suggested limitation of learning models is that they do not explain why the way partners are matched matters [41], although more sophisticated learners can address this deficiency [32].

CBL may better explain why matching matters directly through the information vector and the similarity function. This is the biggest difference between CBL and other algorithms: the formulation of how information enters into decision making, which is systematic, follows what we know from psychology about decision making of individuals, and it is shown to be important through numerous experimental investigations.

CBL also has the ability to incorporate fictitious play, although we do not pursue this in the current paper. As mentioned above, although typically an agent's memory is the list of cases she has experienced—which is what we assume here—it is possible for cases in memory to come from some other sources. The agent could add fictitious play cases to memory and thereafter use those fictitious cases to calculate CBL. Moreover, this agent could distinguish fictitious from real cases if she so desired by adding a variable to the information vector denoting whether the case was fictitious; then fictitious cases could have less—but not zero—importance relative to real cases. Bayesian learning can also be tractable in the case of 2×2 games, where the dimensions of the state space are small. We did not consider forgone payoffs in CBL and, therefore, did not compare CBL to other Belief models. We can imagine that comparing CBL to Bayesian learning models that would have priors defined over the complete state space is a natural next step in this line of research. We leave this for future investigation.

10. Conclusions

In this work, we demonstrate the estimation of a new learning model based on an existing decision theory, Case-based Decision Theory. This form of decision-making under uncertainty when applied to game theoretic experiments performs well when compared to two other leading learning models: Reinforcement Learning and self-tuning Experience Weighted Attraction. An important feature of Case-Based Learning is the ability to systematically incorporate information that is available to subjects into choice decisions. Real people condition their behavior on their observations of their environments, and the case-based approach incorporates this in a natural way.

The parameters of Case-Based Learning indicate a relationship between recency and the type of game played. Constant sum games exhibit a smaller recency effect than non-constant sum games. This indicates that subjects weight experience with a opponents differently, depending on the type of game played. Significant attention is given to the average rates of play, and changes to those average rates of play, through the inclusion of moving averages in the definition of the 'Problem'.

Further work in applying CBL to other decision making environments is important in understanding its limits and sensible parameterizations of information vectors, hypothetical references, and how deliberate thoughtful decisions are affected by institutions that include information. CBL could also be used and adapted to predict behavior across different types of games, or more generally changes to decision-making environments, to understand the influence of previous game play on decisions in new environments and how subjects encode information across games. A natural extension of this type of investigation are repeated prisoner dilemma games or repeated coordination games, which allow more complex equilibria, but also allow for a greater freedom to explore the primitives of learning across environments.

Author Contributions: T.G. performed the formal analysis; T.G. and A.D.P. contributed equally to all other parts of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proofs

Proposition A1. *Let N(t)=N(t+1) and* $H = 0$, $A_0 = 0$ *for all choices, only time is in the definition of the problem for CBL, and the similarity function is simple inverse weighted exponential given in Equation (10). Then there exists* $\phi > 0$ *such that RL attraction RLA with parameter* ϕ *and the case-based attraction implied by* *the similarity function S are equal, and therefore the attractions decay at the same rate and ϕ and w are related in Equation (A1).*

$$\phi = S(x_t, x_{t-1}) = \frac{1}{w} \tag{A1}$$

For Proposition 1 we provide the following proof. When considering attractions from RL and CBL with the following simplifications: $N(t) = N(t+1), H = 0, A_0 = 0$, and only time is in definition of the problem for CBL. Under the condition that the similarity function is an inverse exponential of the difference in the time index, RL attraction degrades at the rate ϕ while CBL degrades at the rate defined by the following similarity function,

$$\phi = \frac{1}{w^{|t-t-1|}}. \tag{A2}$$

Then, for all past attractions more generally, $\phi^k = \frac{1}{w_t{}^{|k|}}$. As past attractions are discounted in RL they get discounted each time period by ϕ so to is the weight in CBL by an equivalent adjustment in the distance between time periods. This rate is held constant across time in both models since $N(t) = N(t+1)$.

The typical similarity and distance functions used in the literature do not have equivalence between the similarity and recency in RL, and therefore this may be seen as a special case of the relationship between RL and CBL. We do not use these specialized forms to estimate against the data, but rather use them to demonstrate the simple similarities and differences in how CBL and RL are constructed.

Appendix B. Definition of the Problem

In this section we discuss in greater detail the definition of the problem set, $\mathcal{P}$, or, equivalently, the definition of the information vector. Experiments are very helpful for the researcher to define the information vector used in CBL since information is experimentally controlled and limited compared to observed behavior in the 'wild.' Here we describe the decisions we made to define the problem set $\mathcal{P}$ in this series of 2×2 games.

In contructing a measure of recency we assume that rounds are considered as simple vectors of whole numbers and do not consider additional non-linearities in this information, such as squares or other transformations of the data. Perhaps more difficult in this setting is the definition of opposing player behavior since actions are anonymous in all games of this experiment. The history of the opposing player's action could be incorporated in many ways into the problem. We use a moving average of the past play for all group members encountered by a subject, so as the recent trend of play changes, agents adapt to those trends. Another possibility is that agents use rules that specify the ordering of past play instead of a moving average. We can accommodate this definition in CBL by using binary indicators for the lag in observed play, This would make sense if subjects used strategies similar to Tit-for-Tat or more complex patterns that incorporate how the last three rounds of play occurred. For completeness we estimate CBL under these different information vectors and find small improvements in goodness of fit with more parameters. Table A1 shows the results of different definitions of the problem.

While the use of additional parameters improves the goodness of fit of the model, we choose to use the simpler and possibly more conservative moving average measure in the main models. We argue that our main estimates are more conservative based on fit, but are preferred because the parameters for the weights on the lags are negative in some cases which violate our understanding of the reasonable parameters for this model.

Table A1. Information Vector Definitions.

	(1) Mean Quadratic Score
One lag	0.677
Two lags	0.684
Three lags	0.682
MA-3	0.678
MA-5	0.678

Note: Each lag represents a extra parameter on whether a partner in the past played either Up or Left, depending on the roll of the subject. Abbreviations: MA-3 moving average for past three rounds, MA-5 moving average for past five rounds.

Appendix C. Normalization of Weights

To compare the relevance of the information vectors we need to transform the weights into comparable units. Therefore, we use a simple form of transformation by multiplying the estimated coefficients by the standard deviation of the data. The transformed coefficient would approximate how much a standard deviation in the data affects the similarity function. This does not change the interpretation of statistical significance, but does provide a way to assess the economic significance of the information vectors to subjects.

In Table A2, we find that the moving average captures more relative weight in constant sum games while recency is weighed more heavily in non-constant sum games. While both weights appear to be economically significant to define similarity of cases across all games, a standard deviation change in the moving average of opposing players garners much more weight on choices on average. Consistent with our previous interpretation, the non-constant sum games tend to discount cases in the relatively far past heavily compared to constant sum games.

Table A2. CBL Normalized weights.

	W_1	W_2
Constant sum games	0.544	2.435
Non-constant sum games	14.245	1.539
All games	0.622	1.414

Appendix D. Individual Game Results

In this section we provide the detailed results of the learning algorithms on each individual game.

Table A3. In-sample Fit by Game: Mean Quadratic Score * denotes the best-fitting model based on the mean quadratic score. CBL has five estimated parameters, RL has four estimated parameters, and ST EWA has one estimated parameter. ST EWA refers to the self-tuning EWA, RL refers to reinforcement learning, and CBL refers to case-based learning.

	CBL	ST EWA	RL
Game 1	0.809 *	0.785	0.808
Game 2	0.667 *	0.661	0.667
Game 3	0.768 *	0.750	0.767
Game 4	0.654 *	0.636	0.650
Game 5	0.630	0.608	0.633 *
Game 6	0.594	0.568	0.598 *
Game 7	0.741	0.735	0.751 *
Game 8	0.637 *	0.629	0.636
Game 9	0.743 *	0.706	0.738
Game 10	0.639 *	0.607	0.637
Game 11	0.616 *	0.597	0.615
Game 12	0.593 *	0.573	0.590

Table A4. Out-of-sample Fit by Game Note: * denotes the best fitting model based on the mean quadratic score. ST EWA refers to the self-tuning EWA, RL refers to reinforcement learning, and CBL refers to case-based learning.

	A: Out-of-Sample: Predict Last 60%			B: Out-of-Sample: Predict Last 50%		
	CBL	ST EWA	RL	CBL	ST EWA	RL
Game 1	0.827	0.811	0.838 *	0.845 *	0.817	0.844
Game 2	0.674	0.669	0.675 *	0.671	0.671	0.677 *
Game 3	0.793 *	0.771	0.793	0.782	0.778	0.798 *
Game 4	0.658 *	0.637	0.653	0.659 *	0.637	0.655
Game 5	0.636 *	0.607	0.632	0.641 *	0.607	0.634
Game 6	0.601 *	0.568	0.600	0.603 *	0.567	0.599
Game 7	0.768	0.766	0.783 *	0.784	0.769	0.785 *
Game 8	0.641	0.638	0.644 *	0.637	0.637	0.642 *
Game 9	0.743	0.715	0.747 *	0.746	0.711	0.747 *
Game 10	0.632	0.605	0.638 *	0.645 *	0.610	0.641
Game 11	0.616	0.598	0.617 *	0.618	0.601	0.624 *
Game 12	0.593 *	0.576	0.593	0.579	0.574	0.593
	C: Out-of-Sample: Predict Last 40%			**D: Out-of-Sample: Predict Last 30%**		
	CBL	**ST EWA**	**RL**	**CBL**	**ST EWA**	**RL**
Game 1	0.843	0.818	0.847	0.854	0.82	0.852
Game 2	0.665	0.672	0.680 *	0.669	0.671	0.680 *
Game 3	0.787	0.779	0.800 *	0.792	0.780	0.802 *
Game 4	0.661 *	0.636	0.655	0.664 *	0.641	0.659
Game 5	0.641 *	0.607	0.634	0.643 *	0.611	0.635
Game 6	0.594	0.569	0.600 *	0.606 *	0.574	0.603
Game 7	0.789	0.775	0.792 *	0.791	0.777	0.795 *
Game 8	0.635	0.635	0.640 *	0.639	0.634	0.639 *
Game 9	0.750 *	0.715	0.748	0.737	0.714	0.748 *
Game 10	0.651 *	0.613	0.650	0.643	0.614	0.651 *
Game 11	0.621	0.605	0.626 *	0.621	0.603	0.624 *
Game 12	0.596 *	0.574	0.594	0.599	0.579	0.599 *
	E: Out-of-Sample: Predict Last 20%					
Game 1	0.866 *	0.821	0.862			
Game 2	0.694	0.686	0.697 *			
Game 3	0.804	0.789	0.806 *			
Game 4	0.660	0.642	0.661 *			
Game 5	0.637 *	0.612	0.636			
Game 6	0.609 *	0.574	0.604			
Game 7	0.794 *	0.776	0.794			
Game 8	0.634	0.637	0.643 *			
Game 9	0.746	0.726	0.757 *			
Game 10	0.663	0.626	0.666 *			
Game 11	0.620	0.608	0.629 *			
Game 12	0.591	0.575	0.595 *			

References

1. Gilboa, I.; Schmeidler, D. Case-based decision theory. *Q. J. Econ.* **1995**, *110*, 605–639. [CrossRef]
2. Von Neumann, J.; Morgenstern, O. *Theory of Games and Economic Behavior*; Princeton University Press: Princeton, NJ, USA, 1944.
3. Savage, L.J. *The Foundations of Statistics*; Wiley: New York, NY, USA, 1954.
4. Pape, A.D.; Kurtz, K.J. Evaluating case-based decision theory: Predicting empirical patterns of human classification learning. *Games Econ. Behav.* **2013**, *82*, 52–65. [CrossRef]
5. Guilfoos, T.; Pape, A.D. Predicting human cooperation in the Prisoner's Dilemma using case-based decision theory. *Theory Decis.* **2016**, *80*, 1–32. [CrossRef]
6. Luce, R.D. *Individual Choice Behavior*; Dover Publications Inc.: Menola, NY, USA, 1959.

7. Erev, I.; Roth, A.E. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* **1998**, *88*, 848–881.
8. Selten, R.; Chmura, T. Stationary concepts for experimental 2 × 2-games. *Am. Econ. Rev.* **2008**, *98*, 938–966. [CrossRef]
9. Chmura, T.; Goerg, S.J.; Selten, R. Learning in experimental 2× 2 games. *Games Econ. Behav.* **2012**, *76*, 44–73. [CrossRef]
10. Ho, T.H.; Camerer, C.F.; Chong, J.K. Self-tuning experience weighted attraction learning in games. *J. Econ. Theory* **2007**, *133*, 177–198. [CrossRef]
11. Gayer, G.; Gilboa, I.; Lieberman, O. Rule-based and case-based reasoning in housing prices. *BE J. Theor. Econ.* **2007**, *7*. [CrossRef]
12. Kinjo, K.; Sugawara, S. Predicting empirical patterns in viewing Japanese TV dramas using case-based decision theory. *BE J. Theor. Econ.* **2016**, *16*, 679–709. [CrossRef]
13. Golosnoy, V.; Okhrin, Y. General uncertainty in portfolio selection: A case-based decision approach. *J. Econ. Behav. Organ.* **2008**, *67*, 718–734. [CrossRef]
14. Guerdjikova, A. Case-based learning with different similarity functions. *Games Econ. Behav.* **2008**, *63*, 107–132. [CrossRef]
15. Ossadnik, W.; Wilmsmann, D.; Niemann, B. Experimental evidence on case-based decision theory. *Theory Decis.* **2013**, *75*, 1–22. [CrossRef]
16. Grosskopf, B.; Sarin, R.; Watson, E. An experiment on case-based decision making. *Theory Decis.* **2015**, *79*, 639–666. [CrossRef]
17. Bleichrodt, H.; Filko, M.; Kothiyal, A.; Wakker, P.P. Making case-based decision theory directly observable. *Am. Econ. J. Microecon.* **2017**, *9*, 123–151. [CrossRef]
18. Radoc, B.; Sugden, R.; Turocy, T.L. Correlation neglect and case-based decisions. *J. Risk Uncertain.* **2019**, *59*, 23–49. [CrossRef]
19. Bhui, R. Case-Based Decision Neuroscience: Economic Judgment by Similarity. In *Goal-Directed Decision Making*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 67–103.
20. Gayer, G.; Gilboa, I. Analogies and theories: The role of simplicity and the emergence of norms. *Games Econ. Behav.* **2014**, *83*, 267–283. [CrossRef]
21. Bordalo, P.; Gennaioli, N.; Shleifer, A. Memory, Attention, and Choice. *Q. J. Econ.* **2020**, *135*, 1399–1442. [CrossRef]
22. Argenziano, R.; Gilboa, I. Similarity & Nash Equilibria in Statistical Games. Available online: https://www.researchgate.net/publication/334635842_Similarity-Nash_Equilibria_in_Statistical_Games (accessed on 23 July 2019).
23. Gilboa, I.; Lieberman, O.; Schmeidler, D. Empirical Similarity. *Rev. Econ. Stat.* **2006**, *88*, 433–444. [CrossRef]
24. Billot, A.; Gilboa, I.; Schmeidler, D. Axiomatization of an Exponential Similarity Function. *Math. Soc. Sci.* **2008**, *55*, 107–115. [CrossRef]
25. Matsui, A. Expected Utility and Case-Based Reasoning. *Math. Soc. Sci.* **2000**, *39*, 1–12. [CrossRef]
26. Charness, G.; Gneezy, U. What's in a name? Anonymity and social distance in dictator and ultimatum games. *J. Econ. Behav. Organ.* **2008**, *68*, 29–35. [CrossRef]
27. Guilfoos, T.; Kurtz, K.J. Evaluating the role of personality trait information in social dilemmas. *J. Behav. Exp. Econ.* **2017**, *68*, 119–129. [CrossRef]
28. Cerigioni, F. *Dual Decision Processes: Retrieving Preferences When Some Choices are Intuitive; Economics Working Papers*; Department of Economics and Business, Universitat Pompeu Fabra: Barcelona, Spain, 2016.
29. Fudenberg, D.; Levine, D.K. *The Theory of Learning in Games*; MIT Press: Cambridge, MA, USA, 1998; Volume 2.
30. Camerer, C.; Ho, T.H. Experience-weighted attraction learning in coordination games: Probability rules, heterogeneity, and time-variation. *J. Math. Psychol.* **1998**, *42*, 305–326. [CrossRef] [PubMed]
31. Camerer, C.; Ho, T.H. Experience-weighted attraction learning in normal form games. *Econometrica* **1999**, *67*, 827–874. [CrossRef]
32. Camerer, C.F.; Ho, T.H.; Chong, J.K. Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *J. Econ. Theory* **2002**, *104*, 137–188. [CrossRef]
33. Shepard, R. Toward a Universal Law of Generalization for Psychological Science. *Science* **1987**, *237*, 1317. [CrossRef]

34. Dayan, P.; Niv, Y. Reinforcement learning: The good, the bad and the ugly. *Curr. Opin. Neurobiol.* **2008**, *18*, 185–196. [CrossRef]
35. Glimcher, P.W. Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 15647–15654. [CrossRef]
36. Harley, C.B. Learning the evolutionarily stable strategy. *J. Theor. Biol.* **1981**, *89*, 611–633. [CrossRef]
37. Roth, A.E.; Erev, I. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games Econ. Behav.* **1995**, *8*, 164–212. [CrossRef]
38. Brier, G.W. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* **1950**, *78*, 1–3. [CrossRef]
39. Selten, R. Axiomatic characterization of the quadratic scoring rule. *Exp. Econ.* **1998**, *1*, 43–61. [CrossRef]
40. Vuong, Q.H. Likelihood ratio tests for model selection and non-nested hypotheses. *Econom. J. Econom. Soc.* **1989**, *57*, 307–333. [CrossRef]
41. Andreoni, J.; Miller, J.H. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *Econ. J.* **1993**, *103*, 570–585. [CrossRef]

games

MDPI

Article

With Potential Games, Which Outcome Is Better?

Santiago Guisasola [1,*] and Donald Saari [2,3,4]

[1] Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, RJ 22460-320, Brazil
[2] Department of Mathematics, University of California, Irvine, CA 92617, USA; dsaari@uci.edu
[3] Department of Economics, University of California, Irvine, CA 92617, USA
[4] Institute for Mathematical Behavioral Sciences, University of California, Irvine, CA 92617, USA
* Correspondence: santiago.guisasola@impa.br

Received: 27 June 2020; Accepted: 6 August 2020; Published: 17 August 2020

Abstract: Lower one- or two-dimensional coordination, or potential games, are popularly used to model interactive behavior, such as innovation diffusion and cultural evolution. Typically, this involves determining the "better" of competing solutions. However, examples have demonstrated that different measures of a "good" choice can lead to conflicting conclusions; a fact that reflects the history of game theory in equilibrium selection. This behavior is totally explained while extending the analysis to the full seven-dimensional class of potential games, which includes coordination games.

Keywords: potential games; social welfare; risk dominance; payoff dominance; innovation diffusion; externalities; decomposition

1. Introduction

When Schelling (1960) wrote *Strategy of Conflict*, it pivoted attention from zero-sum games to the more general behavior allowed by games with mutually beneficial outcomes (which was appropriate during this Cold War period) [1]. Specifically, Schelling made a case for coordination games, which Lewis (1969) used to discuss culture and convention [2]. This behavioral notion of mutually beneficial outcomes was further explored by Rosenthal (1973) with his development of the congestion game [3]. Monderer and Shapley (1996) built on the congestion game with "common interest" games; namely the potential games (which include coordination games) [4]. More recently, Young (2006, 2011) and Newton and Sercombe (2020) took this analysis a step further by modeling, with potential games, how populations on networks evolve from one convention to another [5–7]. The natural question in this work is to discover whether the status quo or an innovation will be accepted.

This issue of finding the "better outcome" (e.g., an innovation or the status quo), which is a theme of this paper, is a fundamental and general concern for game theory; answers require selecting a measure of comparison. A natural choice is to prefer those outcomes where the players receive larger payoffs. Rather than payoff dominance, another refinement of Nash equilibria offered by Harsanyi and Selton (1988) is risk dominance [8].[1] The choice used by Young (2006, 2011) and later by Newton and Sercombe (2020) is to maximize social welfare. Still, other measures can be constructed [5–7].

With the social welfare measure, Young constructed a clever ad-hoc example where, although it is seemingly profitable to adopt the innovation, the innovation is worse than the status quo [5]. Young's observation underscores the important need to understand when and why a model's conclusions can change. This includes his concerns of identifying when and why a new cultural convention is "better". Is there a boundary between the quality of innovations?

[1] A payoff dominant Nash cell is where each agent does at least as well as in any other Nash cell, and at least one does better. A risk-dominant Nash cell is less costly should coordination be mistakenly expected; see Section 3.1.

Conflicting conclusions must be anticipated because different measures emphasize different traits of games. Thus, answering the "better" question requires determining which aspects of a game a given measure ignores or accentuates. The approach used here to address this concern is new; it appeals to a recent decomposition (or coordinate system) created for the space of games [9–12] (in what follows, the [9–11] papers are designated by [JS1], and the book [12] by [JS2]). There are many ways to decompose games, where the emphases reflect different objectives. An early approach was to express a game in terms of its zero sum and identical play components, which plays a role in the more recent Kalai and Kalai bargaining solution [13]. Others include examining harmonic and potential games [14] and strategic behavior such as in [15,16]. While some overlap must be expected, the material in [JS1] and [JS2] appears to be the first to strictly separate and emphasize Nash and non-Nash structures.

Indeed, in [JS1], [JS2], and this paper, appropriate aspects of a game are used to extract all information, and only this information, needed for the game's Nash structures; this is the game's "Nash" component. Other coordinate directions (orthogonal to, and hence independent of the Nash structures) identify those features of a game that require interaction among the players, such as coordination, cooperation, externalities, and so forth. By isolating the attributes that induce behavior among players, these terms define the game's "Behavioral" component. The final component, the kernel, is akin to adding the same value to each of a player's payoffs. While this is a valuable variable with transferable assets, or to avoid having games with negative payoffs, it plays no role in the analysis of most settings. In [JS2], the [JS1] decomposition is extended to handle more player and strategies.[2]

One objective of this current paper is to develop a coordinate system that is more convenient to use with a wide variety of choices that include potential games (a later paper extends this to more players and strategies). An advantage of using these coordinates is that they intuitively organize the strategic and payoff structures of all games. This is achieved by extracting from each payoff entry the behavioral portions that capture ways in which players agree or disagree (e.g., in accepting or rejecting an innovation) and affect payoff values. Of interest is how this structure applies to all 2×2 normal form games. When placing the emphasis on potential games, these coordinates cover their full seven-dimensional space, so they subsume the lower dimensional models in the literature.

By being a change of basis of the original decomposition, this system still highlights the unexpected facts that Nash equilibria and similar solution concepts (e.g., solution notions based on "best response" such as standard Quantal Response Equilibria) ignore nontrivial aspects of a game's payoff structure; see [11]. In fact, this is the precise feature that answers certain concerns in the innovation diffusion literature. Young's example [5], for instance, turns out to combine disagreement between two natural measures of "good" outcomes: one measure depends on unilateral deviations; the other aggregates the collective payoff. Newton and Sercombe re-parametrize Young's model to further explore this disagreement [7]. As we show, Young's example and the Newton and Sercombe arguments stem from a game's tension between group cooperative behavior and individualistic forces.

Other contributions of this current paper are to

- describe the payoff structure of these games;
- characterize the full seven-dimensional space of 2×2 potential games;
- analyze the behavioral tension between individual and cooperative forces in potential games;
- explain why different measures can reach different conclusions; and,
- relate the results to risk-dominance when possible.

---

[2] Experimental work has been done by Jessie and Kendall [17] by building on the decomposition in [JS1]. More precisely and as given in this paper, the separation aspect of the decomposition permits constructing large classes of games with an identical Nash component (or, the strategic component), but with wildly different externalities components (or, the behavioral component). As they showed, the choice of the behavioral term influenced an agent's selection. Section 2 discusses these components.

The paper begins with an overview of the coordinate system for all 2×2 normal form games. After identifying the source of all conflict with symmetric potential games, the full seven-dimensional class is described.

2. Overview of the Coordinate System

As standard with coordinate systems, the one developed for games in [JS1] can be adjusted to meet other needs. The choice given here [JS2] reflects central features of potential games.

Consider an arbitrary 2×2 game $\mathcal{G}$ (Table 1), where each agent's strategies are labeled $+1$ and -1 (cells also will be denoted by TL (top-left), BR (bottom-right), etc.).

Table 1. Arbitrary Game $\mathcal{G}$ in Normal-Form.

	$+1$		-1	
$+1$	a_1	b_1	a_3	b_2
-1	a_2	b_3	a_4	b_4

A weakness of this representation is captured by the Table 2 game. Information about which strategy each player prefers, whether they do, or do not want to coordinate with the other player, and where to find group opportunities is packed into the entries. Yet, in general, this structure is not readily available from the Table 2 form.

Table 2. An Example.

	$+1$		-1	
$+1$	-3	7	7	3
-1	-1	5	-1	-3

The coordinate system described here significantly demystifies a game by unpacking its valued behavioral information. This is done by decomposing a game into four orthogonal components, where each captures a specified essential trait: individual preference, individual coordinative pressures, pure externality (or Behavioral), and kernel components (see Table 3) (the orthogonality comment follows by identifying games with vectors). The original game is the sum of the four parts.

Table 3. The Four Components of a 2×2 Game.

	$+1$		-1	
$+1$	α_1	α_2	α_1	$-\alpha_2$
-1	$-\alpha_1$	α_2	$-\alpha_1$	$-\alpha_2$

a. *Individual Preference Component*

	$+1$		-1	
$+1$	γ_1	γ_2	$-\gamma_1$	$-\gamma_2$
-1	$-\gamma_1$	$-\gamma_2$	γ_1	γ_2

b. *Coordinative Pressure Component*

	$+1$		-1	
$+1$	β_1	β_2	$-\beta_1$	β_2
-1	β_1	$-\beta_2$	$-\beta_1$	$-\beta_2$

c. *Pure Externality Component*

	$+1$		-1	
$+1$	κ_1	κ_2	κ_1	κ_2
-1	κ_1	κ_2	κ_1	κ_2

d. *Kernel Component*

To associate these components with behavioral traits of any 2×2 game, the individual preference component identifies an agent's inherent preference for strategy $+1$ or -1. If $\alpha_i > 0$, then agent i prefers strategy $+1$ to -1 *independent* of what the other agent plays. In turn, $\alpha_i < 0$ means that agent i's individual preference is for strategy -1. The Table 2 values will turn out to be $\alpha_1 = 2, \alpha_2 = 3$, which indicates that both players prefer $+1$ (or TL).

The coordinative pressure component γ_j reflects a conforming stress a game imposes on agent j. Independent of the α_i sign, a $\gamma_j > 0$ value rewards agent j a positive payoff by conforming with agent i's preferred α_i choice. Conversely, when $\gamma_j < 0$, agent j's payoff is improved by playing a strategy different than what agent i wants. With Table 2, $\gamma_1 = -3$ while $\gamma_2 = -1$, so neither player is strategically supportive of personally reinforcing the other agent's preferred choice.

The pure externality component represents consequences that an agent's action imposes on the other agent.[3] If agent i plays $+1$, for instance, then, independent of what agent j does, agent j receives an extra β_j payoff! Acting alone, however, agent j is powerless to change this portion of the payoff. To see why this statement is true, should Column select L in Table 3c, then no matter what strategy Row chooses, this extra advantage remains β_1. The sign of β_j indicates which of agent i's strategies contributes to, or detracts from, agent j's payoff. In Table 2, the $\beta_1 = -3, \beta_2 = 2$ values convert TR into a potential *group* opportunity.

A subtle but important behavioral distinction is reflected by the γ and β terms. The γ_j values capture whether, in seeking a personally preferred (Nash) outcome, an agent should, or should not, coordinate with the other agent's preferred interests. In contrast, the β_j values identify externalities and opportunities to encourage both to cooperate. For a supporting story, suppose the strategies are to take highway 1 or -1 to drive to a desired location. A $\gamma_1 < 0$ value indicates the first agent's personal preference to avoid being on the same highway as the second. However, it should it be winter time, then the second agent, who always has a truck with a plow when driving on highway 1, creates a positive externality that can be captured with a β value.

The final component is the kernel, which for Table 2 is $\kappa_1 = 1, \kappa_2 = 3$. This can be treated as an inflationary term that adds the same κ_i value to each of the *ith* agent's payoffs. Methods that compare payoffs cancel the kernel value, so, as in this paper, the kernel can be ignored.

It is important to point out that the individual and coordinative pressure components contain all information from a game that is needed to compute the Nash equilibrium and to analyze related strategic solution concepts [JS1]. To appreciate why this is so, recall that the Nash information relies on payoff comparisons with unilateral deviations. But with the pure externality and kernel components, all unilateral payoff differences equal zero, so they contain no Nash information. This also means that "best response" solutions and methods ignore, and are not affected by the wealth of a game's β information (for explicit examples, see [11]).

By involving eight orthogonal directions and independent variables, these components span the eight-dimensional space of all 2×2 games. Consequently, any 2×2 game can be expressed and analyzed in terms of these eight coordinates. The equations converting a game into this form are

$$\begin{array}{lll} \alpha_1 = \frac{1}{4}[(a_1 + a_3) - (a_2 + a_4)], & \alpha_2 = \frac{1}{4}[(b_1 + b_3) - (b_2 + b_4)], & \gamma_1 = \frac{1}{4}[(a_1 + a_4) - (a_2 + a_3)], \\ \gamma_2 = \frac{1}{4}[(b_1 + b_4) - (b_2 + b_3)], & \kappa_1 = \frac{1}{4}[a_1 + a_2 + a_3 + a_4], & \kappa_2 = \frac{1}{4}[b_1 + b_2 + b_3 + b_4], \\ \beta_1 = \frac{1}{2}[a_1 + a_2] - \kappa_1, & \beta_2 = \frac{1}{2}[b_1 + b_2] - \kappa_2. & \end{array} \tag{1}$$

For interpretations, κ_j is agent j's average payoff, β_j is agent j's average payoff should the other agent play 1 minus the inflationary κ_j value, α_j is half the difference of agent j's average payoff if the other agent plays 1 and the average if the other agent plays -1, and γ_j is half the difference of the *jth* agent's average TL, BR payoff, and average BL and TR payoff.

To illustrate the derivation of Equation (1), the α_1 value of the Table 4a game is computed. All that is needed is a standard vector analysis to find how much of game $\mathcal{G}$ is in the α_1 coordinate direction, which is denoted by $\mathcal{G}^{\alpha_1}$, where $\alpha_1 = 1$ and $\alpha_2 = 0$. The sum of the squares of the $\mathcal{G}^{\alpha_1}$ entries (which in the following notation equals $[\mathcal{G}^{\alpha_1}, \mathcal{G}^{\alpha_1}]$) is $1^2 + 0^2 + 1^2 + 0^2 + (-1)^2 + 0^2 + (-1)^2 + 0^2 = 4$, so, according to vector analysis, $\alpha_1 = \frac{1}{4}[\mathcal{G}, \mathcal{G}^{\alpha_1}]$. Here, $[\mathcal{G}, \mathcal{G}^{\alpha_1}]$ is the sum of the products of

[3] This is the behavioral component in [JS1].

corresponding entries from each cell. (Identifying a game's payoffs with components of a vector in $\mathbb{R}^8$, $[\mathcal{G}_1, \mathcal{G}_2]$ is the scalar product of the vectors.) In this example, $[\mathcal{G}, \mathcal{G}^{\alpha_1}] = (12)(1) + (10)(0) + (2)(1) + (2)(0) + (0)(-1) + (4)(0) + (6)(-1) + (0)(0) = 8$,, so $\alpha_1 = \frac{1}{4}[\mathcal{G}, \mathcal{G}^{\alpha_1}] = 2$.. Similarly, by defining a corresponding $\mathcal{G}^{\alpha_2}, \mathcal{G}^{\gamma_1}, \ldots$,, the remaining values are $\alpha_2 = 3, \gamma_1 = 4, \gamma_2 = 1, \beta_1 = 1, \beta_2 = 2, \kappa_1 = 5$, and $\kappa_2 = 4$. The Equation (1) expressions can be recovered in this manner.

Table 4. Decomposing a game.

a. $\mathcal{G}$: *A special case*

	+1		−1	
+1	12	10	2	2
−1	0	4	6	0

b. $\mathcal{G}^{\alpha_1}$: *Where* $\alpha_1 = 1, \alpha_2 = 0$

	+1		−1	
+1	1	0	1	0
−1	−1	0	−1	0

This decomposition simplifies the analysis by extracting the portion from each payoff that contributes to these different attributes of a game. Illustrating with Table 1, rather than handling each entry separately, behavior can be analyzed with the separated impact of the components. For instance, the a_1 entry is $a_1 = \alpha_1 + \gamma_1 + \beta_1 + \kappa_1$, which, for Table 2, is $a_1 = -3 = 2 - 3 - 3 + 1$.

To connect this notation with [JS1], [JS2], a game's Nash component, denoted by $\mathcal{G}^N$, is the sum of the individual preference and coordinative pressure components as given in Table 5.

Table 5. The $\mathcal{G}^N$ Nash component.

	+1		−1	
+1	$\eta_{1,1} = \alpha_1 + \gamma_1$	$\eta_{1,2} = \alpha_2 + \gamma_2$	$\eta_{2,1} = \alpha_1 - \gamma_1$	$-\eta_{1,2} = -\alpha_2 - \gamma_2$
−1	$-\eta_{1,1} = -\alpha_1 - \gamma_1$	$\eta_{2,2} = \alpha_2 - \gamma_2$	$-\eta_{2,1} = -\alpha_1 + \gamma_1$	$-\eta_{2,2} = -\alpha_2 + \gamma_2$

Principal facts about 2×2 games follow.[4] As a reminder, a game is a potential game if there exists a global payoff function that aggregates the unilateral incentive structure of the game. More precisely, the payoff difference obtained by an agent unilaterally deviating is reflected in the change of the potential function. Potential games are often called "common interest" games. *Coordination games* have pure strategy Nash equilibria precisely where the agents play the same strategy (i.e., the strategy profiles $(+1, +1)$ and $(-1, -1)$). On the other hand, *anti-coordination games* have pure strategy Nash equilibria when the agents play different strategies (i.e., the strategy profiles $(+1, -1)$ and $(-1, +1)$).

Theorem 1. *Generically (that is, all* $\mathcal{G}^N$ *and* β_j *entries are nonzero), the following hold for* 2×2 *normal form games* $\mathcal{G}$.

1. *A* $\mathcal{G}$ *cell is pure Nash if and only if all of the cell's* $\mathcal{G}^N$ *entries are positive.*
2. $\mathcal{G}$ *is a potential game if and only if* $\gamma_1 = \gamma_2$.
3. $\mathcal{G}$ *is a coordination game if and only if* $|\gamma_1| > |\alpha_1|, |\gamma_2| > |\alpha_2|$ *and* $\operatorname{sgn} \gamma_1 = \operatorname{sgn} \gamma_2 = 1$. *If* $\operatorname{sgn} \gamma_1 = \operatorname{sgn} \gamma_2 = -1$, *then* $\mathcal{G}$ *is an anti-coordination game.*
4. *All of the payoffs in a normal-form game, as in Table 3, can be expressed with the utility functions for agents 1 and 2 given by, respectively,*

$$\pi_1(t_1, t_2) = \alpha_1 t_1 + \gamma_1 t_1 t_2 + \beta_1 t_2 + \kappa_1 \text{ and } \pi_2(t_1, t_2) = \alpha_2 t_2 + \gamma_2 t_1 t_2 + \beta_2 t_1 + \kappa_2,$$

   where t_1 *and* t_2 *represent the strategy choice (either* $+1$ *or* -1*) of agents 1 and 2,*
5. *A potential function for a game with components described in Table 3 can be transformed into*

$$P(t_1, t_2) = \alpha_1 t_1 + \alpha_2 t_2 + \gamma t_1 t_2, \text{ where } \gamma = \gamma_1 = \gamma_2. \tag{2}$$

[4] All of these results extend to $2 \times \ldots \times 2$ games.

6. *A potential game's potential function is invariant to the pure externality and kernel components.*

To explain certain comments, recall that a potential game has a potential function; if an agent changes a strategy, the change in the agent's payoff equals the change in the potential function. To illustrate, suppose the first agent changes from strategy $t_1 = 1$ to $t_1 = -1$, while agent 2 remains at $t_2 = 1$. According to Table 2, the change in the first agent's payoff is

$$[-\alpha_1 - \gamma_1 + \beta_1 + \kappa_1] - [\alpha_1 + \gamma_1 + \beta_1 + \kappa_1] = -2[\alpha_1 + \gamma_1],$$

or $-2[\alpha_1 + \gamma]$ for potential games. According to Equation (2), the change in the potential is the same value

$$P(-1, 1) - P(1, 1) = [-\alpha_1 + \alpha_2 + (1)(-1)\gamma] - [\alpha_1 + \alpha_2 + (1)(1)\gamma] = -2[\alpha_1 + \gamma].$$

Statement 1 is proved in [JS1]. To prove the second assertion, in (Chap. 2 of [JS2]) it is shown that game $\mathcal{G}$ is a potential game if and only if it is orthogonal to the 2×2 matching pennies game $\mathcal{G}^{pennies}$; [5]this orthogonality condition is $[\mathcal{G}, \mathcal{G}^{pennies}] = 0$. A direct computation shows that the individual preference, pure externality, and kernel components always satisfy this condition. The coordinative pressure component satisfies the condition if and only if $\gamma_1 = \gamma_2$.

Statement 4 is a direct computation. Statement 5 is a direct computation showing that changes in an agent's strategy have the same change in the potential function as in the player's payoff. Statement 6 follows, because $P(t_1, t_2)$ (Equation (2)) does not have β or κ values. A proof of the remaining statement 3 is in [JS2]. For intuition, the players in a coordination game coordinate their strategies, which is the defining feature of the coordinative pressure component. Thus, for $\mathcal{G}$ to be a coordination game, the γ values must dominate the game's Nash structure.

3. Disagreement in Potential Games

These coordinates lead to explanations why different behavioral measures can differ about which is the "better" outcome for certain games. This discussion is motivated by innovation diffusion, which is typically modeled by using coordination games with two equilibria, so a key step is to identify the preferred equilibria. A common choice is the risk-dominant pure Nash equilibria. In part, this is because these equilibria have been connected to the long-run behavior of dynamics, such as log-linear learning [18]. Because a coordination game is a potential game, the potential function's global maximum is the risk-dominant equilibrium (Theorem 2).

However, as developed next, there are many games where a maximizing strategy for the potential function differs from the profile that maximizes social welfare. This difference is what allows agents to do "better" by using a profile other than the one leading to the risk-dominant equilibrium. Young [5,6] creates such an example using the utilitarian measure of social welfare, which sums all of the payoffs in a given strategy profile, and Newton and Sercombe [7] discuss similar ideas. A first concern is whether their examples are sufficiently isolated that they can be ignored, or whether they are so prevalent that they must be taken seriously. As we show, the second holds.

An explanation for what causes these conflicting conclusions emerges from the Table 2 decomposition and Theorem 1. A way to illustrate these results is to create any number of new, more complex examples. To do so, start with the fact (Theorem 2.5 in [JS2]) that with 2×2 games and two Nash equilibria, a Nash cell is risk dominant over the other Nash cell if and only if the product of the η values (see Table 4) of the first is larger than the product of the η values of the second. (To handle

[5] This makes sense; "matching pennies" is antithetical (orthogonal) to the cooperative spirit of potential games. The space of matching pennies is the harmonic component discussed in [14].

some of our needs, this result is refined in Theorem 2.) Of significance is that, although the β terms obviously affect payoff values, they play no role in this risk-adverse dominance analysis.

To illustrate, let $\gamma_1 = \gamma_2 = 4$ (Theorems 1-2; to have a potential game) and $\alpha_1 = \alpha_2 = 1$ (Theorem 1-3; to have a coordination game). This defines the Table 6a game where the TL Nash cell (with $(\eta_{1,1})(\eta_{1,2}) = 25$) is risk dominant over the BR Nash cell (with $(-\eta_{1,1})(-\eta_{1,2}) = 9$). There remain simple ways to modify this game that make the payoffs of any desired cell, say BR, more attractive than the risk dominant TL. All that is needed is to select β_j values that increase the payoffs for the appropriate Table 3c cell; for the BR choice, this requires choosing negative β_1, β_2 values (Table 3c). Although these values never affect the risk-dominant analysis, they enhance each player's BR payoff while reducing their TL payoffs. The $\beta_1 = \beta_2 = -3$ choices define the Table 6b game where each player receives a significantly larger payoff from BR than from the risk dominant TL! In both games, the Nash and risk dominance structures remain unchanged.

Table 6. Conflicting measures.

	+1		−1	
+1	5	5	−3	−5
−1	−5	−3	3	3

a. *First example*

	+1		−1	
+1	2	2	0	−8
−1	−8	0	6	6

b. *Conflicting behavior*

These coordinates make it possible to easily create a two-dimensional family of games with such properties. To do so, add Table 3c to the Table 6a game, and then select appropriate β_1, β_2 values to emphasize the payoffs of different cells. For instance, using Young's welfare measure (the sum of a cell's payoffs), no matter which cell is selected, suitable β values exist to make that cell preferred to TL. It follows from the Table 3c structure, for instance, that a way to enhance the TR payoffs is to use $\beta_1 < 0$ and $\beta_2 > 0$ choices. Adding these values to the Table 6a game defines the TR cell values of $-3 + |\beta_1|$ and $-5 + \beta_2$ while the TL values are $5 - |\beta_1|$ and $5 + \beta_2$. Thus, the sum of TR cell values dominates the sum of TL values if and only if

$$(-3 + |\beta_1|) + (-5 + \beta_2) > (5 - |\beta_1|) + (5 + \beta_2), \text{ or iff } -\beta_1 = |\beta_1| > 9. \tag{3}$$

$$\begin{array}{|cc|cc|}\hline 10 & 11 & 2 & 7 \\ \hline 8 & 4 & 8 & 6 \\ \hline \end{array} = \begin{array}{|cc|cc|}\hline 1 & 2 & -3 & -2 \\ \hline -1 & -1 & 3 & 1 \\ \hline \end{array} + \begin{array}{|cc|cc|}\hline 2 & 2 & -2 & 2 \\ \hline 2 & -2 & -2 & -2 \\ \hline \end{array} + \begin{array}{|cc|cc|}\hline 7 & 7 & 7 & 7 \\ \hline 7 & 7 & 7 & 7 \\ \hline \end{array}. \tag{4}$$

Conflict among 50-50, payoff, and risk dominance

The coordinates also make it possible to compare other measures by mimicking the above approach. Games with payoff dominant strategies that differ from the risk adverse ones, for example, require appropriate β values. To explain, if BR is risk dominant, then, as in the Equation (4) game (from Section 2.6.4 in [JS2]), the product of its η values from BR in $\mathcal{G}^N$ (first bimatrix on the right) is larger than the product of the TL $\mathcal{G}^N$ values. This product condition ensures that the only way the payoff and risk dominant cells can differ is by introducing TL β components; this is illustrated with $\beta_1 = \beta_2 = 2$ in the second bimatrix in Equation (4). More generally and using just elementary algebra as in Equation (3), the regions (in the space of games) where the two concepts differ now can finally be determined.

For another measure, consider the 50–50 choice. This is where, absent any information about an opponent, it seems reasonable to assume there is a 50–50 chance the opponent will select one Nash cell over the other. This assumption suggests using an expected value analysis to identify which strategy a player should select. To discover what coordinate information this measure uses, if TL and BR are the two Nash cells, then for Row and this 50–50 assumption, the expected return from playing T is $\frac{1}{2}[|\eta_{1,1}| - |\eta_{2,1}|] + \frac{1}{2}[\beta_1 - \beta_1] + \frac{1}{2}[\kappa_1 + \kappa_1] = \frac{1}{2}[|\eta_{1,1}| - |\eta_{2,1}|] + \kappa_1$. Similarly, the expected value of playing B is $\frac{1}{2}[-|\eta_{1,1}| + |\eta_{2,1}|] + \frac{1}{2}[\beta_1 - \beta_1] + \frac{1}{2}[\kappa_1 + \kappa_1] = \frac{1}{2}[-|\eta_{1,1}| + |\eta_{2,1}|] + \kappa_1$. Consequently, an agent's larger η value completely determines the 50-50 choice. However, if the risk adverse cell of $\mathcal{G}^N$ is not also $\mathcal{G}^N$ payoff dominant, as true with the first bimatrix on the right in Equation (4), and if both players adopt the 50-50 measure, they will select a non-Nash outcome. Indeed, in Equation (4), BR is risk dominant, TL is payoff dominant, and BL, which is not a Nash cell (and Pareto inferior to both Nash cells), is the 50–50 choice. Again, elementary algebra of the Equation (3) form identifies the large region of games where this behavior can arise. (By using appropriate β values, it is easy to create 50–50 outcomes that are disastrous.)

3.1. The Potential and Welfare Functions

Moving beyond examples, these coordinates can fully identify those games for which different measures agree or disagree, which is one of the objectives of this paper. The importance of this analysis is that it underscores our earlier comment that this conflict between the conclusions of measures is a fundamental concern that is suffered by a surprisingly large percentage of all 2×2 games.

Our outcomes are described using maps of the space of 2×2 games. The maps show where the potential and social welfare functions (e.g., the ones used by Young [5,6] and Newton and Sercombe [7]) agree, or disagree, on which is the "better" choice of two equilibria. Not only do these diagrams demonstrate the preponderance of this conflict, but they identify which behavior a specific game will experience. As an illustration, the dark regions of Figure 1 single out those potential games (so $\gamma_1 = \gamma_2$), where $\alpha_1 = \alpha_2 = \alpha$ and $\beta_1 = \beta_2 = \beta$, which are without conflict; for games in the unshaded regions, different measures support different outcomes.

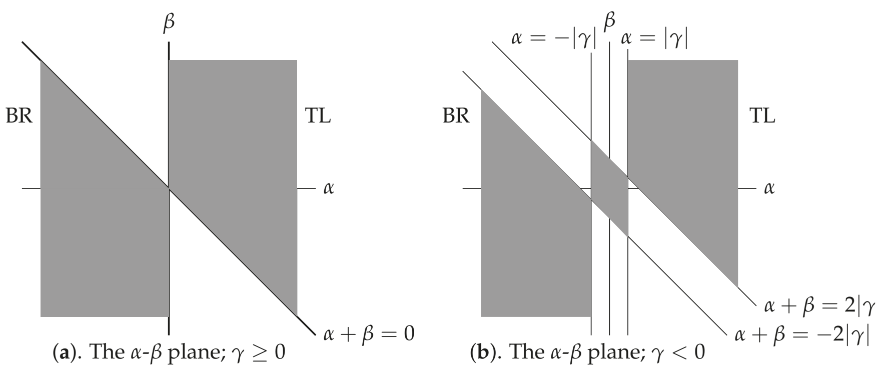


Figure 1. Conflict and agreement with α, β values.

Thanks to the coordinate system for games, the game theoretic analysis is surprisingly simple; it merely uses a slightly more abstract version of the Equation (3) analysis. To illustrate with the above 50-50 discussion, if the Nash cells are TL and BR, then $\eta_{1,1}, -\eta_{2,1}, \eta_{1,2}, -\eta_{2,2}$ are all positive (if the Nash cells are BL and TR, all of these entries are negative). Consequently, the two surfaces $\eta_{1,1} + \eta_{2,1} = 0$ and $\eta_{1,2} + \eta_{2,2} = 0$ separate which one of an agent's η values is larger. Even though the discussion applies to the four dimensional space of η values, one can envision the huge wedges these surfaces define where the η values force the 50–50 approach to select a non-Nash outcome.

A similar approach applies to all of the maps derived here. In Figure 1a, the dividing surface separating which Nash cell is selected by potential function outcomes is $\alpha = 0$; if $\alpha > 0$, the game's top Nash choice for the potential function is TL; if $\alpha < 0$, the top Nash choice is BR. However, the social welfare conclusion is influenced by β values, so it will turn out that the separating line between a social welfare function selecting TL or BR is the $\alpha + \beta = 0$ line. Above this line, TR is the preferred Nash choice; below it is BR. Given this legend, Figure 1a demonstrates those games for which the different measures agree or disagree about the top choices, and the magnitude of the problem. Stated simply, regions that emphasize behavioral terms place emphasis on payoff and social welfare dominant measures; regions that emphasize Nash strategic terms emphasize risk dominant measures.

Stated differently, difficulties in what follows do not reflect the game theory; the coordinate system handles all of these problems. Instead, all of the complications (there are some) reflect the geometric intricacy of the seven-dimensional space of 2×2 potential games. Consequently, readers that are interested in applying this material to specific games should emphasize the maps and their legends (given in the associated theorems). Readers that are interested in the geometry of the space of potential functions will find the following technical analysis of value. However, first, a fundamental conclusion about potential games is derived.

3.2. A Basic "Risk Dominant" Theorem

The coordinates explicitly display a tension between what individuals can achieve on their own (Nash behavior) and with cooperative forces. With a focus on individualistic forces, the potential function is useful because its local maxima are pure Nash equilibria. Even more, as known, the potential's global maximum is the risk dominant equilibrium. This fact is re-derived for 2×2 potential games in a manner that now highlights the roles of a potential game's coordinates.

Theorem 2. *For a 2×2 potential game $\mathcal{G}$, its potential function has a global maximum at the strategy profile (t', t'') if and only if (t', t'') is $\mathcal{G}$'s risk-dominant Nash equilibrium. With two Nash equilibria where one is risk dominant, (t', t'') is the risk dominant strategy if and only if the following inequalities hold*

$$\textbf{a}.\,(\alpha_1 t' + \alpha_2 t'') > 0, \quad \textbf{b}.\,|\gamma| > |\alpha_1|, |\alpha_2|, \quad \textbf{c}.\,t't''\gamma > 0. \tag{5}$$

As shown in the proof, inequality c identifies the Nash cells; e.g., if $\gamma < 0$, then $t't'' = -1$, so the Nash cells are BL and TR. With two Nash cells, the inequality a identifies which one is a global maximum of the potential function. Similarly, inequality b requires γ to have a sufficiently large value to create two Nash cells. Of importance, Equation (5) does not include β values!

To illustrate these inequalities, let $\alpha_1 = 1, \alpha_2 = -2$, and $\gamma = 3$. By satisfying Equation (5b), there are two Nash cells. According to Equation (5c), $t't'' = 1$, so $t' = t''$, which positions the Nash cells at TL and BR. From Equation (5a), $t' - 2t' > 0$, or $t' < 0$, so the risk dominant strategy is the $t' = t'' = -1$ BR cell. Conversely, to create an example where a desired cell, say TR, is risk dominant, the $t' = 1, t'' = -1$ values require (Equation (5c)) $\gamma < 0$ and (Equation (5a)) $\alpha_1 > \alpha_2$. Finally, select $\gamma < 0$ that satisfies Equation (5b); e.g., $\alpha_1 = 1, \alpha_2 = -1$ and $\gamma = -2$ suffice.

The Equation (5) inequalities lead to the following conclusion.

Corollary 1. *If a 2×2 potential game $\mathcal{G}$ has two pure Nash equilibria where one is risk dominant, then $\mathcal{G}$ is a coordination game. If $\gamma > 0$, the Nash cells are at TL and BR, where $\mathcal{G}$ is a coordination game. If $\gamma < 0$, the Nash cells are at BL and TR, where $\mathcal{G}$ is an anti-coordination game.*[6]

Proof of Corollary 1. According to Theorem 2, the Corollary 1 hypothesis ensures that Equation (5) hold. With the b inequality, it follows from Theorem 1-3 that $\mathcal{G}$ is a coordination game. In a 2×2 game, pure Nash cells are diagonally opposite. If $\gamma > 0$, it follows from Equation (5), c that the Nash strategies satisfy $t't'' = 1$, so the Nash cells are at TL (for $t' = t'' = 1$) and BR (for $t' = t'' = -1$), and that this is a coordination game. Similarly, if $\gamma < 0$, then $t't'' = -1$, so the Nash cells are at BL (for $t' = -1, t'' = 1$) and TR (for $t' = 1, t'' = -1$) to define an anti-coordination game. □

[6] As pointed out by a referee, the case where $\gamma < 0$ appears to be related to the notion of self-defeating externalities, making the potential game in this case a stable game, as defined in [19].

Proof of Theorem 2. In a non-degenerate case (i.e., $P(t_1, t_2)$ is not a constant function), P has a maximum, so there exists at least one pure Nash cell. If a game has a unique pure strategy Nash equilibrium, then, by default, it is risk-dominant and P's unique maximum.

Assume there are two Nash cells; properties that the potential, P, must satisfy at a global maximum are derived. Pure Nash equilibria must be diametrically opposite in a normal form representation, so if $\mathcal{G}$ has two pure strategy Nash equilibria where one is (t', t''), then the other one is at $(-t', -t'')$. Consequently, if P has a global maximum at (t', t''),, then P has a local maximum at $(-t', -t'')$, so $P(t', t'') > P(-t', -t'')$. According to Equation (2), this inequality holds iff $[\alpha_1 t' + \alpha_2 t'' + \gamma(t')(t'')] - [\alpha_1(-t') + \alpha_2(-t'') + \gamma(-t')(-t'')] = 2[\alpha_1 t' + \alpha_2 t''] > 0$, which is inequality Equation (5)a.

The local maximum structure of $P(-t', -t'')$ requires that $P(-t', -t'') > P(t', -t'')$ and $P(-t', -t'') > P(t', -t'')$. Again, according to Equation (2), the first inequality is true if

$$\alpha_1(-t') + \alpha_2(-t'') + \gamma(-t')(-t'') > \alpha_1 t' + \alpha_2(-t'') + \gamma(t')(-t''),$$

or $\gamma t' t'' > \alpha_1 t'$. Similarly, the second inequality is true iff $\gamma t' t'' > \alpha_2 t''$. Thus, for a potential game with two Nash cells, P has a global maximum at (t', t'') iff

$$[\alpha_1 t' + \alpha_2 t''] > 0, \quad \gamma t' t'' > \alpha_1 t', \quad \gamma t' t'' > \alpha_2 t'', \quad \gamma t' t'' > 0. \tag{6}$$

The last inequality follows from the first one, which requires at least one of $\alpha_1 t', \alpha_2 t''$ to be positive. Thus, the $\gamma t' t'' > 0$ inequality follows from either the second or third inequality.

All that is needed to establish the equivalence of the Equation (6) inequalities and those of Equation (5) is that Equation (5b) is equivalent to the two middle inequalities of Equation (6). Equation (5b) implies the two middle inequalities of Equation (6) is immediate. In the opposite direction, the first Equation (6) inequality requires at least one of $\alpha_1 t'$ or $\alpha_2 t''$ to be positive. If it is $\alpha_j t$, then because $|\gamma| = \gamma t' t''$, this positive term requires the appropriate middle inequality of Equation (6) to be $|\gamma| > |\alpha_j|$. If it holds for both terms, the proof is completed. If it holds for only one term, say $\alpha_1 t' > 0$, but $\alpha_2 t'' < 0$, then the first Equation (6) inequality requires that $|\alpha_1| > |\alpha_2|$, which completes the proof.

The second step requires showing that (t', t'') is a risk-dominant Nash equilibrium if Equation (5) holds. According to Harsanyi and Selten (1988), (t', t'') is a game's risk-dominant Nash equilibrium if

$$(P(-t',t'')-P(t',t''))(P(t',-t'')-P(t',t''))>(P(t',-t'')-P(-t',-t''))(P(-t',t'')-P(-t',-t'')), \tag{7}$$

which is $(-2\alpha_1 t' - 2\gamma t' t'')(-2\alpha_2 t' - 2\gamma t' t'') > (2\alpha_1 t' - 2\gamma t' t'')(2\alpha_2 t' - 2\gamma t' t'')$. This inequality reduces to

$$\gamma t' t''(\alpha_1 t' + \alpha_2 t'') > 0. \tag{8}$$

If $t' t'' = 1$, the Nash cells are at TL and BR, so both entries of these two Table 4 cells must be positive (Theorem 1-1). For the TL cell, this means that $\gamma > -\alpha_1, -\alpha_2$,, while for the BR cell it requires $\gamma > \alpha_1, \alpha_2$. Consequently, $\gamma > |\alpha_1|, |\alpha_2|$ and $\gamma > 0$: inequalities b and c of Equation (5) are satisfied. That inequality a of Equation (5) holds follows from $\gamma t' t'' > 0$ and Equation (8).

Similarly, if $t' t'' = -1$, then BL and TR are the Nash cells. Again, each entry of each of these Table 4 cells must be positive: from BL, we have that $-\alpha_1, \alpha_2 > \gamma$,, while from TR we have that $\alpha_1, -\alpha_2 > \gamma$. Consequently, $\gamma < 0$, $\gamma t' t'' > 0$ and $|\gamma| > |\alpha_1|, |\alpha_2|$; these are inequalities b and c of Equation (5). That inequality a holds again follows from Equation (8). □

A consequence of Theorem 2 is that the potential function can serve as a comparison measure of Nash outcomes. Other natural measures reflect the overall well-being of all agents, such as the utilitarian social welfare function that sums each strategy profile's payoffs. To obtain precise conclusions, our results use this social welfare function. However, as indicated later, everything extends to several other measures.

3.3. Symmetric Games

Without question, it is difficult to understand the structures of a four-dimensional object, leave alone the seven-dimensions of the space of 2×2 potential games. Thus, to underscore the ideas, we start with the simpler (but important) symmetric games (all symmetric games are potential games); doing so reduces the dimension of the space of games from seven to four (with the kernel). This is where $\alpha_1 = \alpha_2 = \alpha$, $\beta_1 = \beta_2 = \beta$, $\gamma_1 = \gamma_2 = \gamma$, and $\kappa_1 = \kappa_2 = \kappa$. Ignoring the kernel term, the coordinates are given in Table 7.

Table 7. Decomposition for a symmetric potential game.

a. *Nash terms*

	+1		−1	
+1	$\alpha+\gamma$	$\alpha+\gamma$	$\alpha-\gamma$	$-\alpha-\gamma$
−1	$-\alpha-\gamma$	$\alpha-\gamma$	$-\alpha+\gamma$	$-\alpha+\gamma$

b. *Behavioral, externality terms*

	+1		−1	
+1	β	β	$-\beta$	β
−1	β	$-\beta$	$-\beta$	$-\beta$

To interpret Table 7, the Nash entries combine terms, where each agent prefers a particular strategy independent of what the other agent does (if $\alpha > 0$, each agent prefers to play 1; if $\alpha < 0$ each agent prefers -1) with terms that impose coordination pressures. That is, if $\gamma > 0$, the game's Nash structure inflicts a pressure on both agents to coordinate; if $\gamma < 0$, the game's joint pressure is for anti-coordination. With symmetric games, the payoffs for both agents agree at TL and at BR, so if one of these cells is social welfare top-ranked (the sum of the entries is larger than the sum of entries of any other cell), the cell also has the stronger property of being payoff dominant.

All 2×2 symmetric games can be expressed with these four $\alpha, \beta, \gamma, \kappa$ variables (Equation (1)). Ignoring κ, the remaining variables define a three-dimensional system. To tie all of this in with commonly used symmetric games, if BR is the sole Nash point, then the above defines a Prisoner's Dilemma game if $-\alpha + \gamma > 0$ and $\alpha + \gamma + \beta > -\alpha + \gamma - \beta$, the second of which reduces to $\beta + \alpha > 0$. Similarly, the Hawk–Dove game with Nash points at BL and TR, and "hawk" strategy as $+1$, has $-\alpha - \gamma > 0$ and $\alpha - \gamma > 0$ for the Nash points, and $\beta < 0$ to enhance the BR payoffs, so $\gamma < 0$, $\alpha < -|\gamma|$, and $\beta < 0$. A coordination game simply has $\gamma > |\alpha|$ (these inequalities allow the different games to easily be located in Figure 1 and elsewhere). As stated, because $\gamma_1 = \gamma_2$, it follows (Theorem 1-2) that all 2×2 symmetric games are potential games with the potential function (Equation (2))

$$P(t_1, t_2) = \alpha(t_1 + t_2) + \gamma t_1 t_2 \tag{9}$$

A computation (Table 7) proves that the social welfare function (the sum of a cell's entries) is

$$w(t_1, t_2) = (\alpha + \beta)(t_1 + t_2) + 2\gamma t_1 t_2. \tag{10}$$

Our goal is to identify those games for which the potential and social functions agree, or disagree, on the ordering of strategies. Here, the following results are useful.

Theorem 3. *A 2×2 symmetric potential game $\mathcal{G}$ satisfies*

$$w(t_1, t_2) = 2P(t_1, t_2) + (\beta - \alpha)(t_1 + t_2). \tag{11}$$

If $\alpha = \beta$, the potential function and the welfare function rankings of $\mathcal{G}$'s four cells agree.

Both the potential and welfare functions are indifferent about the ranking of the BL and TR cells, denoted as BL $\sim$ TR. If one of these cells is a Nash cell for $\mathcal{G}$, then so is the other one, but neither is risk dominant.

Equation (11) explicitly demonstrates that β values—the game's behavioral or externality values—are solely responsible for all of the differences between how the potential and social welfare functions rank $\mathcal{G}$'s cells.

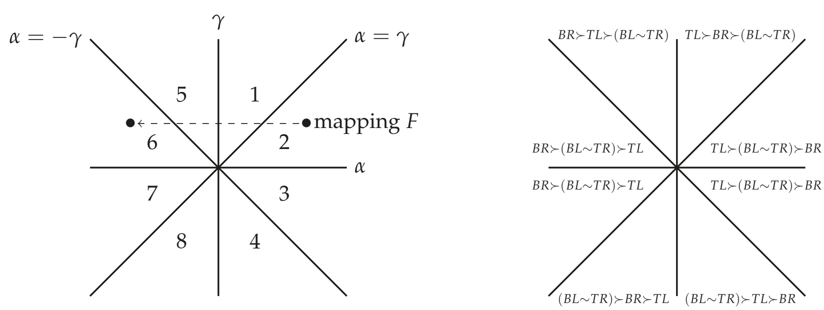
Proof. Adding and subtracting $\alpha(t_1 + t_2)$ to Equation (10) leads to Equation (11). Thus, if $\alpha = \beta$, then $\frac{1}{2}w(t_1, t_2) = P(t_1, t_2)$, so both functions have the same ranking of $\mathcal{G}$'s four cells.

The BL and TR cells correspond, respectively, to $(t_1 = -1, t_2 = 1)$ and $(t_1 = 1, t_2 = -1)$, where $t_1 + t_2 = 0$. Thus, (Equation (11)), the $\frac{1}{2}w$, and P values for each of these cells is γ. As both measures have the same value for each cell, both have a tie ranking for the two cells denoted by $BL \sim TR$.

The Nash entries of BL and TR are the same but in different order (Table 6a). Thus, if both entries of one of these cells are positive, then so are both entries of the other. This means that both are Nash cells (Theorem 1-1). The risk dominant assertion follows because inequality a of Equation (5) is not satisfied; it equals zero. Equivalently, P does not have a unique global maximum in this case. □

3.3.1. A Map of Games and Symmetries

A portion of a map that describes the structure of all 2 × 2 games (Chapter 3, [JS2]) is expanded here to concentrate on the symmetric games. While variables α, γ, β require a three-dimensional space, the potential function does not depend on β, so Figure 2a highlights the α-γ plane. Treat the positive β axis as coming out orthogonal to the page.



(**a**). The α-γ plane

(**b**). The P rankings

Figure 2. The α, β, γ structures.

Changes in the potential game and P (Equation (9)) depend on $\alpha, \gamma, \gamma - \alpha$, and $\gamma + \alpha$ values, which suggests dividing the α-γ plane into sectors where these terms have different signs. That is, divide the plane into eight regions (Figure 2a) with the lines $\alpha = 0, \gamma = 0, \alpha = \gamma$, and $-\alpha = \gamma$. The first two lines represent changes in a game's structure by varying the α and γ signs. For instance, reversing the sign of α changes which strategy the agents prefer; swapping the γ sign exchanges a game's coordination, anti-coordination features. The other two lines are where certain payoffs (Table 6a) change sign, which affects the game's Nash structure. The labelling of the regions follows:

$$\begin{array}{llllllll} (1) & \gamma > \alpha > 0 & (2) & \alpha > \gamma > 0 & (3) & \alpha > -\gamma > 0 & (4) & -\gamma > \alpha > 0 \\ (5) & \gamma > -\alpha > 0 & (6) & -\alpha > \gamma > 0 & (7) & -\alpha > -\gamma > 0 & (8) & -\gamma > -\alpha > 0 \end{array}$$

A natural symmetry simplifies the analysis. In Table 6, interchanging each matrix's rows and then columns creates an equivalent game, where the t', t'' cell in the original becomes the $-t', -t''$ cell in the image. This equivalent game is identified in Figure 2a with the mapping

$$F(\alpha, \gamma, \beta) = (-\alpha, \gamma, -\beta). \tag{12}$$

Geometrically, F flips a point on the right (a symmetric potential game) about the γ axis to a point on the left (which corresponds to the described changing of rows and columns of the original game); e.g., in Figure 2a, the bullet in region 2 is flipped to the bullet in region 6. Similarly, the original β value is flipped to the other side of the α-γ plane. Consequently, anything stated about a (t', t'') strategy or cell for a game in region k holds for a $(-t', -t'')$ strategy or cell of the corresponding game in region $(k+4)$. Thanks to this symmetry, by determining the P ranking of region k to the right of the γ axis, the P ranking of region $k+4$ is also known.

The following theorem describes each region's P ranking. The reason the decomposition structure simplifies the analysis is that all of the comments about Nash cells follow directly from Table 6a. If, for instance, $\gamma > \alpha > 0$ (region 1), then only cells TL and BR have all positive entries (Table 6a), so they are the only Nash cells. Similarly, if $-\gamma > \alpha > 0$ (region 4), only cells BL and TR have all positive entries, so they are the Nash cells (Theorem 1-1). Each cell's P value is specified in matrix Table 8a, so the P ranking of the cells follows immediately. Region 2, for instance, has $\alpha > \gamma > 0$, so TL is P's top-ranked cell. Whether BL is P-ranked over BR holds (Table 8a) iff $-\gamma > \gamma - 2\alpha$, or $\alpha > \gamma$, which is the case. This leads to P's ranking of $TL \succ (BL \sim TR) \succ BR$ for all region 2 games (here, $A \succ B$ means A is ranked over B and $A \sim B$ means they are tied). Each cell's $\frac{1}{2}w$'s value (half the social welfare function), which comes from Equation (11), is given in Table 8b.

Table 8. A symmetric potential game's P and $\frac{w}{2}$ values.

a. *P values*

	$+1$	-1
$+1$	$2\alpha+\gamma$	$-\gamma$
-1	$-\gamma$	$\gamma-2\alpha$

b. $\frac{w}{2}$ *values*

	$+1$	-1
$+1$	$\gamma+[\alpha+\beta]$	$-\gamma$
-1	$-\gamma$	$\gamma-[\alpha+\beta]$

Theorem 4. *The following hold for a* 2×2 *symmetric potential game.*

1. *Region 1 has Nash equilibria at TL and BR, where TL is risk dominant. The P ranking of the cells is* $TL \succ BR \succ (BL \sim TR)$*. The region 5 P ranking is* $BR \succ TL \succ (BL \sim TR)$*; BR is risk dominant.*
2. *Regions 2 and 3 have a single Nash cell at TL, where, for each region, the P ranking of the cells is* $TL \succ (BL \sim TR) \succ BR$*. The P ranking in regions 6 and 7 is* $BR \succ (BL \sim TR) \succ TL$*.*
3. *Region 4 has Nash cells at BL and TR, where neither is risk dominant. The P ranking is* $(BL \sim TR) \succ TL \succ BR$*. The region 8 P ranking is* $(BL \sim TR) \succ BR \succ TL$*.*

The content of this theorem is displayed in Figure 2b. Notice, with $\alpha > 0$, the potential function has BR bottom ranked unless BR is a Nash cell. This makes sense; $\alpha > 0$ means (Table 3a) that both agents prefer a "+1" strategy, so they prefer T and L. In fact, with $\alpha > 0$, the only way TL loses its top-ranked P status is with a sufficiently strong negative γ value (region 4 of Figure 2a). This also makes sense; a negative γ value (Table 3b) captures the game's anti-coordination flavor, which, if strong enough, can crown BL and TR as Nash cells.

Similar comments hold for $\alpha < 0$; this is because TL and BR reverse roles in the P rankings (properties of F (Equation (12)). Thus, the P ranking of region 1 is $TL \succ BR \succ (BL \sim TR)$, so the P ranking of region 5 is $BR \succ TL \succ (BL \sim TR)$. Accordingly, for $\alpha < 0$, P always bottom-ranks TL unless TL is a Nash cell, which reflects that $\alpha < 0$ is where the players have a preference for B and R.

The next theorem describes where the potential and social welfare function rankings agree or disagree. As its proof relies on Table 8b values, it is carried out in the same manner as for Theorem 4. Namely, to determine whether TL is ranked above BL or TR, it must be determined (Table 8b) whether $\gamma + (\alpha+\beta) > -\gamma$, or whether $2\gamma > \alpha + \beta > 0$. Next, according to (Table 8b), the social welfare function ranks BR above BL (or TR) iff $\gamma - (\alpha+\beta) > -\gamma$, or $2\gamma > \alpha+\beta > 0$. Thus, if $2\gamma > \alpha+\beta > 0$, the social welfare ranking is $TL \succ BR \succ (BL \sim TR)$.

Theorem 5. *For a* 2×2 *symmetric potential game with a coordinative flavor of* $\gamma > 0$*, the social welfare function (Equation (10)) ranks the cells in the following manner:*

1. *If* $(\alpha + \beta) > 2\gamma$*, the ranking is* $TL \succ (BL \sim TR) \succ BR$ *and* TL *is payoff dominant,*
2. *if* $2\gamma > (\alpha + \beta) > 0$*, the ranking is* $TL \succ BR \succ (BL \sim TR)$ *and* TL *is payoff dominant,*
3. *if* $0 > (\alpha + \beta) > -2\gamma$*, the ranking is* $BR \succ TL \succ (BL \sim TR)$ *and* BR *is payoff dominant, and*
4. *if* $-2\gamma > (\alpha + \beta)$*, the ranking is* $BR \succ (BL \sim TR) \succ TL$ *where* BR *is payoff dominant.*

For games with an anti-coordinative flavor of $\gamma < 0$*, the social welfare rankings are*

5 *if* $(\alpha + \beta) > -2\gamma$*, the ranking is* $TL \succ (TR \sim BL) \succ BR$ *and* TL *is payoff dominant;*
6 *if* $-2\gamma > (\alpha + \beta) > 0$*, the ranking is* $(BL \sim TR) \succ TL \succ BR$*,*
7 *if* $0 > (\alpha + \beta) > 2\gamma$*, the ranking is* $(BL \sim TR) \succ BR \succ TL$*, and*
8 *if* $0 > 2\gamma > (\alpha + \beta)$*, the ranking is* $BR \succ (BL \sim TR) \succ TL$ *and* BR *is payoff dominant.*

As with Theorem 4, this theorem ignores certain equalities, such as $\alpha + \beta = 2\gamma > 0$, but the social welfare ranking is the obvious choice. This equality captures the transition between parts 1 and 2, so its associated ranking is $TL \succ BR \sim (BL \sim TR)$.

A message of these theorems is to anticipate strong differences between the potential and social welfare rankings. Each game in region 1 of Figure 2a, for instance, has the unique P ranking of $TL \succ BR \succ (BL \sim TR)$, so TL is P's "best" cell. In contrast, each game in each region has four different social welfare rankings[7] where most of them involve ranking conflicts! In region 1 of Figure 2a, for instance, an admissible social welfare ranking (Theorem 5-4) is $BR \succ (BL \sim TR) \succ TL$, where the payoff dominant BR is treated as being significantly better than P's top choice of TL.

The coordinates explicitly identify why these differences arise: The potential function ignores β, while the β values contribute to the social welfare rankings. By influencing a game's payoffs and identifying (positive or negative) externalities that players can impose on each other, the β values constitute important information about the game. To illustrate, Table 5 has two different symmetric games; they differ only in that the first game has $\beta = 0$ with no externalities while the second has $\beta = -3$, which is a sizable externality favoring BR payoffs (Table 6b). Both of the games are in region 1 of Figure 2, so both have the same P ranking of $TL \succ BR \succ (BL \sim TR)$ (Theorem 4-1), where TL is judged the better of the two Nash cells. However, the social welfare ranking for the Table 5b game is $BR \succ TL \succ (BL \sim TR)$, which disagrees with the P ranking by crowning BR as the superior cell. By examining this Table 5b game, which includes externality information, it would seem to be difficult to argue otherwise.

Viewed from this externality perspective, Theorem 5 makes excellent sense. It asserts that, with sufficiently large positive β values, the social welfare function favors TL, which must be expected. The decomposition (Table 6b) requires $\beta > 0$ to favor TL payoffs. Conversely, $\beta < 0$ enhances the BR payoffs.

Table 9. Conflicting measures.

a. *Example with* $\beta = 0$

	$+1$		-1	
$+1$	4	4	2	-4
-1	-4	2	-2	-2

b. *Example with* $\beta = -6$

	$+1$		-1	
$+1$	-2	-2	8	-10
-1	-10	8	4	4

[7] If ties, such as $TL \sim BR$ or $BR \sim (BL \sim TR)$ are included, there are seven distinct social welfare rankings for each game in each Figure 2a region.

The potential and social welfare rankings can even reverse each other. According to Theorem 4-2, all games in region 2 of Figure 2a have a single Nash cell with the P ranking of $TL \succ (BR \sim TR) \succ BR$. This region requires $\alpha > \gamma > 0$, so the Table 9a example is constructed with $\alpha = 3, \gamma = 1$. For this game, where $\beta = 0$, the P and social welfare rankings agree. To modify the game to obtain the reversed social welfare ranking of $BR \succ (BL \sim TR) \succ TL$, where the non-Nash cell BR will be the social welfare function's best choice, Theorem 5 describes precisely what to do; select β values that satisfy $(\alpha + \beta) < -2\gamma$. For Table 9, this means that $\beta < -2(1) - 3 = -5$. The $\beta = -6$ choice leads to the Table 9b game, where the social welfare ranking reverses that of the potential function. Again, it is difficult to argue against this game's social welfare ranking.

4. Conflict and Agreement

These negative conclusions, where potential and social welfare rankings disagree, can be overly refined for many purposes. Similar to an election, the interest may be in the winner rather than who is in second, third, or fourth place. Thus, an effective but cruder measure is to determine where potential and social welfare functions have the same top-ranked cell.

All of the conflict in potential and social welfare rankings are strictly caused by β values, which suggests identifying those β values that allow the same potential and social welfare preferred cell. It is encouraging how answers follow from α and β comparisons.

Corollary 2. *For symmetric 2×2 games, the following hold for $\gamma \geq 0$:*

1. *The potential and social welfare functions have TL as the top-ranked and payoff dominant cell for $\alpha + \beta > 0$ and $\alpha > 0$ (shaded Figure 1a region on the right).*
2. *If $\alpha + \beta < 0$ and $\alpha > 0$ (unshaded region region on the right of Figure 1a), then BR is the social welfare top-ranked and payoff dominant cell, but the top-ranked P cell is TL.*
3. *If $\alpha + \beta < 0$ and $\alpha < 0$ (shaded Figure 1a region on the left), both functions have the BR cell top-ranked. BR also is the payoff dominant cell.*
4. *If $\alpha + \beta > 0$ and $\alpha < 0$ (unshaded region on the left of Figure 1a), TL is the social welfare top-ranked and payoff dominant cell, while the P top-ranked cell is BR.*

The content of this corollary serves as a legend for the Figure 1a map; the shaded regions are where both measures have the same top-ranked cell. A simple way to interpret this figure is that for all games to the right of the β axis ($\alpha > 0$), P's top-ranked cell is TL, while to the left it is BR. In contrast, above the $\alpha + \beta = 0$ slanted line, the social welfare's top-ranked cell is TL, while below it is BR. Thus, in the unshaded regions, one measure has BR top-ranked, while the other has TL.

This corollary and Figure 1a show that if the α value (indicating a preference of the agents for T and L or B and R) is not overly hindered by the externality forces (e.g., if $\alpha > 0$ and $\beta > -\alpha$) then the potential and social welfare functions share the same top ranked cell. But should conflict arise between these two fundamental variables, where the α and β values favor cells in opposite directions, disagreement arises between the choices of the top ranked P and social welfare cells.

Proof. The proof follows directly from Table 8. With $\gamma \geq 0$, P's top-ranked cell is TL for $\alpha > 0$ (to the right of the Figure 1a β axis), and BR for $\alpha < 0$ (Table 8a). According to Table 8b, the social welfare's top-ranked cell is TL iff $\gamma + [\alpha + \beta] > \gamma - [\alpha + \beta]$, or iff $\alpha + \beta > 0$; this is the region above the Figure 1a slanted line. The same computation shows that the social welfare's top-ranked cell is BR for the region below the slanted line. This completes the proof. □

Everything becomes slightly more complicated with $\gamma < 0$. The reason is that this $\gamma < 0$ anti-coordination factor permits BL and TR to become Nash cells. This characteristic is manifested in Figure 1b, where the Figure 1a $\alpha = 0$ and $\alpha + \beta = 0$ lines are separated into strips.

The content of the next corollary is captured by Figure 1b, where the potential and social welfare's top-ranked cells agree in the three shaded regions. To interpret Figure 1b, P's top-ranked cell is BR for all games to the left of the vertical strip ($\alpha < -|\gamma|$), cells BL and TR (or $BL \sim TR$) in the vertical strip ($-|\gamma| < \alpha < |\gamma|$), and cell TL to the right of the vertical strip ($\alpha > |\gamma|$). Similarly, the social welfare's top-ranked cell is BR below the slanted strip, $BL \sim TR$ in the slanted strip, and TL above the slanted strip. As $\gamma \to 0$, the width of the strips shrink and Figure 1b merges into Figure 1a.

Corollary 3. *For symmetric* 2×2 *potential games, the following hold for* $\gamma < 0$*:*

1. *The top-ranked P cell is TL iff* $\alpha > |\gamma|$ *(the Figure 1b region to the right of the* $\alpha = |\gamma|$ *vertical line). In this region, the social welfare ranking is TL (to agree with P) for* $(\alpha + \beta) > 2|\gamma|$ *(shaded region on the right of Figure 1b), but conflicts with P's choice with the social welfare top-ranking of* $BL \sim TR$ *if* $2|\gamma| > (\alpha + \beta) > -2|\gamma|$ *(the portion of the strip below the shaded region on the right of Figure 1b), and with the top-ranked BR for* $-2|\gamma| > (\alpha + \beta)$ *(the unshaded region below the strip on the right of Figure 1b).*
2. *For* $-|\gamma| < \alpha < |\gamma|$ *(the vertical strip of Figure 1b), both BL and TR are P's top-ranked cells with the ranking* $BL \sim TR$*. In this strip, the social welfare function has the same* $BL \sim TR$ *ranking only if* $-2|\gamma| < (\alpha + \beta) < 2|\gamma|$ *(the shaded trapezoid). Outside of this region in the strip, the social welfare top-ranked cell differs from P's* $BL \sim TR$ *choice by being TL for* $(\alpha + \beta) > 2|\gamma|$ *(above the trapezoid) and BR for* $(\alpha + \beta) < -2|\gamma|$ *(below the trapezoid).*
3. *The top-ranked cell for P is BR for* $\alpha < -|\gamma|$ *(to the left of the vertical strip). The social welfare function's top-ranked cell also is BR for* $\alpha + \beta < -2|\gamma|$ *(shaded Figure 1b region on the left). However, in this region, the social welfare function has BL and TR top ranked, or* $BL \sim TR$*, for* $-2|\gamma| < (\alpha + \beta) < 2|\gamma|$ *(the portion of the slanted strip above the shaded region) and TL top ranked for* $(\alpha + \beta) > 2|\gamma|$ *(above the slanted strip).*

Proof. The proof follows directly from Table 8. With $\gamma < 0$, it follows from Table 8a that P's top-ranked cell is TL if it is preferred to either BL or TR, which is if $2\alpha + \gamma > -\gamma$ or if $\alpha > |\gamma|$. This is the Figure 1b region to the right of the $\alpha = |\gamma|$ vertical line. Similarly, P's top-ranked cell is BR if $\gamma - 2\alpha > -\gamma$, or if $-\alpha > |\gamma|$; this is the region to the left of the $\alpha = -|\gamma|$ vertical line. The same computation shows that in the vertical strip $-|\gamma| < \alpha < |\gamma|$, P's top-ranked cells are the two Nash cells BL and TR, where P's ranking is $BL \sim TR$.

Using the same approach with Table 8b, it follows that the social welfare's top-ranked cell is TL if it has a higher score than BL or TR, which is if $-|\gamma| + (\alpha + \beta) > |\gamma|$, or if $\alpha + \beta > 2|\gamma|$. This is the region above the $\alpha + \beta = |\gamma|$ slanted line. Similarly, the social welfare top ranked cells are $BL \sim TR$ for $-2|\gamma| < (\alpha + \beta) < 2|\gamma|$, which is the slanted strip (which expanded the Figure 1a slanting line), and BR for $(\alpha + \beta) < -2|\gamma|$, which is the region below the slanting strip. This completes the proof. □

4.1. Changing β_1, β_2

The cause of conflict between potential and social welfare rankings now is clear; the first ignores β values while the second depends upon them. However, a feature of the previous section is that if TL or BR ended up being the social welfare top-ranked cell, it also was the payoff dominant cell. This property is a direct consequence of the symmetric game structure where the behavioral terms (Table 6b) always favored one of these two cells.

To recognize the many other possibilities, change the β structure from $\beta = \beta_1 = \beta_2$ to $\beta = \beta_1 = -\beta_2$. This affects Table 6b by changing the sign of player 2's entries, so the game's externality features now emphasize either BL or TR. The social welfare function becomes

$$w = 2\gamma t_1 t_2 + \alpha(t_1 + t_2) + \beta(t_2 - t_1) \tag{13}$$

A reason for considering this case is that any (β_1, β_2) can be uniquely expressed as

$$(\beta_1, \beta_2) = (b_1, b_1) + (b_2, -b_2) \text{ where } b_1 = \frac{\beta_1 + \beta_2}{2}, b_2 = \frac{\beta_1 - \beta_2}{2}. \tag{14}$$

Thus, combining Figure 1 with the impact of $(\beta, -\beta)$ captures the general complexity.

Because the decomposition isolates appropriate variables for each measure, Table 8 is the main tool to derive the Figure 3 results. In this new setting, Table 8 is replaced with Table 10, where part a restates the potential function values for each cell and b gives half of the social welfare function's values.

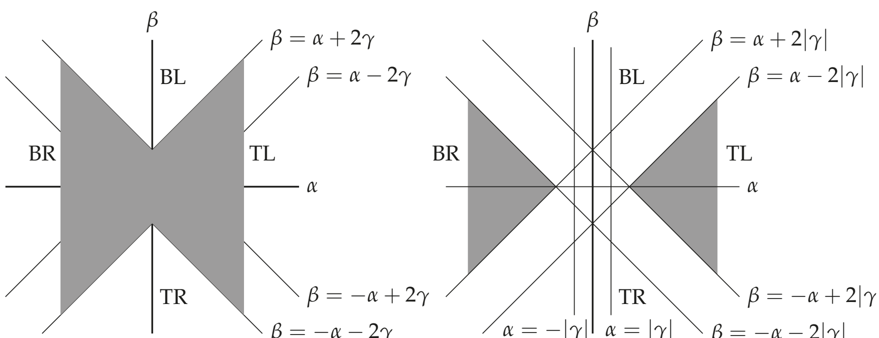
Table 10. A quasi-symmetric potential game's P and $\frac{w}{2}$ values, with $(\beta, -\beta)$.

	+1	−1
+1	$2\alpha + \gamma$	$-\gamma$
−1	$-\gamma$	$\gamma - 2\alpha$

a. *P values*

	+1	−1
+1	$\gamma + \alpha$	$-\gamma - \beta$
−1	$-\gamma + \beta$	$\gamma - \alpha$

b. $\frac{w}{2}$ *values*



(**a**). The α-β plane; $\gamma \geq 0$ (**b**). The α-β plane; $\gamma < 0$

Figure 3. More conflict and agreement with α, β values.

As with Figure 1a, if $\gamma \geq 0$, then P's top-ranked cell is TL for $\alpha > 0$ and BR for $\alpha < 0$. The same holds for Figure 3a. According to Table 10b, the social welfare function ranks TL over BL if $\gamma + \alpha > -\gamma + \beta$, or if $\beta < 2\gamma + \alpha$. In Fig. 3a, this is the region below the $\beta = \alpha + 2\gamma$ slanted line. Similarly, the social welfare function ranks TL over TR if $\alpha + 2\gamma > -\beta$, or $\beta > -\alpha - 2\gamma$, which is the Figure 3a region above the $\beta = -\alpha - 2\gamma$ line. P's top ranked cell is BR if $\alpha < 0$, which is the Figure 3a region to the left of the β axis. A similar analysis shows that the social welfare function ranks BR above TR if $\beta > \alpha - 2\gamma$, or the region above the Figure 3a $\beta = \alpha - 2\gamma$ line. Finally, this function ranks BR above BL if $\beta < -\alpha + 2\gamma$, which is the region below the $\beta = -\alpha + 2\gamma$ line.

Consequently, agreement between the two measure's top-ranked cell is in Figure 3a shaded regions, where BR is the common choice to the left of the β axis and TL is for the region to the left. Conflict occurs in the unshaded region where BL is the welfare's top-cell on the top and TR is for the region below. Again, these outcomes capture the β structure where, now, positive β values emphasize the BL cell and negative values enhance the TR entries. Contrary to Figure 1a, the upper unshaded region now is where BL, rather than TL, is the welfare function's top-ranked cell.

Consistent with Figure 1b, the situation becomes more complicated with the anti-coordination $\gamma < 0$. Again, P's top ranked cell is BR to the left of the vertical strip, $BL \sim TR$ in the strip, and TR to the right of the strip. Similar algebraic comparisons show that the social welfare's top-ranked cell is BL in the upper unshaded region of Figure 3b, including the portion of the β axis. Similarly, TR is the welfare function's top-ranked cell in the lower unshaded region. Consequently, the two large shaded regions are where agreement occurs (going from left to right, BR, TL).

With the Table 6a Nash structure, outcomes for all possible (β_1, β_2) values can be computed from Figures 1 and 3. To illustrate with $\gamma = -0.5, \alpha = -2, \beta_1 = 11, \beta_2 = 1$, it follows from $|\gamma| < -\alpha$ that BR is P's top-ranked cell. The information for the welfare function comes from Equation (14) where $b_1 = 6, b_2 = 5$. To find half of the welfare functions value, substitute $\alpha = 2\gamma = -0.5, \beta = 6$ in Table 8b, substitute $\alpha = 2\gamma = -0.5, \beta = 6$ in Equation (13b), and add the values. It already follows from plotting these values in Figures 1b and 3b that the outcome is either TL or BL.

4.2. Changing α_1, α_2

The general setting for a potential game involves variables $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma$. This suggests mimicking what was done with β by carrying out an analysis using $\alpha = \alpha_1 = -\alpha_2$; this is simple, but not necessary. The reason is that most needed information about P's top-ranked cell comes from Equation (5). As this expression shows, with appropriate choices of α_1, α_2, and γ, any cell can be selected to be P's risk-dominant, top-ranked choice, any admissible pair of cells can be Nash cells, where a designated one is risk dominant, and any cell can be selected to be the sole Nash cell. Finding how the behavioral terms (the β_1, β_2 values) can change which cell is the welfare function's top-ranked cell has been reduced to elementary algebra.

All that is needed to obtain answers is to have a generalized form of Tables 8 and 10, which is given in Table 11. The Table 11a values come from the general form of the potential function in Equation (2). The Table 11b values for the welfare function come from a direct computation of its equation

$$w(t_1, t_2) = (\alpha_1 + \beta_2)t_1 + (\alpha_2 + \beta_1)t_2 + 2\gamma t_1 t_2. \tag{15}$$

Table 11. A potential game's P and w values.

	+1	−1
+1	$\alpha_1 + \alpha_2 + \gamma$	$\alpha_1 - \alpha_2 - \gamma$
−1	$-\alpha_1 + \alpha_2 - \gamma$	$-\alpha_1 - \alpha_2 + \gamma$

a. *P values*

	+1	−1
+1	$\alpha_1 + \alpha_2 + \beta_1 + \beta_2 + 2\gamma$	$\alpha_1 - \alpha_2 - \beta_1 + \beta_2 - 2\gamma$
−1	$-\alpha_1 + \alpha_2 + \beta_1 - \beta_2 - 2\gamma$	$-\alpha_1 - \alpha_2 - \beta_1 - \beta_2 + 2\gamma$

b. *w values*

In the manner employed above, Theorem 6 is a sample of results. Here, use Equation (5) to determine the potential function structure, and Theorem 6 to compare social welfare (and β) values.

Theorem 6. *The social welfare function (Equation (15)) is maximized at TL if and only if $\alpha_1 + \beta_2 + \alpha_2 + \beta_1 > 0$ and $\beta_i + \alpha_{\neg i} > -2\gamma$, where $i = 1, 2$ and $\neg i$ denotes the agent who is not i. The welfare function is maximized at TR if and only if $\alpha_2 + \beta_1 < -2\gamma, \alpha_1 + \beta_2 > 2\gamma$, and $\alpha_1 + \beta_2 > \alpha_2 + \beta_1$. The welfare function is maximized at BL if and only if $\alpha_1 + \beta_2 < -2\gamma, \alpha_2 + \beta_1 > 2\gamma$, and $\alpha_2 + \beta_1 > \alpha_1 + \beta_2$. Finally, the welfare function is maximized at BR if and only if $\alpha_1 + \beta_2 + \alpha_2 + \beta_1 < 0$ and $\beta_i + \alpha_{\neg i} < 2\gamma$, for $i = 1, 2$.*

5. Discussion

Questions regarding the various measures for game theory have proved to be difficult to analyze. There is an excellent reason for this complexity; answers must depend upon the particular payoffs of a game, but it was not clear what portions of each payoff contribute to which aspects of a game. As such, a surprising and welcomed property of the coordinate system is how it identifies how all of a game's entries interact; the coordinates precisely dissect and extract from each payoff entry its contribution to the different attributes of a game.

Support for these comments come from equations such as Equation (9) for the potential function and Equation (10) for the welfare function. The different signs of t_1t_2 and t_i coefficients, for instance, nicely capture the complexity of a standard approach; it indicates there exists a twisting of certain portions of the payoff entries that are needed to carry out an analysis. The decomposition's separation of which parts of a payoff entry affect Nash structures and which affect payoff and externality factors explain why different measures of a game can have different conclusions. What illustrates the power of doing so is how the discovery and proofs of many subtle results now reduce to elementary algebraic computations.

Our analysis described how and why differences can arise among potential function, payoff dominance, and social welfare conclusions about games. Everything extends more generally. As the decomposition demonstrates, expect methods, learning approaches, and measures that emphasize "best response", comparisons of individual payoff differences, and obtaining Nash equilibria to ignore behavioral terms. Should the objective be to identify properties of Nash structures, doing so simplifies the analysis by eliminating the redundant (for a Nash analysis) β_j variables. However, by not including β terms, it must be expected that answers from these approaches about games will differ from those measures that capture the value of payoffs, such as the social welfare function and payoff dominance. They must; the two different classes of measures depend upon different information about the games.

Author Contributions: Conceptualization, all authors; investigation, all authors; methodology, all authors; supervision, all authors; validation, all authors; writing-original draft, all authors; writing-review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Our thanks for useful comments from two referees and an editor.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schelling, T. C. *The Strategy of Conflict*; Oxford University Press: New York, NY, USA, 1960.
2. Lewis, D. K. *Convention: A Philosophical Study*; Harvard University Press: Cambridge, MA, USA, 1960.
3. Rosenthal, R. W. A class of games possessing pure-strategy Nash equilibria. *Int. J. Game Theory* **1973**, *2*, 65–67. [CrossRef]
4. Monderer, D.; Shapley, L. S. Potential games. *Games Econ. Behav.* **1996**, *14*, 124–143. [CrossRef]
5. Young, P. The Diffusion of Innovations in Social Networks; In *The Economy as an Evolving Complex System*; Oxford University Press: New York, NY, USA, 2006; Volume III, pp. 267–281.
6. Young, P. The dynamics of social innovation. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 21285–21291 [CrossRef] [PubMed]
7. Newton, J.; Sercombe, D. Agency, potential and contagion. *Games Econ. Behav.* **2020**, *119*, 79–97. [CrossRef]
8. Harsanyi, J.C.; Selten, R. *A General Theory of Equilibrium Selection in Games*; MIT Press: Cambridge, MA, USA, 1988.
9. Jessie, D.T.; Saari, D.G. Cooperation in n-player repeated games. In *The Mathematics of Decisions, Elections, and Games;*American Mathematical Society: Providence, RI, USA, 2014; pp. 189–206.
10. Jessie, D.T.; Saari, D.G. *Strategic and Behavioral Decomposition of Games*; IMBS Technical Report 15-05; University of California: Irvine, CA, USA, 2015.
11. Jessie, D.T.; Saari, D.G. From the Luce Choice Axiom to the Quantal Response Equilibrium. *J. Math. Psychol.* **2016**, *75*, 1–9. [CrossRef]
12. Jessie, D.T.; Saari, D. G. *Coordinate Systems for Games: Simplifying the "Me" vs. "We" Interactions*; Springer: Berlin/Heidelberg, Germany, 2019.
13. Kalai, A.; Kalai, E. Cooperation in strategic games revisited. *Q. J. Econ.* **2013**, *128*, 917–966. [CrossRef]
14. Candogan, O.; Menache, I.; Ozdaglar, A.; Parrilo, P. Flows and Decompositions of Games: Harmonic and Potential Games. *Math. Oper. Res.* **2011**, *36*, 474–503. [CrossRef]
15. Abdou, J.; Pnevmatikos, N.; Scarsini, M.; Venel, X. Decomposition of Games: Some Strategic Considerations. *arXiv* **2019**, arXiv:1901.06048.

16. Hwang, S-H.; Rey-Bellet, L. Strategic Decompositions of Normal Form Games: Zero-sum Games and Potential Games. *Games Econ. Behav.* **2020**, *122*, 370–390. [CrossRef]
17. Jessie, D. T.; Kendall, R. *Decomposing Models of Bounded Rationality*; Technical Report, IMBS Technical Reports, 15-06; MBS: Irvine, CA, USA, 2015.
18. Blume, L.E. The Statistical Mechanics of Strategic Interaction. *Games Econ. Behav.* **1993**, *5*, 387–424. [CrossRef]
19. Hofbauer, J.; Sandholm, W.H. Stable Games and their Dynamics. *J. Econ. Theory* **2009**, *144*, 1665–1693. [CrossRef]

games

MDPI

Article

Valuable Cheap Talk and Equilibrium Selection

Julian Jamison

Department of Economics, University of Exeter Business School, Rennes Drive, Exeter EX4 4PU, UK; J.Jamison@exeter.ac.uk

Academic Editor: Ulrich Berger
Received: 8 June 2020; Accepted: 5 August 2020; Published: 20 August 2020

Abstract: Intuitively, we expect that players who are allowed to engage in costless communication before playing a game would be foolish to agree on an inefficient outcome amongst the set of equilibria. At the same time, however, such preplay communication has been suggested as a rationale for expecting Nash equilibria in general. This paper presents a plausible formal model of cheap talk that distinguishes and resolves these possibilities. Players are assumed to have an unlimited opportunity to send messages before playing an arbitrary game. Using an extension of fictitious play beliefs, minimal assumptions are made concerning which messages about future actions are credible and hence contribute to final beliefs. In this environment, it is shown that meaningful communication among players leads to a Nash equilibrium (NE) of the action game. Within the set of NE, efficiency then turns out to be a consequence of imposing optimality on the cheap talk portion of the extended game. This finding contrasts with previous "babbling" results.

Keywords: strategic communication; two-stage games; pareto efficient equilibria; belief formation

---

1. Introduction

Self-enforcing agreements—those for that no party has any incentive to break given that all others comply—should be carried out even if they are not binding in a formal sense. This is in fact the defining characteristic of the standard Nash equilibrium concept, and thus, one of the common justifications for this concept is that if players are allowed to communicate before playing a game, they could hardly reasonably agree on an outcome not satisfying this criterion. Recall that a Nash equilibrium constitutes for each player a set of strategies and beliefs (about other players' strategies), such that the strategies are the best responses to beliefs and the beliefs are correct (see e.g., Osborne 2004 [1]). We assume that there is no recourse to court-enforceable contracting, or equivalently that any such interactions have already taken place. Unfortunately, while intuitively pleasing, this justification for the use of a Nash equilibrium has been characterized by a shortage of formal models.

On a related, but distinctly different track of reasoning, it is natural to wonder why agents would ever agree on an inefficient outcome, assuming that they had the chance to talk in the first place. In other words, why would players agree ahead of time to an inefficient outcome of a game if there were another potential outcome, also an equilibrium, that gave strictly greater payoffs to all of them? Once again, the challenge has lied in constructing a realistic, but necessarily simplified, formal model of the agents' communication process. Among other problems, this inefficient result appears to be incompatible with the arguments outlined above, in which Nash equilibria in general are justified.

This type of preplay communication is often called *cheap talk* , which may be roughly defined as nonbinding, nonpayoff relevant, preplay communication. Although cheap talk has indeed received attention as a potential solution to these questions surrounding the equilibrium concept, in practice, it has been mostly used in the study of signaling games, in repeated environments (often in connection with learning), and in certain applied settings. These are of course all important applications, but these leave the original ambiguities unresolved. This paper, then, returns to the goal of constructing a more comprehensive model of pure cheap talk and explores its relationship with equilibria and equilibrium selection.

This paper develops a model of cheap talk that involves an unlimited communication session, called a *conversation* , before the play of a standard game begins. Players announce in advance what actions they plan to take in the upcoming game, and taken together, these announcements form one possible prediction of what they may actually do. On the other hand, there is also a common prior forecast, given exogenously, of what each player will do; this forecast is updated as the conversation proceeds. An announcement is defined to be credible only if it is close to the best response to one or the other of these two predictions about the rest of the players. Otherwise, an announcement has no external justification, so it is deemed unbelievable and disregarded. The conversation continues indefinitely in this manner, possibly, but not inexorably toward some limit. Realistically, it will rarely if ever go on for very long (although, for complicated games, it may take some time), since if it is going to converge, it will do so rapidly. However, it is important not to have an artificial limit imposed externally, just as long finitely repeated games behave very differently than infinitely repeated games.

The paper's first main result is that if the conversation converges toward a limit, then this limit must be a Nash equilibrium of the underlying action game in which payoffs are determined.[1] Conversely, any Nash Equilibrium forms a possible limit of the conversation. This result can be interpreted as saying that meaningful communication before a game can only lead to Nash outcomes. Since the cheap talk is the initial interaction between the players, we assume that they cannot be sure of the strategies that their opponents will follow in the communication stage. Any strategy in this phase that is weakly dominated by another is clearly not optimal; anything else is potentially the preferred choice and is therefore, given the lack of information, one possible optimal choice.[2] The paper's second result then states that optimal pregame play in the conversation stage leads to an efficient outcome and that any efficient final outcome is a possible result of such a strategic conversation.[3] This can be interpreted as saying that rational, or thoughtful, speech leads to efficiency. This completes the connection between cheap talk (as modeled here, i.e., in an environment where rationality and utilities are common knowledge), Nash equilibria, and Pareto optimality. The first result applies to all games (at least those with a Nash equilibrium), while the second result only has bite in games with multiple equilibria.

The conclusion derived from the second main result contrasts with previous "babbling" results, in which it is impossible to select among the set of Nash equilibria because all pregame communication is ignored. The main reason for the difference is that those previous studies looked for equilibria of the extended communication game as a whole—for instance, by assuming that the full strategies of all players are known. This allows equilibrium strategies in which no value is placed even on seemingly mutually informative communication, whereas the model below presupposes the impossibility of ignoring beneficial interchange. Thus, the present paper takes a more primitive view of pregame strategies, especially since in part it is attempting to justify the equilibrium concept in the first place. Naturally, although the model does not impose beliefs about the cheap talk stage, it still must make some assumption about beliefs held upon entering the action game. Another approach that will destroy the babbling equilibria is to assume an arbitrarily small, but positive cost to sending messages—this is a restriction on the environment rather than on the structure of equilibrium or on belief formation. While this limitation is plausible in reality, it is, strictly speaking, no longer a model of cheap talk, even if the total sum spent on sending messages is always lower than the game's smallest payoff differential.

The paper proceeds as follows. Section 2 provides a brief survey of some of the relevant literature. In Section 3, some motivation is given for the specific assumptions made in this conversational model

1 The limit is an ε-Nash equilibrium.
2 This is discussed in further detail in Section 3.
3 The notion of efficiency used here is *stable efficiency*, a concept that is equivalent to Pareto efficiency in generic two person games.

of cheap talk. Section 4 lays out the formal model, stating and proving the paper's two main results. Several examples are detailed in Section 5 in order to illustrate both the cheap talk process and the implications of the theorems. Finally, Section 6 concludes the paper by summarizing the model and discussing some possible extensions of its implications.

2. Previous Literature

The concept of cheap talk was introduced into the economics literature by Crawford and Sobel (1982) and Farrell (1987) [2,3]. Since then, a sizable literature has developed related to this topic, with such examples as Farrell and Gibbons (1989), Forges (1990), Farrell (1993), Aumann and Hart (1993), Blume and Sobel (1995), and a survey in Farrell and Rabin (1996) [4–9]. The paper that perhaps is closest to the present one is Rabin (1994) [10]. It models a finite instead of an infinite opportunity for communication, but also seeks a notion of optimality rather than equilibrium in the analysis of the extended game. The specific form of cheap talk assumed by Rabin is different from the one presented below, in particular with respect to the element of choice between strategies against which to credibly best respond. The results can be framed in terms of the two central questions posed here, but are generally less conclusive in either. Both papers adhere to the full rationality paradigm of classical game theory and previous work on cheap talk, as opposed to, say, the evolution literature.

There are a number of papers that study a more limited class of games. For instance, Matsui (1991) [11] applied cheap talk to common interest games, and in this context, his notion of *cyclic stability* yields efficiency. Canning (1997) [12] studied signaling games of common interest, although the messages do not necessarily constitute cheap talk per se. He found that off-path beliefs are vital to the question of whether or not efficiency is eventually realized; randomly drawn off-path beliefs encourage experimentation and lead to efficiency. Finally, Sandroni (2000) [13] studied two person repeated coordination games without cheap talk. He introduced the concept of *blurry beliefs*, which is a less restrictive (that is, more fully rational) belief dynamic than those used in evolutionary game theory, although it is stronger than anything used here. Sandroni showed that if the belief classes of the players satisfy *reciprocity*, then cooperation will be achieved. Overall, the current paper pins down the link between communication and (efficient) equilibrium outcomes more concretely than the previous literature. In particular, it explores a specific empirically-consistent model of belief formation and shows a two way equivalence between that process and the optimality of the resulting behaviors.

A fairly large class of papers has studied repeated games and the emergence of Nash equilibria without introducing cheap talk, including Crawford and Haller (1990), Young (1993), and Kalai and Lehrer (1993) [14–16]. Finally, there have been some experimental studies of communication and equilibrium selection in various coordination games; see, for example,Cooper et al (1992), Brandts and Cooper 125 (2007), and Cachon and Camerer (1996) [17–19]. The results can be summarized (and oversimplified) as finding that two way pregame communication greatly increases the chances of observing efficient equilibrium outcomes. Pertinently, this holds even if the efficient equilibrium is not risk-dominant, in contradistinction to some previous results. Meanwhile, some experimental studies found that preplay communication can actually induce fewer choices consistent with Nash equilibria, e.g., Boulu-Reshef et al (2020) [20] in the context of public goods games. This could either be due to the limited opportunity for communication and/or the possibility of social preferences (which would change the set of NE).

3. Motivation

This section provides some intuition and justification for the structure of the model that follows; the impatient reader can skip to the next section. The model assumes that there is an action game to be played, about which the players are assumed to have full information (in order to abstract away from any signaling incentives during the conversation). Each player begins with a common forecast about what actions he or she will take in the upcoming game. These expectations can be interpreted as vague initial ideas about how the game might be played, arising perhaps from societal conventions or from focal points (hence the assumption that the forecasts are common and known). They are not beliefs in the formal sense, although they will be updated throughout the conversation.[4] Since a priori, nothing can be absolutely ruled out by any of the players, the prior forecasts are totally mixed.[5] Needless to say, the forecasts are not in any way binding: players ignore what they themselves are "expected" to do, although they can take into account the influence this expectation has on their opponents.[6] The key distinction between forecasts and standard beliefs is that the forecasts are about the general environment (how might this game typically be played by others?), whereas beliefs are about the actual decisions by the specific players interacting in a given concrete situation. Thus, among other implications, it makes sense to reason about players trying to influence the beliefs that their opponents have about them, whereas they cannot influence the more broadly prevalent forecasts. Of course, then, we need to model from whence the forecasts come (social norms, news media, evolutionary psychology, etc), but that is outside the scope of the present paper.

During the conversation stage, before playing the action game, players send public messages to each other. Since we are attempting to understand what such preplay communication can achieve, we assume that there is an unlimited (but countable) opportunity to send these messages. For simplicity and without loss of generality, the messages are taken simply to be announcements of a player's own expected actions in the game. One could assume instead that players announce mixtures of their possible actions, but this is an unnecessary complication. Essentially, given infinite riskless communication, this slight limitation on the flexibility of messages imposes no loss in the long run. Implicitly, we are assuming that players can understand one another and that they take messages at face value (not in a strategic sense, but in a linguistic sense). If the message "action L" is sent, everyone understands that to mean "action L" and not "action R". Thus, there is a *natural language* for speech; the players share enough common history or cultural affiliation that they are able to talk and understand one another in a previously unencountered situation.

Of course, not all announcements should be considered seriously. We need to define a notion of *credibility* or believability. The first requirement is that a player's announced action should be *self-committing*, in the sense that if it were believed and best responded to, the original announcer would still be willing to carry through with it (within the confines of the action game). This requirement is equivalent, then, to being in the support of some Nash equilibrium of the action game. At the beginning of the preplay conversation, any self-committing action is credible, so players have a chance to guide the discussion. In general, there will be some tradeoff between allowing the players leeway to influence the conversation at the beginning, but requiring them at some point to pay attention to what the others are saying and to reflect that updated information in their own announcements. Unlike in the deterministic best response dynamics of evolutionary models, it is important in this model that players have a choice over what to say; this is the hallmark of a conversation. It is this choice, along with the lack of payoffs until the action game is at the very end, that differentiates this paper's model from an evolutionary learning model.

The common forecast is very slowly updated by each credible announcement. We can think of the prior forecast as the result of a long, but finite fictitious history of credible announcements, with each new stated action adding to the average.[7] As beliefs get updated, the initial forecast can be ignored and only the actual credible announcements counted toward an average forecast: this forecast constitutes a

---

4 The players do not have beliefs about the full strategies of their opponents, only ideas about what might actually occur in the game. Thus, the preplay forecasts are distributions over actions, not distributions over mixed strategies (which themselves are distributions over actions). This is not crucial to the conclusions reached.

5 It is not strictly necessary for the results that the priors be totally mixed.

6 The author performed the analysis under the seemingly weaker assumption that all that is known about the prior forecasts is that they place a certain minimum weight on each action, but the results carry over. Since this assumption adds complexity, but is no sounder in justification (Why can the entire distributions not be known if the minimum weights are?), it has been left out.

7 Recall that the average of multiple sets of actions is equivalent to a mixed strategy.

player's *appearance*. In general, we recursively define an announced action to be credible if it is the best response (within ε) to either the current forecast of an opponent's behavior or to an opponent's appearance.[8] If there are more than two players, either the common updated forecast or a player's appearance may be substituted for each. The intuition here is that a player can either say something like, "This is what I think you are going to do, and if so then I would plan to do such-and-such," or something along the lines of, "Okay—for the moment I'll take you at your face value, and in that case I'll want to do so-and-so." Of course, he or she only needs to consider credible announcements in making these plans.

At any time during the preplay conversation, a player can make any announcement desired, but only those statements that are credible will have an impact on the conversation. Since all players know the prior forecasts and all previous announcements, they can calculate which of these announcements were actually credible and hence also which of their own announcements will be perceived as credible by others. If at any point, there is but a single action that is credible for a particular player, it must be that this player can only seriously be considering that action (at that point in time). Therefore, in effect, it does not matter whether or not he or she actually announces that action; everyone knows that it is being considered, and hence, it should count toward the forecast and appearance of that player, regardless of what may or may not be announced. This argument implies that without loss of generality, we may assume that all players make credible announcements during each round of the conversation.[9] Finally, we assume that at each point in time, any player can start over; that is, declare a clean slate and remake their appearance anew. This is the equivalent of declaring that the conversation has broken down from his or her perspective and, among other implications, allows the players to attempt to coordinate. Although it may seem like an overly strong possibility, in fact, a player's appearance is a powerful commitment device, and so, giving up on it involves a significant loss.[10] In any case, of course, the clean slate option is available equally to all of the players. This completes the description of the cheap talk conversation.

One last remaining question about the credibility concept concerns the infinite durability of credible announcements. That is, a credible announcement always "counts" even if it is no longer credible. The reason for this is that any credible announcement indicates evidence of a desire for that action if possible, and there is no reason to think that the desire will change or that the desired action may not once again become plausible. In effect, each announcement has a small impact that builds toward the whole impression, rather than the fads of currently credible actions. In fact, if only those actions that are credible at the moment are averaged into the player's appearance, at each communication stage, one can observe swings back and forth of what is and is not believable. Furthermore, in this updated setting, eventually, only one pure strategy will be credible, and so, it is essentially impossible for players to converge to a mixed strategy.

Once the preplay conversation is complete, we have a countably infinite sequence of announcements for each player, with an associated sequence of appearances (the average credible announcement to date). This latter sequence may or may not have a limit.[11] Because of the infinite horizon and the nature of the updating process, if the limit does exist for a given player, then the forecasts made by the other players about this player will also converge, and to the same point. In this case, we specify that entering the concrete action game, the beliefs held by the other players about this player are also this same point in the strategy space. In this way, the conversation is a model of belief formation. If the appearance does not converge, then the appropriate forecast will not converge either,

8 We assume that players only care about payoffs up to some arbitrarily small constant ε, either because they cannot perceive finer differences or because they are indifferent over this range.
9 We make the standard assumptions on the action game so that a best response always exists.
10 In particular, continually starting over inhibits convergence, in which case, the player has no influence on the ultimate course of the discussion. This is never optimal, as shown below.
11 If no credible announcements were made after some finite stage, this is taken to mean that the limit does not exist. However, as above, we may assume that this does not occur.

and beliefs remain open for the time being. Of course, it may be true in general that appearances have a limit only for some (possibly empty) subset of the group of players.

If the appearances of all players converge, then we say that the conversation itself converges. However, in this case, every player continues to make credible announcements, and hence, at the limit, these announcements must be near the best responses to the actions stated by the other players, and hence to the limits of the other players. Since by definition, the latter are the beliefs held by the given player upon entering the action game, his or her limit must be an action that is (near) the best response to his or her beliefs and is therefore one optimal strategy to pursue in the action game. Therefore, we may assume that this limit action is indeed chosen, validating the beliefs of the other players. Of course, since this is true for all players, the limits must be mutual best responses, and thus, the play arrives at a Nash equilibrium. This is Theorem 1 below.

We next turn our attention to the question of optimality in the cheap talk stage of the overall game. Stepping back for a moment, we consider the question of whether or not to participate in the conversation at all, given the opportunity to do so. Since there is a natural language with which to communicate, any player can initiate a conversation. Whether or not they choose to participate, other players will hear and be influenced by the announcements of this player. Therefore, if they do not also make announcements, this player (or players) will have free reign to drive beliefs toward the equilibrium of their choosing (by announcing it ad infinitum). Since this outcome is at least weakly bad for other players, it cannot hurt them to also join in the conversation and attempt to guide the discussion in a direction favorable for them. For instance, in the Battle of the Sexes game, played between one man and one woman, Player 1 conversing with himself will continuously announce the equilibrium that he prefers. Entering the action game, the other player believes these announcements and best responds to them, so that the play will in fact be at that equilibrium. In this case, it would have been a good idea for Player 2 to at least try to promote her favored outcome, that is to participate in the conversation. Thus, we may assume, without any loss of generality, that all players converse.

Players do not know the cheap talk strategies employed by their opponents (if they did, we should instead be modeling what occurred before this conversation in order for that knowledge to be gained), so these players must consider all strategies to be possible. Thus, if a cheap talk strategy for one player never performs better (in terms of the payoffs ultimately realized in the action game, of course) than another competing strategy and does strictly worse against at least one possible strategy profile of the opponents, then the original strategy should be discarded as suboptimal. Anything that is not weakly dominated is optimal.[12] This is intentionally a broad definition of a strategy; it is meant to be as loose as possible and yet at least minimally capture the requirements of optimality. Theorem 2 below proves that if all players employ communication strategies that are optimal in this loose sense, then the conversation must converge to a stably efficient equilibrium of the game. This class of equilibria, defined below, is essentially those Nash equilibria for which no coalition can break away and, on their own, force the other players to follow them to some other equilibrium that is preferred by the coalition. In two person games with distinct payoffs (a property that holds generically), this result is equivalent to Pareto optimality.

4. Model

Consider a game **G** with n players and finite action spaces S_i for $i = 1, ..., n$.[13] Payoffs are given by u_i for $i = 1, ..., n$. It will be simplest to think of **G** in normal form. **G** is played exactly once, though **G** itself may be a repeated game. Before this happens, there is a **conversation** C(**G**), defined as follows. Each player begins the pregame conversation with a totally mixed prior **forecast** $\pi_i = \pi_i^1 \in \Delta(S_i)$ about his or her behavior. The forecasts are common knowledge among all the players. At each

[12] Naturally, since full rationality is assumed, we could endlessly iterate the process, but there is no need.
[13] The assumption of finiteness can be weakened.

round $t = 1, 2, 3...$ of the conversation, player i announces $m_i^t \in S_i$. The announcements are made simultaneously by all players in each round.[14]

Let $NE(G) \subseteq \bigtimes_{i=1}^{n} \Delta(S_i)$ be the set of Nash equilibria of **G**, and define $E_i \subseteq S_i$ by:

$$E_i = \{s_i \in S_i \mid \exists \sigma \in NE(G) \text{ with } s_i \in supp(\sigma_i)\}.$$

This set constitutes the self-committing actions for player i. At $t = 1$, any $m_i^1 \in E_i$ is said to be **credible**. If m_i^1 was credible, then we define:

$$\pi_i^2 = (T\pi_i^1 + m_i^1)/(T+1)$$

for some fixed T, which can be chosen to be large relative to the scale of the strategy space and payoffs in the underlying game. This captures the slow updating process of prior forecasts by credible announcements. In a similar fashion, the **appearance** is given by $p_i^2 = m_i^1$. If the initial announcement was not credible, then the forecast is not updated, and the appearance is undefined. Recursively, we now define m_i^t to be credible when:

$$m_i^t \in \varepsilon BR_i(\bigtimes_{j \neq i} q_j^t) \text{ with } q_j^t = \pi_j^t \text{ or } p_j^t \forall j,$$

where $\varepsilon BR_i(\sigma_{-1})$ denotes:

$$\left\{ s_i \in S_i \mid \max_{s_i' \in S_i} u_i(s_i', \sigma_{-1}) - u_i(s_i, \sigma_{-i}) < \varepsilon \right\}$$

for some arbitrarily small $\varepsilon > 0$. If m_i^t is not credible,[15] then $\pi_i^{t+1} = \pi_i^t$ and $p_i^{t+1} = p_i^t$. If m_i^t is credible, then we define:

$$\pi_i^{t+1} = ((T+t-1)\pi_i^1 + m_i^1)/(T+t) \text{ and } p_i^{t+1} = ((t-1)p_i^t + m_i^t)/t.$$

Say that player i's appearance *converges* if player i never entirely stops making credible announcements and if $\lim_{t\to\infty} p_i^t$ exists. If this happens, it is clear that $\lim_{t\to\infty} \pi_i^t$ also exists and is the same; call it b_i for the belief about player i. If the limit exists for all players, then the conversation converges. In this case, we assume that beliefs after the conversation and entering **G** are given by $\mu_i = \bigtimes_{j \neq i} b_j$.

Definition 1. *An **acceptable equilibrium** (of **G**) is a profile $\sigma \in \bigtimes_{i=1}^{n} \Delta(S_i)$ such that $\sigma = b$ for some belief vector b resulting from a convergent conversation starting at some prior forecasts π; the set of acceptable equilibria is denoted $AccE(G)$.*

Theorem 1.

1. $NE(G) \subseteq AccE(G)$
2. $AccE(G) \subseteq \varepsilon NE(G)$

Proof. (1) Let $\sigma \in NE(G)$, and consider prior forecasts π very close to σ. By the definition of a Nash equilibrium, any $s_i \in supp(\sigma_i)$ is in $\varepsilon BR_i(\pi_{-i})$. Now, let the players announce actions in the support

---

14 Sequential announcements lead to a forced asymmetry regarding who speaks when. The effects of this generalized first-mover advantage are irrelevant for the present discussion.

15 Unless player i has only one possible credible announcement, as discussed in Section 3.

of σ in such a way as to match as nearly as possible the actual distribution prescribed by σ. Initially, all these actions will be credible as stated. Of course, the forecasts will change over time, but since the updating process is slow and the cheap talk announcements are matching the given distribution, the forecasts will always stay near σ. Hence, the actions in the support of the announcements will remain credible forever. In this manner, $\lim_{t\to\infty} p_i^t$ exists $\forall i$, and moreover, $\lim_{t\to\infty} p_i^t = \sigma_i$. Thus, σ is indeed an acceptable equilibrium.

(2) If $\sigma \in AccE(G)$ and so is the limit of a convergent conversation, it must be that all $s_i \in \text{supp}(\sigma_i)$ are credibly announced infinitely often during the preplay cheap talk stage.[16] Since in the limit, both the forecasts and the appearances are arbitrarily near σ, each such s_i must be in $\varepsilon BR_i(\sigma_{-1})$, and therefore, $\sigma_i \in \varepsilon BR_i(\sigma_{-i}) \forall i$. □

Among other things, this result justifies the possibility that after a convergent conversation, players both rationally and self-consistently hold the beliefs that are given by the model. Theorem 1 in some sense clarifies the relationship between cheap talk (as has been modeled here) and Nash equilibrium. If the communication is meaningful, that is if the cheap talk has a limit, then it must lead to a Nash outcome. Of course, there is no guarantee that the conversation will converge, and it is quite possible that it will not.[17] Furthermore, no Nash equilibrium, even if inefficient, can yet be ruled out. Something stronger than an acceptable equilibrium is required.

We next turn to defining the appropriate efficiency concept in this setting.

Definition 2. *Call $\sigma \in NE(G)$* ***directly attainable*** *from $\sigma' \in NE(G)$ by the coalition S if σ_s is a Nash equilibrium in the induced game fixing all players outside of S to play as in σ', and if also $\forall i \notin S$, we have $u_i(\sigma_i, \sigma_S, \sigma'_{-i,S}) > u_i(\sigma'_i, \sigma_S, \sigma'_{-i,S})$.*

This is a strenuous definition: the first condition asks that the members of S be able to "jump" to σ from σ', and the second condition requires that once they have done so, they can force the rest of the players to follow them.

Definition 3. *Call $\sigma \in NE(G)$* ***attainable*** *from $\sigma' \in NE(G)$ by the coalition S if there is a chain of equilibria, each directly attainable by S, leading from σ' to σ; if also, $\forall i \in S \;\; u_i(\sigma) > u_i(\sigma')$; and if finally, there is no similar such chain (for any coalition) leading away from σ.*

These are once again fairly strict requirements. The second one states that all members of S must strictly prefer the new equilibrium, and the third states that the new equilibrium itself is immune to these sorts of deviations.

Definition 4. *A Nash equilibrium of* ***G*** *is* ***stably efficient*** *if nothing is attainable from it; the set of these equilibria is denoted $StEff(G)$.*

By considering the grand coalition of all players, it is clear that an equilibrium exhibiting stable efficiency will tend to be efficient. In games with distinct payoffs, no singleton coalitions can ever attain alternate equilibria (this follows from the first condition of the first definition), and hence, in two-person games, stable efficiency is generically equivalent to efficiency. It is clear that stably efficient equilibria always exist (since whatever is attained must itself be stably efficient). In most games, efficiency and stable efficiency will coincide, but when they do not, it is important that we use the latter concept. Stable efficiency is related to the coalition-proof concept introduced by [21], but is more

16 In particular, since the conversation converges, there must be some round after which nobody ever cleans their slate and starts over.
17 Consider, as one example, fictitious play in the rock-paper-scissors game.

farsighted in that it looks at the full implications of a coalitional deviation; it turns out that neither definition is a refinement of the other.

Recall that a cheap talk strategy is **optimal** if it is not weakly dominated.

Definition 5. *An* ***agreeable equilibrium*** *(of* ***G****) is a profile* $\sigma \in \mathop{\times}_{i=1}^{n} \Delta(S_i)$ *such that* $\sigma = b$ *for some belief vector b resulting from a convergent optimal conversation starting at some prior forecasts* π*; the set of agreeable equilibria is denoted* $AgrE(G)$*.*

Theorem 2.

1. $StEff(G) \subseteq AgrE(G)$
2. $AgrE(G) \subseteq \varepsilon StEff(G)$

Proof. (1) Consider $\sigma \in StEff(G)$, and let the prior forecasts π be very close to σ. Since the forecasts favor σ so heavily, the only way that another equilibrium can ever be reached during the conversation is if it is directly attainable or the result of a chain of directly attainable equilibria. Thus, all of the players know that these are the only feasible outcomes, and in fact (see the strategies below), they can be reached in a conversation. However, since σ is stably efficient, it is not possible for any player (as a member of any coalition) to be sure that by deviating from one of these alternates, a superior payoff can be achieved. It must be the case that either not all members of the coalition will profit by the switch (in which case, those who do not profit will not participate in the deviation) or if they do, that then, there is another coalition that can profitably and successfully deviate away from this new point. Of course, it is possible that one's payoff will be increased by attempting to switch equilibria, but there will always be circumstances in which it is not profitable. Thus, there is no strategy that weakly dominates the strategy "emulate σ", which is always available due to the prior forecasts. This implies that one optimal strategy for all players is to follow σ, and the result of this will be that the conversation converges with σ. There may be other optimal strategies, and there may be other possible results to the conversation; however, this is sufficient to show that $\sigma \in AgrE(G)$, as desired.

(2) Suppose that a conversation is converging toward an equilibrium σ that is not stably efficient (even up to ε-indifference). If there is just one coalition that can attain a superior equilibrium for itself, it can pursue the following strategy: (a) Erase its current appearance and start over, and then, (b) announce the actions that lead to the first equilibrium along the chain. If all members of the coalition have done likewise, then they will be able to credibly repeat those announcements in the next round, since these are mutual best responses given the forecasts near σ for the other players. If the other members have not done this, each individual can start over again and try once more. If eventually they coordinate, then they can continue to make these announcements indefinitely. At some point, the forecasts and appearances will then be very close to this new equilibrium, and the only credible choice for the other players will be to switch to it as well (this follows from the definition of directly attainable). They can continue in this fashion until the final equilibrium in the chain, where the process will conclude (by the argument in Part (1) above).

Of course, this attempted deviation will not always work, but it is safe in that either it works (that is, all members of the coalition coordinate) and a higher payoff is realized or it does not and the conversation stays at σ instead. Therefore, the deviation strategy weakly dominates the "emulate σ and stay where you are" strategy. Since this is true for all members of the coalition, optimality implies that all of them will attempt to force the switch to the preferred attainable equilibrium, and with probability one they will eventually coordinate (since they always have the opportunity to start over). Therefore, σ was not in fact an agreeable equilibrium.

Similarly, if there were several coalitions that could attain superior equilibria, each member of each coalition can start over at each round and attempt to coordinate with his or her coalition. Any player who is a member of several coalitions or who has a choice between attainable equilibria can randomize between these possibilities. If the player puts almost all weight on his or her individually preferred outcome among all these choices and spreads $\sigma(\varepsilon)$ weight across the others, then this will be ε-optimal, but will at the same time guarantee that with probability one, coordination takes place at some point. This weakly dominates "emulate σ" because either the conversation converges to σ anyway (though this never actually happens with optimality), or another coalition coordinates (which could not be helped), or one of the attempted coalitions coordinates first (which increases payoffs). Therefore, once again, no optimal conversation will remain at σ, and thus, it could not have been agreeable. □

The intuition behind Part (1) is particularly simple in two player games. In this case, given a strong prior forecast, either player can insist on the original equilibrium σ for longer than the other player can credibly hold out against it (by the definition of Nash). Therefore, both players must optimally be able to get at least their payoff from σ. However, since σ is efficient, this means that both players get exactly this payoff under any optimal strategies, and thus, staying at σ itself is as good as anything else. The examples in the next section serve to illustrate the mechanisms behind both the definitions and the proof of the theorem. It should be pointed out that in most specific cases, very little of the somewhat complex machinery developed above is necessary or applicable; the process is often hopefully quite natural and intuitive.

5. Examples

The most obvious example of an equilibrium selection problem is posed by the following coordination game:

	A	B
A	$2,2$	$0,0$
B	$0,0$	$1,1$

Of the three Nash equilibria in the game, only one is efficient. There is also an inefficient equilibrium, and this type of coordination problem comes up often in many contexts—including viral pandemics (see, e.g., Jnawali et al. 2017 [22]). Although in scenarios without communication, it is possible for (B,B) to occur, Theorem 2 implies that the efficient equilibrium (A, A) is the only possible outcome after rational nonbinding communication takes place among the players, no matter the prior forecasts. This is easy to see if either of the forecasts puts significant weight on A. In that case, the other player can credibly repeatedly announce A as the best response and, in this manner, eventually force the only credible announcement by either player to be A. Since this yields the highest possible payoff, it is optimal, and the conversation will converge to A.

If instead the prior forecasts are both heavily skewed toward B, then each player can reason as follows: "If I announce B, we will be stuck there forever, and I will get a payoff of one. If I announce A, there is some chance that my opponent will announce B, in which case, we will get stuck, and I will receive one. However, there is also some chance that my opponent will announce A. If we both continue to do this, these will remain credible announcements (since they each best respond to the other's appearance), and we will converge to the efficient equilibrium, delivering me a payoff of two instead of zero or one. I can always go back to announcing B and force that equilibrium (or start over altogether), so there is no risk of ending up at the really inefficient mixed equilibrium. Since there are no instantaneous payoffs lost from miscoordination along the way, the only possible optimal strategy is for me to announce A."

Both players are rational, so they will in fact both announce A at all rounds of the cheap talk communication, and the conversation will end up converging to the efficient equilibrium. Given that the forecasts were heavily skewed toward B, it may be a long time before the two players have truly convinced each other of their intention to play A, but they have all the time in the world and every reason to make use of it. If we looked instead at the pure coordination game in which (A, A) also yields payoffs of one to each player, the analysis is slightly changed. If the prior forecasts lean toward either of the symmetric and efficient pure equilibria, the conversation will converge in that direction. However, if the priors miscoordinate just right (for example, they are completely uniform for both players), it will be necessary for both players to randomize their initial announcement. If they coordinate at that point, successful convergence follows. If not, they simply clean their slate, start over, and try again. At some point, they must (that is, with probability one) both choose the same action (this is why it is necessary to randomize rather than to try to coordinate in some deterministic pattern), and then, they are done.

A less clear-cut example with a unique efficient equilibrium is found in the following version of the "stag-hunt" game:

	S	R
S	5, 5	0, 4
R	4, 0	3, 3

Here, the unique efficient equilibrium involves choosing a risk-dominated action, perhaps making it more difficult to reach. Allowing communication, however, will afford the players an opportunity to convince each other that it is safe to play action S. [23] has argued to the contrary that cheap talk may not help in this game. His reasoning is that since each player would prefer the other to take action S, they should each attempt to convince the other player to choose it. The way to do this is by claiming that you yourself are also going to pick S. Therefore, hearing the other player announce S should be discounted as purely manipulative and ignored.

It seems that Aumann's argument is not self-evident, at least when there is an unlimited chance to communicate. Rational players know that they will eventually agree on a Nash equilibrium; there is zero probability of suckering the other player or miscoordinating. At this point, it comes down to a choice among equilibria. Knowing this perfectly in advance, if a player announces S, it must be because he or she is hoping to eventually end up at the efficient equilibrium, that is to end up playing S. It is, after all, the best response at that point. In any case, the data clearly support the idea that allowing preplay messages increases the probability of observing the efficient, but risk-dominated equilibrium; see Charness (2000) and Miller and Moser (2004) [24,25].

We turn our attention next to the Battle of the Sexes, which is not at all a game with common interests:

	F	B
F	2, 1	0, 0
B	0, 0	1, 2

In this case, it is not immediately obvious that even with communication, efficiency can necessarily be achieved. If the prior forecasts favor either one of the pure equilibria, then the player who prefers that equilibrium will be able to credibly "insist" on it, and it will be the ultimate limit of the conversation. If the forecasts are balanced, however, neither player can be assured of getting his/her preferred outcome. Insisting on it whenever possible may lead the conversation to converge toward the inefficient mixed equilibrium, which is worse for both players. Therefore, this strategy is not optimal. If instead, the players "yield" to the other player with some extremely small probability at each round, this will always be achieved within ε of any other strategy, and since it always leads to one of the efficient equilibria, it weakly dominates the strategy by a player that forever insists on getting his or her way. Thus, under this scenario, the players are behaving optimally and can achieve efficiency with certainty.

As a final example, we turn to games with three players in order to explain some of the added complexity that arises. First, consider the following game in which the matrix player's payoffs are listed last:

A	L	R
U	0, 0, 10	−5, −5, 0
D	−5, −5, 0	1, 1, −5

B	L	R
U	−2, −2, 0	−5, −5, 0
D	−5, −5, 0	−1, −1, 0

This game has two pure Nash equilibria, namely (U, L, A) and (D, R, B), only the first of which is efficient. The second equilibrium is directly attainable from the first through a coalition of the row and column players, but it is not fully attainable because they enjoy a lower payoff in this equilibrium. Thus, the first equilibrium is stably efficient (and hence, the second, dominated one cannot be) and will be the result of rational communication. Nevertheless, since the row and column payoffs would be higher at the intermediate point along the chain fixing the matrix player at A, the original efficient equilibrium is not coalition-proof. Now, modify the payoffs slightly:

A

	L	R
U	$2, 2, 10$	$-5, -5, 0$
D	$-5, -5, 0$	$1, 1, -5$

B

	L	R
U	$-2, -2, 0$	$-5, -5, 0$
D	$-5, -5, 0$	$3, 3, 0$

Only the equilibrium payoffs have been changed, but the analysis has been affected greatly. Both pure equilibria are now efficient, but for exactly the reasons outlined above, only the second one, (D, R, B), is stably efficient and can be the result of cheap talk. On the other hand, the original equilibrium is now coalition-proof, showing the discrepancy between the two concepts.

One of the (unavoidable) limitations of this model is that it can say nothing about zero-sum games, except that communication can only converge to a Nash equilibrium. Other games in which all equilibria are efficient, and so for which Theorem 2 is vacuous, are games with a unique Nash equilibrium. These include matching pennies, rock-paper-scissors (where many of the convergence problems of fictitious play show up), and the game-theoretic standby of the prisoner's dilemma. Of course, we cannot expect that simple communication would lead to cooperation, a strictly dominated strategy. We have assumed throughout that there is only a single (though unlimited) chance for the players to talk for playing a game. If **G** is a repeated game and the players have a full conversation between each stage, then optimal speech should lead to efficient outcomes all along the extensive form game tree, both on and off the equilibrium path. This gives rise to the difficult problem of finding renegotiation-proof equilibria[18].

6. Conclusions

Coordination games of various forms, from actual rendezvous games to super-modular games and complementarity games, have received increasing attention in the game theory literature. Most equilibrium selection in such games, however, has been relatively informal, appealing to such concepts as focal points, initial conditions, or competition (essentially an evolutionary argument). Cheap talk, meaning costless and nonbinding preplay communication, has presented an intuitively pleasing method for formally attacking the equilibrium selection problem. The model of *conversations* presented here attempts to provide one possible resolution to this question of equilibrium selection, as well as to the even older question of justifying the Nash equilibrium concept.

The model assumes that players meet for the first time and communicate in order to allay their uncertainty about the future actions of their opponents. Since they have no knowledge of the cheap talk strategies used by the other players, we do not look for an actual equilibrium of the extended game. Instead, we look for all outcomes that could reasonably occur as the result of rational communication on the part of the players. Messages are defined to be credible in the context of a particular conversation. If at the end of a conversation, a player has put forward a consistent and credible appearance, this is assumed to in fact be the other players' belief about his or her future actions. From this base, it is proven that meaningful communication (that is, in which there is convergence) must end up at a Nash equilibrium. This is a partial justification for the Nash concept. It is then proven that optimal

[18] See, for example, the survey paper by Bergin and MacLeod (1993) [26].

communication, meaning that all players make strategic and rational announcements, leads to the deselection of inefficient equilibria.

A strength of the paper is that it gives a decisive answer to these two issues within the context of a single model. It also applies to games with more than two players or that do not necessarily exhibit common interests. There are, however, several qualifications to the model. First, the results do not prove that convergence must take place, only that if it does, then it takes a certain form. Secondly, since by no means all applications allow the possibility for preplay communication, this cannot be a general justification for the Nash concept. Indeed, this also potentially predicts a distinction between environments where one would expect Nash equilibrium to obtain versus others where one would not necessarily expect it. Finally, the model does put restrictions on the belief formation process, in that it requires some very small amount of faith to be put in credible announcements, at least over the long run. Note that this is not a departure from full rationality; traditional models have simply left this process unmodeled. There are also a number of possible relevant extensions of this model, notably to correlated equilibrium and to introducing a stochastic element in the conversation.

Calvin Coolidge once wisely said, "It is better to remain silent and be thought a fool than to speak and prove it." However, that applies only to fools: the moral of this paper is, "It is worse to remain silent and only be supposed rational than to speak and confirm it."

Funding: This research received no external funding.

Acknowledgments: I would like to thank Max Amarante, In-Koo Cho, Lynn Conell-Price, Peter Diamond, Glenn Ellison, Ehud Kalai, Elizabeth Murry, Tom Palfrey, Lones Smith, Radhanjali Shukla, and the seminar participants at Caltech, MIT, Penn State, Rochester, and the 17th Arne Ryde Symposium for valuable comments. All remaining shortcomings are mine only.

Conflicts of Interest: The author reports no conflict of interest.

References

1. Osborne, M.J. *An Introduction to Game Theory*; Oxford University Press: New York, NY, USA, 2004; Volume 3.
2. Crawford, V.P.; Sobel, J. Strategic information transmission. *Econom. J. Econom. Soc.* **1982**, *50*, 1431–1451. [CrossRef]
3. Farrell, J. Cheap talk, coordination, and entry. *RAND J. Econ.* **1987**, *18*, 34–39. [CrossRef]
4. Farrell, J.; Gibbons, R. Cheap talk can matter in bargaining. *J. Econ. Theory* **1989**, *48*, 221–237. [CrossRef]
5. Forges, F. Universal mechanisms. *Econom. J. Econom. Soc.* **1990**, *58*, 1341–1364. [CrossRef]
6. Farrell, J. Meaning and credibility in cheap-talk games. *Games Econ. Behav.* **1993**, *5*, 514–531. [CrossRef]
7. Aumann, R.; Hart, S. *Polite Talk Isn't Cheap*; Technical Report; Mimeo, Hebrew University of Jerusalem: Jerusalem, Israel, 1993.
8. Blume, A.; Sobel, J. Communication-proof equilibria in cheap-talk games. *J. Econ. Theory* **1995**, *65*, 359–382. [CrossRef]
9. Farrell, J.; Rabin, M. Cheap talk. *J. Econ. Perspect.* **1996**, *10*, 103–118. [CrossRef]
10. Rabin, M. A model of pre-game communication. *J. Econ. Theory* **1994**, *63*, 370–391. [CrossRef]
11. Matsui, A. Cheap-talk and cooperation in a society. *J. Econ. Theory* **1991**, *54*, 245–258. [CrossRef]
12. Canning, D. Learning and efficiency in common interest signaling games. In *The Dynamics of Norms*; Bicchieri, C., Jeffrey, R.C., Skyrms, B., Eds.; Cambridge University Press: Cambridge, UK, 1997; pp. 67–85.
13. Sandroni, A. Reciprocity and cooperation in repeated coordination games: The principled-player approach. *Games Econ. Behav.* **2000**, *32*, 157–182. [CrossRef]
14. Crawford, V.P.; Haller, H. Learning how to cooperate: Optimal play in repeated coordination games. *Econom. J. Econom. Soc.* **1990**, *58*, 571–595. [CrossRef]
15. Young, H.P. The evolution of conventions. *Econom. J. Econom. Soc.* **1993**, *61*, 57–84. [CrossRef]
16. Kalai, E.; Lehrer, E. Rational learning leads to Nash equilibrium. *Econom. J. Econom. Soc.* **1993**, *61*, 1019–1045. [CrossRef]
17. Cooper, R.; De Jong, D.V.; Forsythe, R.; Ross, T.W. Forward induction in coordination games. *Econ. Lett.* **1992**, *40*, 167–172. [CrossRef]

18. Brandts, J.; Cooper, D.J. It's what you say, not what you pay: An experimental study of manager-employee relationships in overcoming coordination failure. *J. Eur. Econ. Assoc.* **2007**, *5*, 1223–1268. [CrossRef]
19. Cachon, G.P.; Camerer, C.F. Loss-avoidance and forward induction in experimental coordination games. *Q. J. Econ.* **1996**, *111*, 165–194. [CrossRef]
20. Boulu-Reshef, B.; Holt, C.A.; Rodgers, M.S.; Thomas-Hunt, M.C. The impact of leader communication on free-riding: An incentivized experiment with empowering and directive styles. *Leadersh. Q.* **2020**, *31*, 101351. [CrossRef]
21. Bernheim, B.D.; Peleg, B.; Whinston, M.D. Coalition-proof nash equilibria i. concepts. *J. Econ. Theory* **1987**, *42*, 1–12. [CrossRef]
22. Jnawali, K.; Morsky, B.; Bauch, C.T. Strategic Interactions in Antiviral Drug Use during an Influenza Pandemic. 2017. Available online: http://currents.plos.org/outbreaks/index.html%3Fp=65448.html (accessed on 3 August 2020).
23. Aumann, C. Nash Equilibria are not Self-Enforcing. In *Economic Decision Making: Games, Econometrics, and Optimization (Contributions in Honor of Jacques Drèze)*; Elsevier: Amsterdam, The Netherlands, 1990.
24. Charness, G. Self-serving cheap talk: A test of Aumann's conjecture. *Games Econ. Behav.* **2000**, *33*, 177–194. [CrossRef]
25. Miller, J.H.; Moser, S. Communication and coordination. *Complexity* **2004**, *9*, 31–40. [CrossRef]
26. Bergin, J.; MacLeod, W.B. Efficiency and renegotiation in repeated games. *J. Econ. Theory* **1993**, *61*, 42–73. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Games Editorial Office
E-mail: games@mdpi.com
www.mdpi.com/journal/games

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-03943-774-0