2006s-12

Who Gets the Last Word? An Experimental Study of the Effect of a Peer Review Process on the Expression of Social Norms

Jim Engle-Warnick, Andreas Leibbrandt

Série Scientifique
Scientific Series

Montréal
Août 2006

CIRANO
Centre interuniversitaire de recherche en analyse des organisations

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du Ministère du Développement économique et régional et de la Recherche, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère du Développement économique et régional et de la Recherche, and grants and research mandates obtained by its research teams.

Les partenaires / The Partners

Partenaire majeur

Ministère du Développement économique,
de l'Innovation et de l'Exportation

Entreprises partenaires

Alcan inc.
Banque du Canada
Banque Laurentienne du Canada
Banque Nationale du Canada
Banque Royale du Canada
Bell Canada
BMO Groupe financier
Bombardier
Bourse de Montréal
Caisse de dépôt et placement du Québec
Fédération des caisses Desjardins du Québec
Gaz Métro
Hydro-Québec
Pratt & Whitney Canada
Raymond Chabot Grant Thornton

Autres partenaires gouvernementaux

Industrie Canada
Ministère des Finances du Québec
Ville de Montréal

Partenaires universitaires

École Polytechnique de Montréal
HEC Montréal
McGill University
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web.

ISSN 1198-8177

Who Gets the Last Word? An Experimental Study of the Effect of a Peer Review Process on the Expression of Social Norms*

Jim Engle-Warnick†, *Andreas Leibbrandt* ‡

Résumé / *Abstract*

Au cours de trois jeux de confiance différents, nous alternons la personne qui aura le dernier mot sur le résultat : le premier joueur, le deuxième joueur, ou un comité (à qui revient la décision) composé du premier et du second joueur. Ce comité fonctionne de manière similaire à la révision conventionnelle par les pairs, où des joueurs expérimentés passent un jugement sur les résultats préalablement réalisés par deux joueurs différents. Étonnamment, donner le dernier mot au premier joueur donne l'avantage au deuxième joueur. D'autre part, laisser le comité prendre la décision augmente la confiance du premier joueur. Finalement, les premiers et seconds joueurs passent différents types de jugements lorsqu'ils font partie d'un comité.

Mots clés : confiance, économie expérimentale, normes sociales, préférences sociales, révision par les pairs, sanctions par une tierce personne

We alter who gets the last word on the outcome in three different types of trust games: the first mover, the second mover, or, a committee comprised of first and second movers. The committee functions in a manner similar to a peer review process, in which experienced subjects pass judgment on the outcome reached by a different pair of subjects. Surprisingly, giving the first mover the last word benefits the second mover. Letting the committee decide increases the first mover's trust. And first and second movers pass different types of judgments when they act as a committee.

Keywords: *experimental economics, peer review, social norm, social preferences, third-party punishment, trust*

Codes JEL : C92

* We thank Julie Héroux, and Alexander Campbell for research assistance and the participants of the Greifensee Seminar for helpful comments. We gratefully acknowledge the Social Science and Humanities Research Council and the ENABLE network for financial support, and the Centre for Interuniversity Research and Analysis on Organizations and the Bell University Laboratory in Electronic Commerce and Experimental Economy for use of the experimental laboratory.

† CIRANO and Department of Economics, 855 Sherbrooke St. W., McGill University, Montreal, Québec, H3A 2T7, Canada, e-mail: jim.engle-warnick@mcgill.ca, tel.: (514) 398-1559; fax: (514) 398-4938.

‡ University of Zurich, Institute for Empirical Research in Economics.

Introduction

A wealth of experimental economics literature has recently developed involving the study of social preferences and social norms (Camerer 2002 provides a survey; Fehr and Gächter, 2002; Henrich et al, 2001). The motivation behind many studies begins with the observation that human societies can hardly exist without social norms, that much field evidence indeed shows their existence (e.g. Elster 1989; Bandeira et al, 2005; Conlin et al, 2003), and that the experimental laboratory provides the control which is lacking in the field to distinguish between conflicting explanations for their existence and effectiveness.

One direction that this research has taken involves the study of third-party sanctions. The idea is that norms, i.e. behavioural regularities, develop when third parties, who do not have a stake in a conflict, inflict punishment upon an individual for the violation of a norm. A typical third-party sanction experiment introduces a third person into an existing game. The outcome of the third party is not dependent upon the relationship between the other players. The third party observes the actions of the other players and is then given an opportunity to punish them for their behavior, i.e. to exercise community enforcement (Kandori, 1992) or indirect reciprocity (Nowak and Sigmund, 2005), usually at some cost to herself.

The evidence in these experiments points strongly to the use of third-party interventions in a variety of contexts. Fehr and Fischbacher (2004), for example, find third-party sanctions in a prisoner's dilemma and dictator game. Seinen and Schram (2006) observe that third parties reward second parties according to developed group norms. Ottone (2004) extends results in the dictator game to show that changes in roles

played by subjects influence decisions to punish, as does the level of the personal cost of the punishment. And Carpenter and Matthews (2004) show that subjects are willing to punish players in groups other then their own in public goods games.

We introduce a type of third-party intervention that involves giving a committee the last word on the outcome in a game. Most of us prefer to receive the last word and folk wisdom supports us: "S/he who laughs last, laughs best", and this is the basis of our study. Standard game theory tells another story in many types of social interactions. Take, for example, the ultimatum game: although the responder moves last, she should accept any offer from the proposer, not putting any weight on the last word. However, economic experiments (Camerer, 2003 provides a survey) have shown that the last word has some power in this game, and a wealth of literature has recently resulted in models that incorporate social preferences and social norms that can account for similar experimental findings in other games (e.g., Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Fehr and Schmidt, 1999).

We report results from an experimental study of three different games, designed to elicit social preferences, in which different players receive the last word: (1) the first mover, (2) the second mover, and (3) a committee comprised of first and second movers. The basic game we study is the Moonlighting Game (Abbink et al, 2000), which is itself a modification of a trust game (Berg et al, 1995). In this two-stage game, an A-player decides whether to send money to or take money away from a B-player. If she sends money, it is multiplied by a constant amount and given to the B-player; if she takes money, it is deducted from the B-player's balance with the same amount added to her own balance. The B-player then decides whether to send money back to or take money

away from (at a cost to herself, possibly representing a costly punishment) the A-player. A feature of this game is that in each stage subjects may take positive or negative actions, thus the game is in some ways more realistic than a trust game (in which the first stage action is restricted to be zero or positive) with regard to some types of relationships, and subjects are not primed to act in the direction of trust or punishment.

For our second game, we add a third stage to the Moonlighting Game, which is a replication of the second stage, but with the A-player making the final decision; we call this the Three-Stage Moonlighting Game. For our third game, we modify the third stage of the Three-Stage Moonlighting Game, replacing the A-player with a committee. This committee is comprised of one randomly selected A-player and one randomly selected B-player, each from a different pair of players who have also made their own decisions in the first and second stages of their own game. The committee views the game history of the other pair, and makes the final decision as if they are the A-player in the Three-Stage Moonlighting game. The average decision is implemented. We call this game the Committee Moonlighting Game. Notice that while players are deciding how to end the game for other pairs, different pairs are also deciding for them.

In our Committee Moonlighting Game punishment reduces the payoff of the *sanctioned* player, but it is not costly for *sanctioning* player, and since A- and B-players decide on sanctions for other A- and B-players, the sanctions are performed by peers. This type of punishment is akin to, say, a peer-review process upon which the progression of much of scientific inquiry is based, or a disciplinary meeting, where peers take decisions in disciplinary hearings, and cast a vote on the outcome of the proceedings.

Peer review heavily relied upon to enforce norms of scientific behavior; we employ it to elicit norms in our game.[1]

Peer review has been studied theoretically in a variety of contexts. Perhaps the closest context to our experiment is Ellison (2002), who presents a theory of evolving standards in academic publishing under a peer review regime. Standards evolve along two dimensions: the quality of the main ideas of the research and other qualities. He shows that researchers with a biased view of the main ideas of their own work, i.e., with a biased view of their own type, and who learn from being reviewed themselves, cause an increase in emphasis on the "other qualities" dimension. Baliga and Sjostrom (2001) show that self assessment, with no peer reporting, is optimal when a principal must decide whether to implement a project originating with one of her employees, several employees know the quality of the project, and the employee's promotion upon a successful project is at the expense of her peers. And Che and Yu (2001) show that teams, in which peer sanctioning is available, can be optimal ways of organizing a workforce in a repeated game.

We find that it matters who has the last word on the outcome of the game. In the Two-Stage Moonlighting Game, A-players earn significantly more than B-players, whereas this difference vanishes in the Three-Stage Moonlighting Game and even reverses in the Committee Moonlighting Game, where B-players are in a much better position. Moreover, we observe that the level of trust increases significantly in the Committee Moonlighting Game, suggesting that our committee decision represents a

[1] Our experiment is comparable to other experiments that seek to understand the effect of institutions on behavior where trust is needed to get transactions to occur (Andreoni, 2005; Bohnet et al, 2001; Bohnet and Zeckhauser, 2004; Falk and Kosfeld, 2005). What is different in our paper is the peer-review mechanism.

unique mechanism where players increase trust even when they do not get additional information or another punishment possibility.

The next section presents the experimental design, which is followed by our presentation of the experimental results. We then discuss the results before concluding.

Experimental Design

The three games we study all consist of identical first and second stages. This part of the game is called the Two-Stage Moonlighting Game (2ML), introduced by Abbink et al (2000). The story of the game goes that an illegal worker (a "moonlighter") chooses either to steal money out of the till or perform at some effort level at work. The employer then decides whether and how much to pay the moonlighter for work, or to try to pursue some kind of costly punishment if she perceives the moonlighter deserves it; the punishment is always costly because the worker is illegal. This game was attractive to us because since it provides the possibility of both positive and negative actions toward the other player at each stage of the game, it reduces the possibility that the framing of the game would prime subjects' behavior in one direction or another.

The Two-Stage Moonlighting Game

In the Two-Stage Moonlighting Game (2ML) we call the first player the A-player and the second player the B-player. The A and B–player both start with the same endowment of $12. In the first stage, the A–player can either choose to increase her own balance by up to $6, at the same time decreasing the B–player's balance by the same amount, or she can choose to decrease her own balance by up to $6 at the same time

increasing the B–player's balance by triple the amount. In the second stage, the B–player can either choose to decrease her own balance by an amount, at the same time decreasing the A–player's balance by triple the amount (this is a costly punishment or a sanction), or she can choose to decrease her own balance by an amount at the same time increasing the A–player's balance by the same amount (this is a costly reward). All strategies are constrained so that neither player's balance can go below zero. A simplified version of the structure of the game is depicted in Figure 1a. In the figure, the strategy space is reduces to sending money to or taking money from the other player. The figure reinforces the fact that both positive and negative actions are available to each player at each stage of the game. Note that players can also choose to do nothing at each stage. The result is a game that is neutral with respect to framing of rewards and punishments.

The Three-Stage Moonlighting Game

Figure 1b presents a simplified version of the Three-Stage Moonlighting Game (3ML). In 3ML we add a third stage to 2ML. In this new third stage, the A–player can either choose to decrease her own balance by an amount, at the same time decreasing the B–player's balance by triple the amount (a costly punishment to the B-player), or she can choose to decrease her own balance by an amount at the same time increasing the B–player's balance by the same amount (a costly reward to the B-player). The third stage is thus identical to the second stage, with the A-player facing the same choices that the B-player faced in the second stage. As in 2ML, the constraint is imposed that no player's balance can go below zero.

The Committee Moonlighting Game

In the Committee Moonlighting Game (CML), each A–player is, in the third stage, randomly paired with a different B–player to form a committee, and both make a decision simultaneously for another pairing that has completed stages one and two. The two members of the committee choose independently either to decrease the balance of the A–player of this pairing by an amount, at the same time decreasing the balance of the B–player of this pairing by triple the amount (a punishment to the B-player), or to decrease the balance of the A–player of this pairing by an amount at the same time increasing the balance of the B–player of this pairing by the same amount (a reward to the B-player that costs the A-player). The decisions of the A- and B-player within the committee do not affect their own balances. The two decisions within the committee are averaged to determine the final balances of the other pairing. The third stage of 3ML is thus identical to the third stage of CML, where the committee faces the same strategy alternatives that the A-player faced in 3ML. Figure 1c presents an example of the way a pairing might work for the composition of the committee.

Overview of the Games

The difference between 2ML and the other two games is the addition of the third stage, which has the same strategy space as the second stage in 2ML. The difference between 3ML and CML is who is making the decision for the player pairs: Player A or a third party consisting of a committee of players who are not themselves affected by their choices. Notice that in addition to making decisions that positively or negatively affect the other player's balance, in all three games each player has the option to not affect the

other's balance at all stages of the game. At no time can the A or B–players make a decision that results in a negative balance for either player.

Predictions

Theoretical Predictions for 2ML and 3ML

Standard game theory assumes that people are self-interested, hence there is no reason for trust or reciprocity in either game. Using backward induction, the player in the last stage whether A or B should always send zero (since any other action is costly), thus the A-players take maximally. This results in payoffs of $18 for A-players and $6 for B-players. Social preference models like the Fehr-Schmidt inequity-aversion model, however, can explain trust and reciprocal behaviour in both games. According to their model, players behave reciprocally if they are sufficiently inequity-averse, making it rational for an A-player to trust if she believes to be paired with sufficiently inequity-averse B-player.

Theoretical Predictions for CML

The experience playing the first two stages of the game (the equivalent of 2ML) should give subjects the opportunity to form an opinion of what the last word should be in the game, thus with this mechanism we allow for A and B-players to express and implement their opinion. We conjecture that players will use the possibility to enforce a social norm, and that this expression may be different for the A-players in CML than in 3ML, since it is costly for the sanctioned, but not for themselves. Using Ellison's (2002) insights, we might expect a subject's decision to depend on her type. This would lead us to expect Players A and B to behave differently when making committee decisions.

Experimental Procedures

We conducted our experiments on a computer network using the software Z-tree (Fischbacher, 1999). Strategies were limited to integer numbers for simplicity, and the program automatically enforced nonnegative balance constraints. At the beginning of each session, each participant was randomly assigned to be an A-player or a B-player. All procedures were common knowledge. The subjects played the game for exactly one round to render impossible repeated-game effects.

Altogether 146 subjects participated in our experiment, which consisted of four sessions of each game type. There were 46 participants in 2ML, and 50 participants each in 3ML and CML. No subject participated in more than one session. All decisions were conducted at the Bell Labs at the Centre for Interuniversity Research and Analysis on Organizations in Montréal, Quebec, Canada. The sessions lasted approximately fifty minutes and subjects received on average $21.80 CAD including a $10 show up fee (which partially compensates for travel costs to the off-campus laboratory). All participants were English-speaking students of different universities in Montréal.

Experimental Results

Aggregate Results

Table 1 presents aggregate results, averaged across subjects and subject pairs. The first column of the table reveals that in 2ML the average first decision by the A-Player is -1.39, i.e., on average the A-player take from the B-player's balance in stage 1. By contrast, the average decision by the B-player in 2ML is a positive 2.09, indicating

that the B-players increase the A-player balances. A-players earn a higher average payoff (12.78 vs. 9.23; t = 1.668, p-value = 0.05), although B-players have the last word.

The second column of Table 1 shows the same results for 3ML. Notice the similarity in the balances of the two player types after the second stage ($12.84 for Player A and $12.04 for Player B), as well as the identical shares at the end of the game ($11.40 for both players types). Giving A-players the last word does not improve their outcomes; to the contrary, A-players earn on average 10% less. Paradoxically, B-players earn about 20% more through adding the third stage in 3ML. This leads us to Result 1:

Result 1: Giving A-players the last word results in a worse outcome for them.

The third column of Table 1 presents the average results for CML. Now A-players send a positive amount to B-players (2.6), and this is statistically different from what they send in 3ML (t = 2.332, p-value = 0.01). This provides us with Result 2:

Result 2: Introducing the committee review mechanism increases trust.

Sadly for the A-players, the extra trust does not pay off, as B-players do not send more back than in 3ML (t = 0.468, p-value = 0.32). This results in B-players having a higher balance at the end of the second stage than A-players (17.12 vs 12.00; t = 2.161, p-value = 0.02). Interestingly, the average committee decision (which is located on the row labelled "A-player second decision") does not correct the imbalance: B-players receive a final share of $16.96 vs. $8.96 for A-players (t = 4.823, p-value = 0.000). Thus B-

players benefit from the peer review process: they earn more in CML than in 3ML (t = 2.624, p-value = 0.006). For the A-players, the situation worsens even more: they earn less in CML than in 3ML (t = -1.357, p-value = 0.09).

Distribution of Play

Figure 2 presents histograms of the A-players' first decisions in 2ML, 3ML and CML. While in the first two games there is almost no difference in A-player behaviour, in CML one can clearly see a significant shift out of the negative strategies and into the positive ones. Figure 3 shows no such discernable shift for the B-players.

Figure 4 shows the decisions in the committees, i.e., the individual averages of the A- and B-player decisions. The average committee decisions differ from the second decisions A-players make in 3ML; this can be seen by comparing Figure 4 with the top panel of Figure 5. Average committee decisions are more dispersed around 0 than A-player second 3ML decisions are. The bottom panel of Figure 5 and Figure 6 investigate whether we can attribute the dispersion to one of the two player types. The answer is that A-player committee decisions exhibit less of a mode at 0, using the committee institution more to change distributions than B-players do (Kolmogorov-Smirnov two-sample test, p-value = 0.01). This is result 3.

Result 3: Introducing the committee review mechanism results in different actions taken by the different player types.

Finally, in Figure 7 we present histograms of decision pairs in all three games. The data are ordered in the figures by A-player decision: bars at the far left represent the

most negative A-player decisions and bars at the far right present the most positive decisions. For the most part these figures confirm that positive and negative actions tend to be paired together, as we would expect to see with reciprocal behaviour.

Regressions

We ran two-sided Tobit regressions on the decisions taken at the second and third stages of our games, taking into account left and right censoring of some of the decisions made by the subjects. The censoring occurs any time a subject makes an extreme decision: for example, a B-player who chooses -6 in stage two plays a left-censored strategy since she cannot choose a more negative decision. Our regression model uses past strategies taken by the players as conditioning variables for the selection of a strategy. The dependent variable is the player's own decision. Note that a positive action by a player represents a positive action toward the other player. Thus a positive coefficient on the other type's decision represents reciprocity behavior.

Table 2 presents results for A-players (and the average committee decision replacing the A-players) and B-players. Table 2a reveals that, for B-players in all three games the coefficient on A-player decision is positive and significant. The marginal effect of a dollar sent by an A-player is lower in 3ML than in the other two games, but the direction of the effect is the same in all three games, i.e. in the direction of reciprocity.

Table 2b presents coefficient estimates for the A-players` second decisions and average committee decisions in 3ML and CML. In 3ML, the marginal effect on the B-player's decision is not significant. By contrast, the average committee decision exhibits

significant reciprocity with regard to the B-player's decision. The peer review committee took reciprocal action where the A-player herself would not. This is interesting: recall that when the committee makes such a decision it can cost the A-player on whose behalf the decision is being made. This is Result 4:

Result 4: Introducing the committee review mechanism resulted in the elicitation of a social norm, which was not elicited in the absence of the committee.

Lastly, we present Table 3 to dig a bit deeper into the individual decisions within the committee. The table reports results from regressions run separately on A-players and B-players, and uses the history of decisions of both player types in both the player's own pairing (in-group) and in the pairing the player is making a committee decision for (out-group). The results show a difference between the two player types. While B-player decisions are significantly affected by all decisions made both within their own group and in the out-group for which they render their judgement, the A-players appear only to be affected by the out-group A-player's first decision.

Discussion of Results

At first blush it seemed to us counterintuitive that giving the A players the last word reduces their payoff in the game (Result 1), but it seems to us consistent with the well-known result in trust games where trust is positive and its expected value is negative. Perhaps the A-players expect the last word to improve their position in the game, but find it not worthwhile to use the costly punishment when it comes time to do so. There is some evidence for this from Result 4, where the A-players express themselves more in committee decisions, when it is not costly for them, than they do

when it is their own decision in game 3ML. The same reasoning holds true for the increase in trust which does not pay off when committees make decisions (Result 2). At any rate, the asymmetry of the roles induces a type of asymmetric behaviour reminiscent of the trust game, which can be viewed as a Moonlighting Game with a restricted strategy space. The fact that different types make different committee decisions (Result 3) is consistent with the conjecture that that their notion of the norm in the game is a function of their type. These conjectures could be tested with a symmetric game.

Conclusion

In this paper we alter who has the last word on the outcome of a type of a trust game: the A-player, B-player or a Committee. Our committee approach is akin to the implementation of a peer review procedure, in which subjects with experience in the game pass a costless judgement (with respect to their own outcome) on the outcome reached by another pair or players. One attractive aspect of our design is that it avoids certain features of alternative methods of investigating third-party behaviour. For example, most existing studies give the third party an endowment, which could be interpreted by the subjects as being meant to be used for the purpose of punishment. Our design avoids these difficulties while allowing subjects, who have experience in the game, express and implement their opinions regarding normative outcomes in the game. And basing our game on the Moonlighting Game provides players with both positive and negative actions towards the other player at each stage of the game.

We report four main findings: (1) giving the A-player the last word can lead to less desirable outcomes for her, (2) introducing committees can increase trust, (3) the

kind of decisions taken in the committees depend on the player type, and (4) introducing the committees can result in enforcement of a social norm where second parties hesitate.

We believe that our results are helpful for better understanding and modelling different peer review processes. Since there is still much to learn regarding third party behaviour in games, controlled laboratory experiments are needed to better understand how such mechanisms function. One can implement our mechanism with virtually any game, complementing existing methods of third-party sanctioning, in an effort to better understand what outcomes subjects believe should occur in a social situation.

References

Abbink, K., B. Irlenbusch and E. Renner, 2000, "The Moonlighting Game: An Experimental Study on Reciprocity and Retribution", *Journal of Economic Behavior and Organization,* 42:2:265-277.

Andreoni, J., 2005, "Trust, Reciprocity, and Contract Enforcement: Experiments on Satisfaction Guaranteed", mimeo University of Wisconsin-Madison.

Baliga, S. and T. Sjostrom, 2001, "Optimal Design of Peer Review and Self-Assessment Schemes", *Rand* 32:27-51.

Bandeira, O., I. Barankay and I. Rasul, 2005, "The Evolution of Cooperative Norms: Evidence from a Natural Field Experiment", CEPR Discussion Paper 5358.

Berg, J., J. Dickhaut and K. McCabe, 1995, "Trust, Reciprocity, and Social History", *Games and Economic Behavior,* 10:122-142.

Bohnet, I., B. Frey, and S. Huck, 2001, "More order with less law: On contract enforcement, trust, and crowding, *American Political Science Review*, 95:135-144.

Bohnet, I., R. Zeckhauser, 2004, "Social Comparisons in Ultimatum Bargaining", *Scandinavian Journal of Economics,* 106:3:495-510.

Bolton, G. and A. Ockenfels, 2000, "ERC: A Theory of Equity, Reciprocity, and Competition", *American Economic Review,* 90:166-193.

Camerer, C., 2003, *Behavioral Game Theory*, Princeton University Press: Princeton.

Carpenter, J. and P. Matthews, 2004, "Social Reciprocity", mimeo Middlebury College.

Charness, G. and M. Rabin, 2002, "Understanding Social Preferences with some Simple Tests", *Quarterly Journal of Economics,* 117:3:775-816.

Che, Y.K. and S.W Yoo, 2001, "Optimal Incentives for Teams", *American Economic Review*, 91:525-541.

Conlin, M., M. Lynn, T. O`Donoghue, 2003, „The norm of restaurant tipping", *Journal of Economic Behaviour & Organization*, 52:3:297-321.

Ellison, G., 2002, "Evolving Standards for Academic Publishing: a q-r Theory", *Journal of Political Economy*, 110:5:994-1034.

Elster, J., 1989, "Social Norms and Economic Theory", *The Journal of Economic Perspectives*, 3:4:99-117.

Falk, A. and U. Fischbacher, 2006, "A Theory of Reciprocity", *Games and Economic Behavior*, 54:2:293-315.

Falk, A. and M. Kosfeld, 2005, "The Hidden Costs of Control", *American Economic Review*, forthcoming.

Fehr, E. and U. Fischbacher, 2004, "Third Party Punishment and Human Behavior", *Evolution and Human Behavior*, 25:63-87.

Fehr, E. and S. Gächter, 2002, "Altruistic Punishment in Humans", *Nature*, 415:137-140.

Fehr, E. and S. Gächter, 2000, "Fairness and Retaliation: The Economics of Reciprocity", *Journal of Economic Perspectives*, 14:159-181.

Fehr, E. and K. Schmidt, 1999, "A Theory of Fairness, Competition and Cooperation", *Quarterly Journal of Economics*, 114: 817-868.

Fehr, E., G. Kirchsteiger and A. Riedl, 1998, "Gift Exchange and Reciprocity in Competitive Experimental Markets", *European Economic Review*, 42:1-34.

Fehr, E., S. Gächter and G. Kirchsteiger, 1997, "Reciprocity as a Contract Enforcement Device", *Econometrica*, 65:833-860.

Fischbacher, U., 1999, "Z-Tree: A Tool for Readymade Economic Economic Experiment", University of Zurich Working Paper 21.

Henrich, J. et al, 2001, "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies", *American Economic Review*, 91:73-78.

Kandori, M., 1992, "Social Norms and Community Enforcement", *Review of Economic Studies*, 59:63-80.

Kreps, D., P. Milgrom, J. Roberts and R. Wilson, 1982, "Rational Cooperation in the finitely repeated prisoners dilemma", *Journal of Economic Theory*, 27:245-252.

Nowak, M. and K.Sigmund, 2005, "Evolution of Indirect Reciprocity", *Nature*, 437:1291-1298.

Ottone, S., 2004, "Transfers and Altruistic Punishments in Third Party Punishment Game Experiments", mimeo Department of Public Policy and Public Choice (POLIS), University of Eastern Piemont "Amedeo Avogadro".

Seinen, I. and Schram, A, 2006, "Social status and group norms: Indirect reciprocity in a repeated helping experiment", *European Economic Review*, 50:581-602.

Table 1: Aggregate Statistics

	Game		
	Two Stage	Three Stage	Committee
A-Player First Decision	-1.39	-0.12	2.6
	(5.06)	(4.78)	(3.21)
B-Player Decision	2.09	2.32	3.08
	(6.10)	(5.94)	(5.32)
A-Player Second Decision	-	0.4	0.64
	-	(2.87)	(4.72)
A-Player Stage 2 Balance	12.78	12.84	12
	(6.85)	(7.14)	(4.96)
B-Player Stage 2 Balance	9.23	12.04	17.12
	(7.57)	(9.28)	(6.92)
A Share	12.78	11.4	8.96
	(6.85)	(7.31)	(5.02)
B Share	9.23	11.4	16.96
	(7.57)	(8.29)	(6.33
Efficiency	22	22.8	25.92
	(11.65)	(10.42)	(6.90)

Notes: A-Player Second Decision in Committee Game is
Average of A- and B-Player decisions.
Standard deviations in parentheses.

Table 2: B-Player and A-Player Regressions

Table 2a: B-Player Decision

	Game		
	Two Stage	Three-Stage	Committee
A-Player First Decision	1.01**	0.66**	0.92**
	(0.166)	('0.251)	(0.275)
Constant	3.36**	2.13*	0.69
	(0.848)	(1.164)	(1.125)
n	23	25	25
no. left-censored	2	3	0
no. right-censored	1	1	0
pseudo r-squared	0.15	0.04	0.06

Table 2b: A-Player Second Decision

	Game		
	Two Stage	Three-Stage	Committee
A-Player First Decision	-	-0.28*	-0.96**
		('0.152)	(0.262)
B-Player Decision	-	0.17	0.38**
		(0.122)	(0.158)
Constant	-	0.13	2.70**
		(0.676)	(0.887)
n	-	25	25
no. left-censored	-	2	1
no. right-censored	-	3	0
pseudo r-squared	-	0.03	0.08

Notes: *significant at 10% level; **significance at 5% level.
Standard errors in parentheses.
Two-sided censored regression.
Committee variable is the average of the A-Player and B-player decision

Table 3: Committee Decision Regressions

Individual Committee Decision

	Committee Game	
	A-Player	B-Player
In-Group A-Player First Decision	-0.12	-0.83**
	-0.3052	(0.374)
In-Group B-Player Decision	0.15	0.51**
	(0.189)	(0.227)
Out-Group A-Player First Decision	-0.92**	-0.98**
	(0.314)	(0.362)
Out-Group B-Player Decision	0.38	0.58**
	(0.229)	(0.203)
Constant	1.84	3.37**
	(1.246)	(1.513)
n	25	25
no. left-censored	0	0
no. right-censored	0	0
pseudo r-squared	0.05	0.07

Notes: Committee variable is the individual decision.
*significant at 10% level; **significance at 5% level.
Standard errors in parentheses.
Two-sided censored regression.

Figure 1a: The Two-Stage Moonlighting Game (2ML)

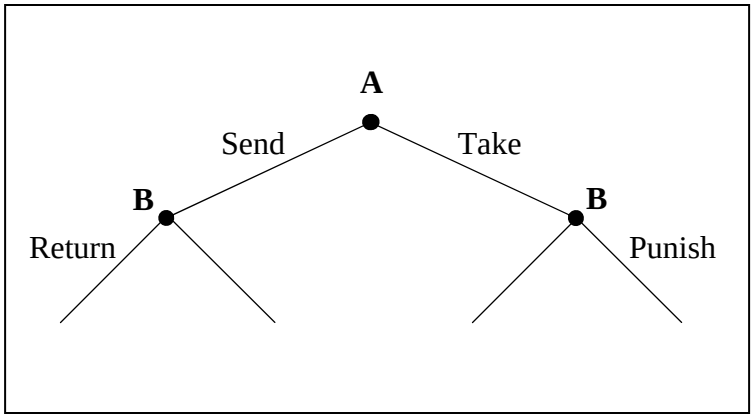


Figure 1b: The Three-Stage Moonlighting Game (3ML)

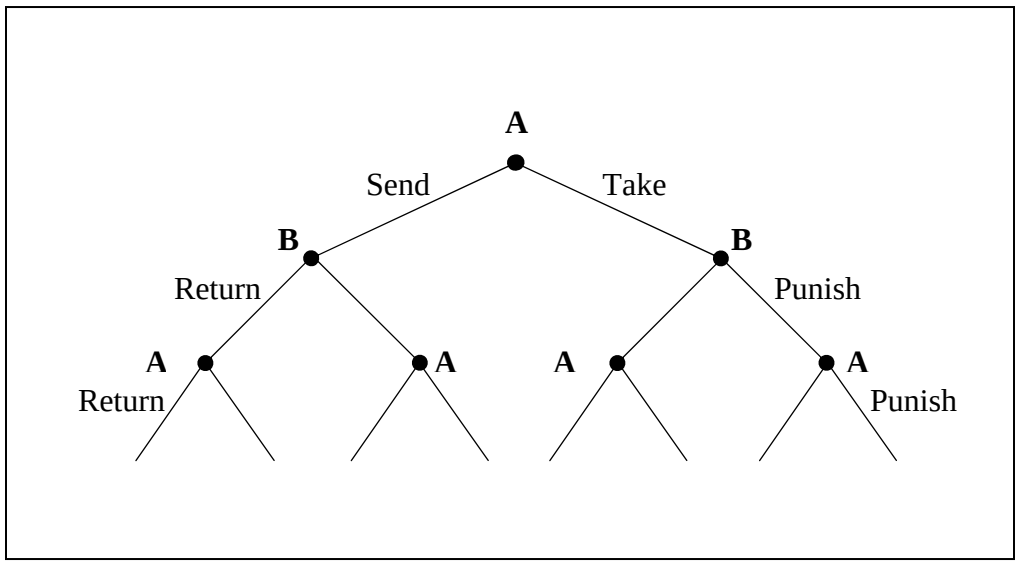
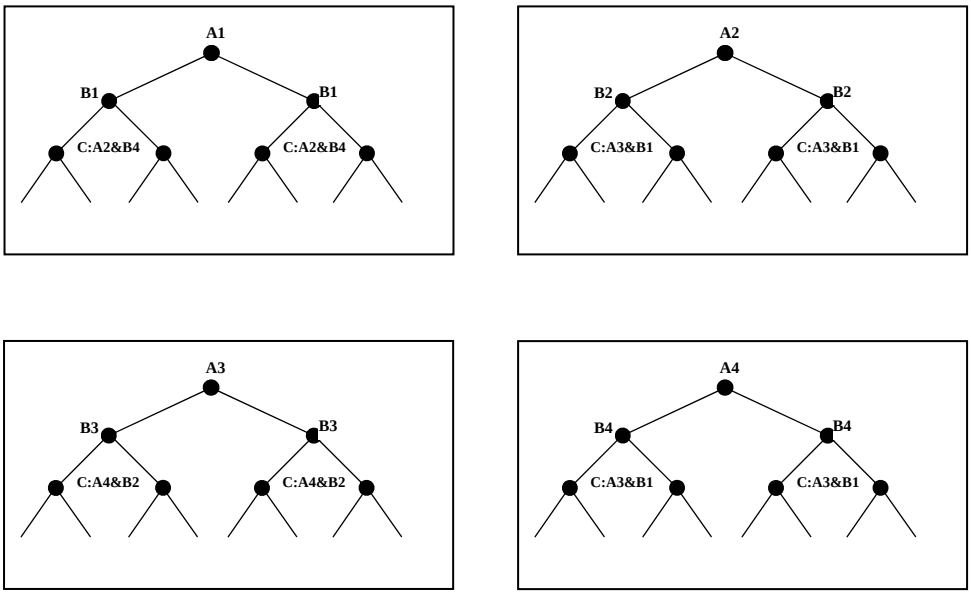
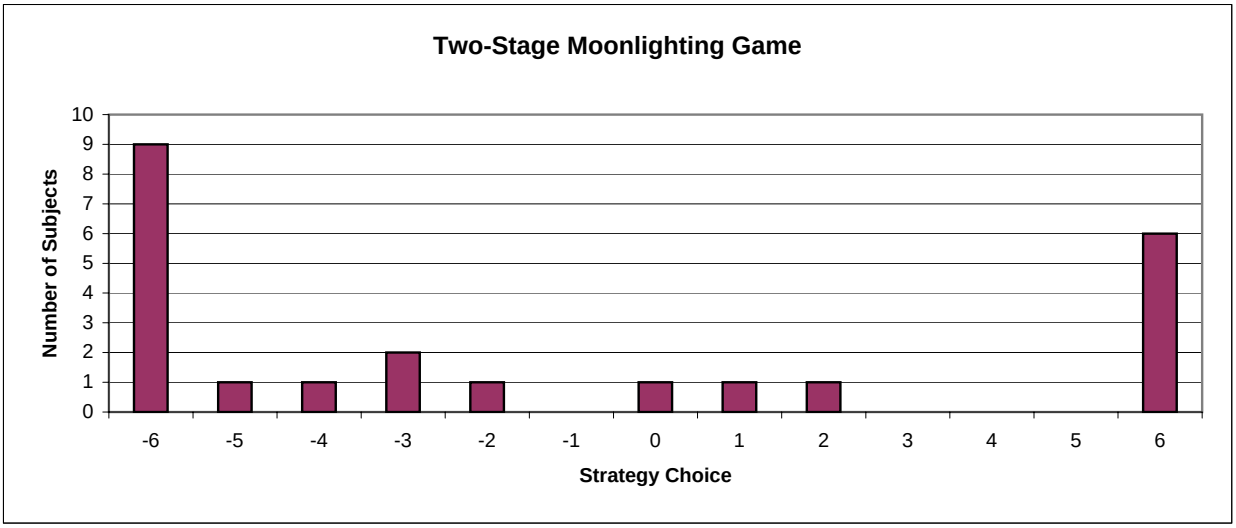


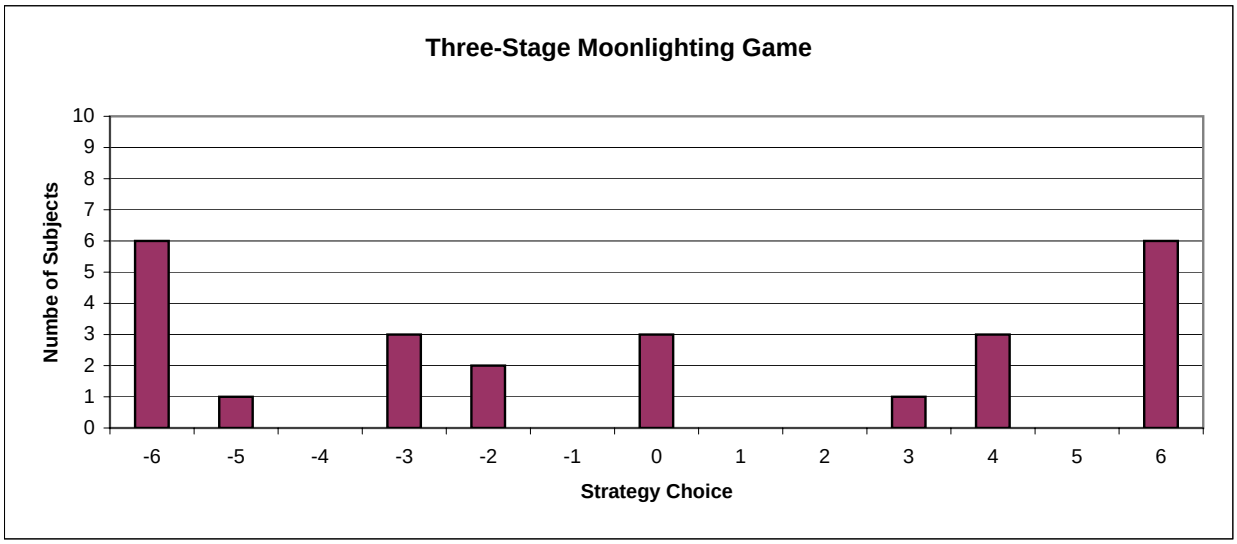
Figure 1c: The Committee Moonlighting Game (CLW)



Note: Figure 1c shows one possibility how committees may be formed, e.g. the A-player from pairing two and the B-player from pairing four form a committee and decide for the pairing 1 in the third stage.

Figure 2: Histogram of A-players' First Decisions





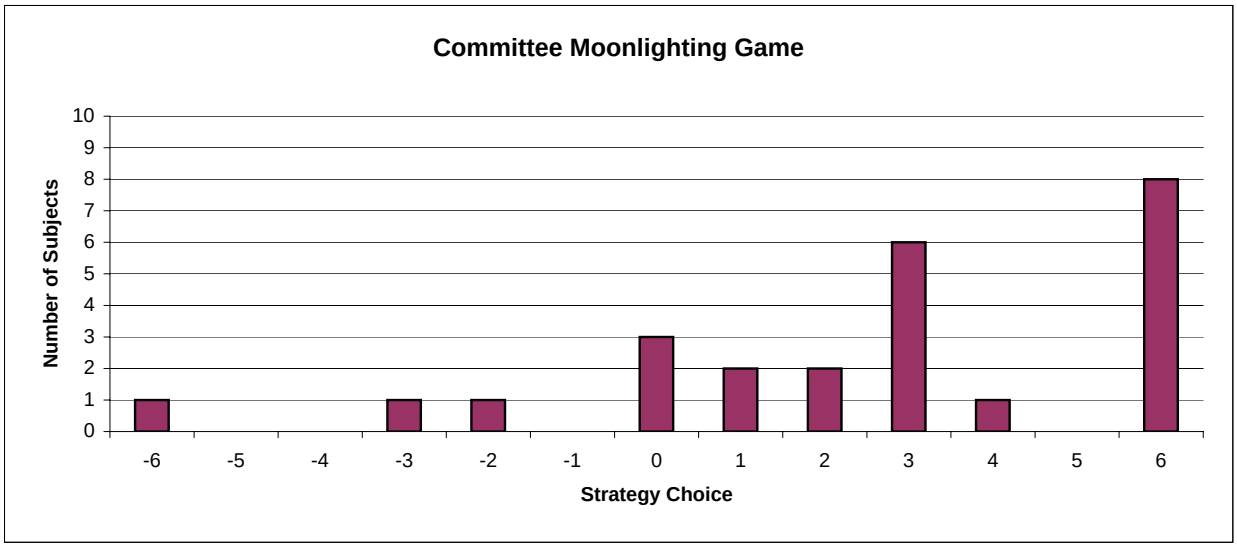
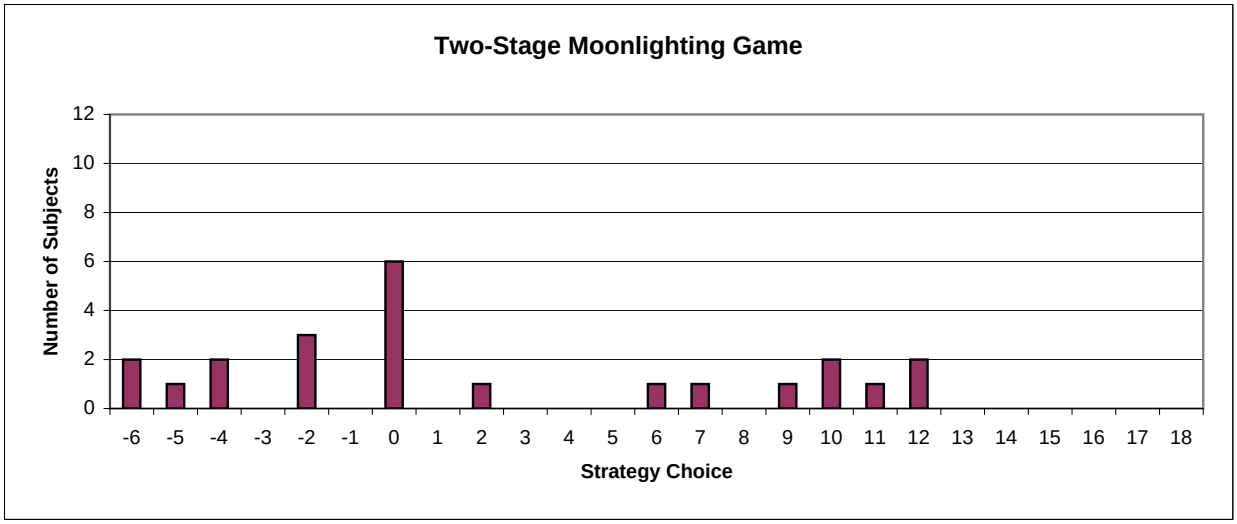
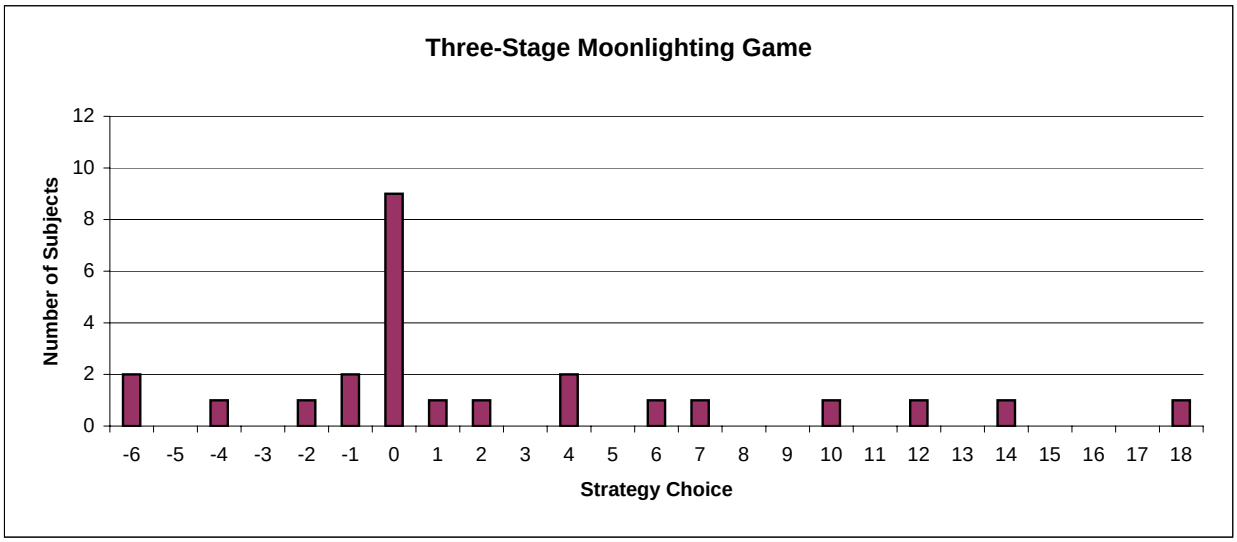


Figure 3: Histogram of B-players' First Decisions





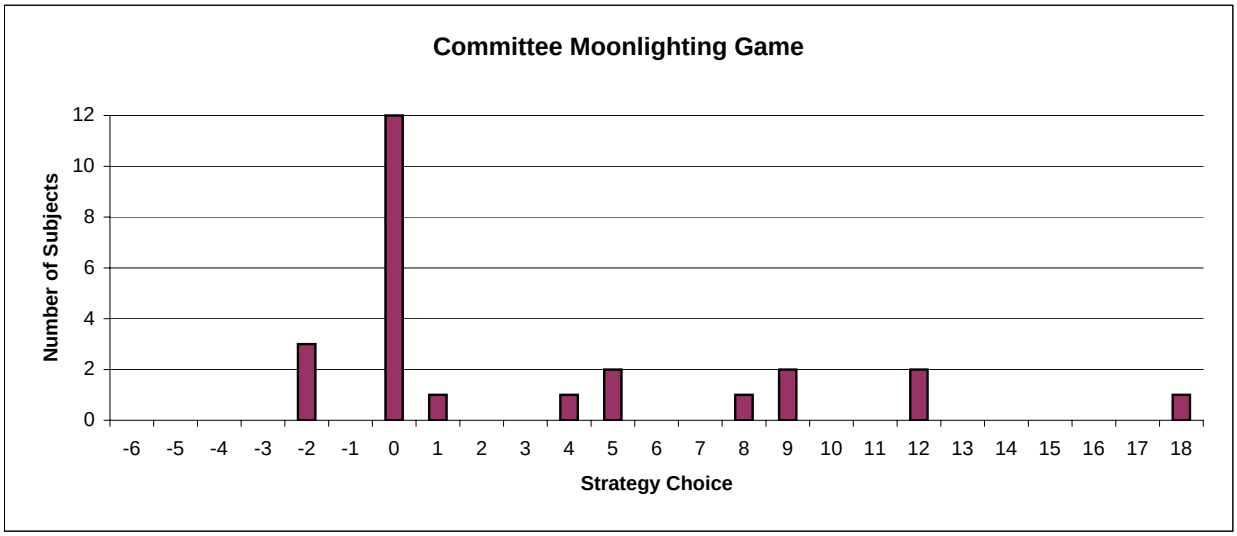


Figure 4: Histogram of Average Committee Decisions

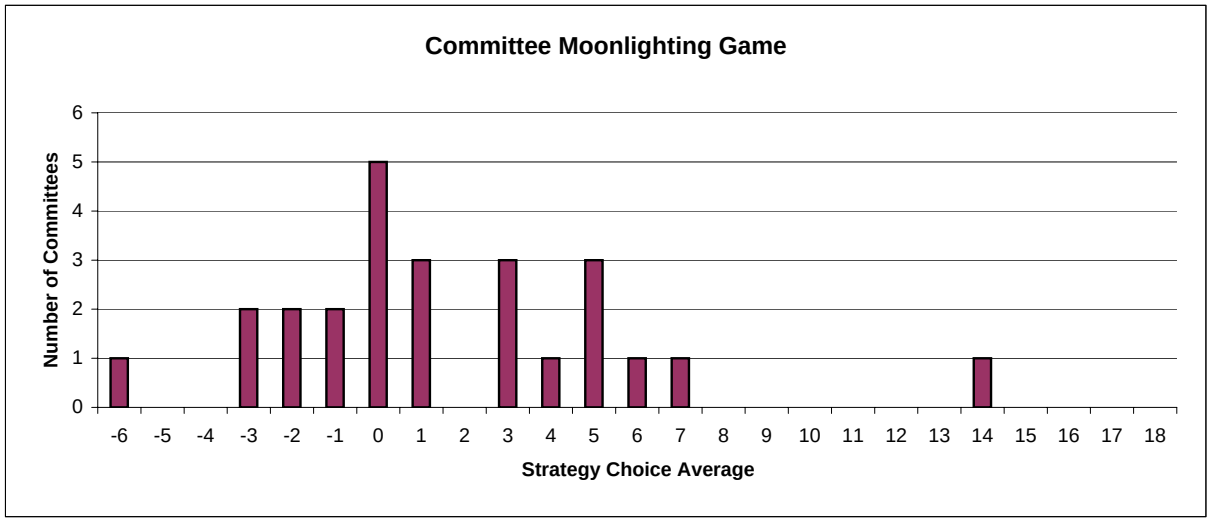
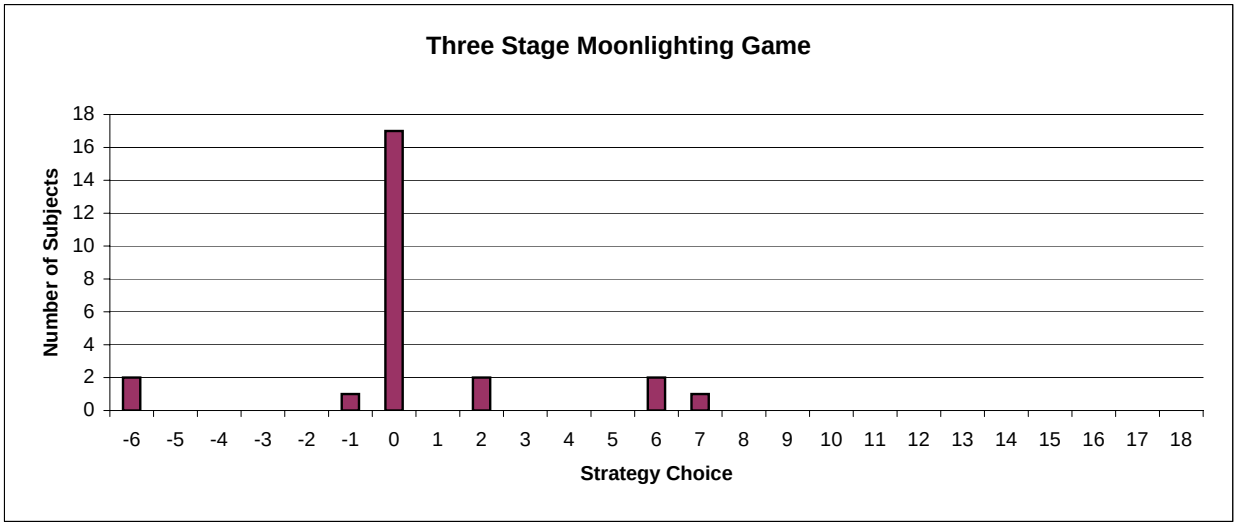


Figure 5: Histogram of A-player Second Decisions



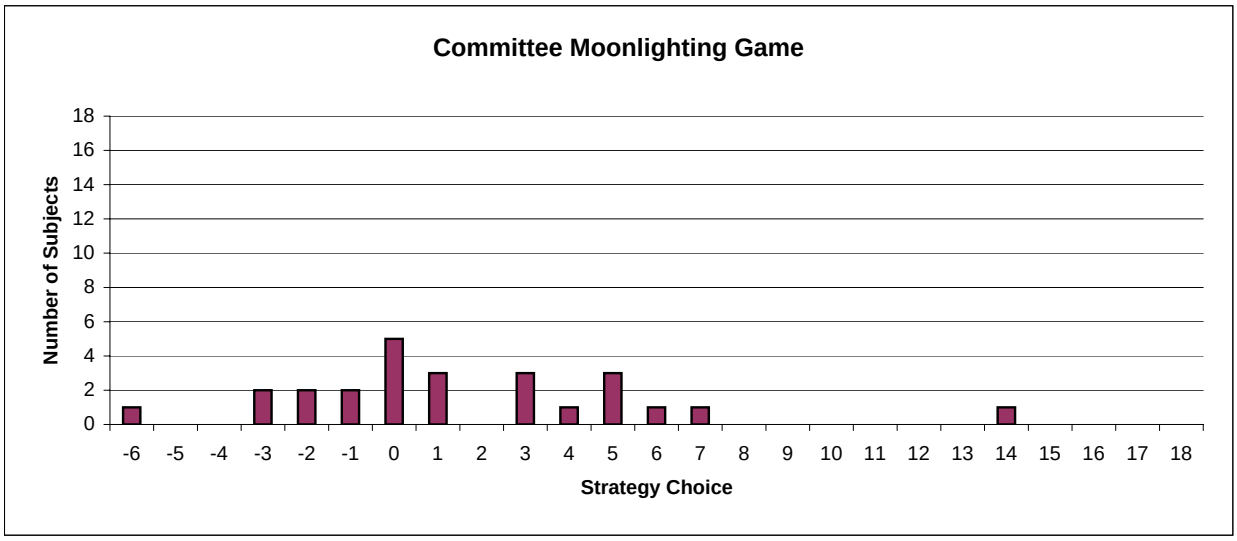


Figure 6: Histogram of B-player Second Decisions

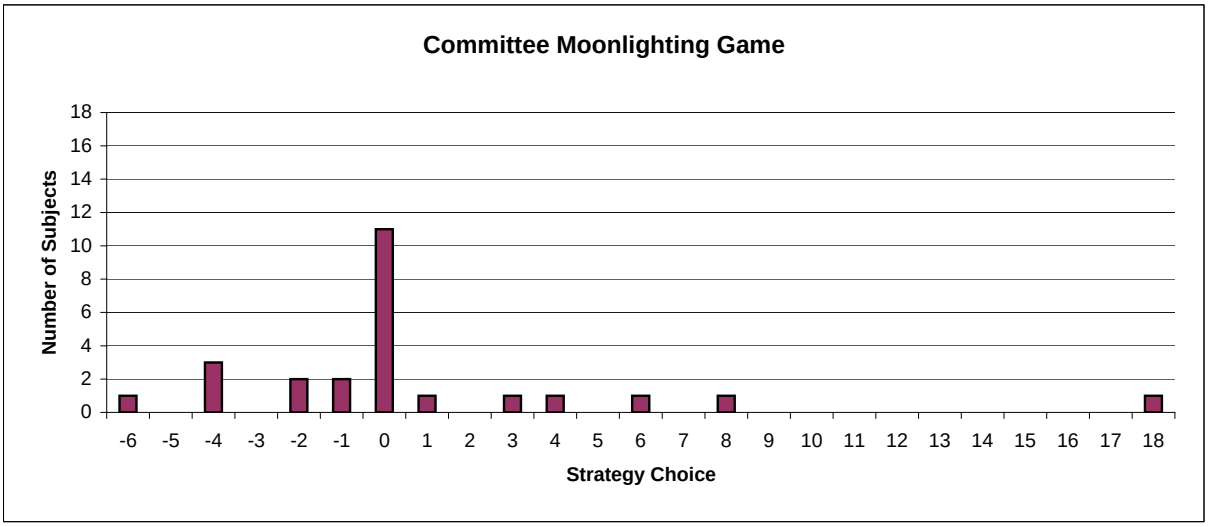


Figure 7: Decision Pairs in the Three Games