LUISS Guido Carli

LIBERA UNIVERSITA' INTERNAZIONALE DEGLI STUDI SOCIALI

Department of Business and Management

Chair of Managerial Decision Making

Differences in the decision-making process between Individuals and Groups

SUPERVISOR

Prof. Luigi Marengo

CANDIDATE

Lorenzo Simoni

ID Number: 697451

CO-SUPERVISOR

Prof. Luca Panaccione

ACADEMIC YEAR 2018/2019

Table of Contents

INTRODUCTION

In this thesis I will analyze the differences in terms of behavior between individuals and groups when it comes to decision making. To do so I will use Game theory concepts and axioms, since with game-theoretical decisions there is generally no space for individual interpretation of the right answer, it is either right or wrong, making the analytic process the most impartial possible. The first part then will be focused on explaining and assembling the Game-theory postulates, such as the definition of rationality and the explanation of utility functions, in order to have a full understanding of the treatments in the following chapters.

The areas of analysis will be on differences in:

- Rational behavior;
- Reciprocal behavior;
- Independence in decision making;
- Outcomes of decision making.

Moreover, at the beginning of each demonstration there will be a brief introduction to summarize and show the previous experiments and conclusions already made by the academy in the same or similar fields. For what concerns the groups it has to be said that the academy has started to investigate their behavior just recently and that’s the reason why many things that might seem obvious had to be tested in order to make the right comparisons and get data for the comparison.

The experiments done and proposed are three, the first is a revised centipede, the second is a classic ultimatum game (those will be really relevant to show differences in rationality and reciprocal behaviors), and the third one is a guessing game, where participants have to guess the right amount of beans in a jar; this will show the differences in terms of independency (intended as the capability to stick to own ideas) in decision making and the accuracy of those decisions.

All the results show with a certain consistency that the groups are less influenceable and more precise than individuals. The explanation for this may be the fact that the groups have an intrinsic mechanism that smooth the components ideas and then it makes the final decision more accurate, whereas the individuals have no one to compare their ideas with.

However, making decisions in groups has its own downsides. In fact, it will be shown that, even if the outcome of the groups is consistently bigger that the individual's outcome, the overall welfare is decreased.

CHAPTER 1

1.1 PREFERENCES

To better understand why people make certain decisions we first need to understand what are the criteria that they use to undertake them. We can say that these criteria are based upon certain preferences.

To give a mathematical explanation of preferences, let's suppose that X is a set of possible finite actions or object where the selection of one excludes the selection of the others. If we imagine that those objects are *a, b,* and *c*, and individual can express preferences among these three, ranking them with a certain order, that order is the representation of preferences.

From a mathematical point of view these preferences can be represented though an inequality relation between the three objects taken as an example *a, b* and c, with the only variation that the symbol used will not be the classic inequality one "≥" but another specific that represents the preferences "≽".

We can further divide preferences into "strict preferences", "indifference" and "weak preferences". Strict preferences happen when an individual will always choose a certain element or a certain action to another one, for example if a is strictly preferred to b then it will be represented as $a \succ b$. Indifference represents the situation when for an individual there's no distinction between two choices, for example if a is indifferent to b then mathematically it will be $a \sim b$. For what concerns weak preferences, a certain object is weakly preferred to another if for the individual is indifferent or preferred a certain object to another, for example if a is weakly preferred to b then $a \succcurlyeq b$.

At the basis of preferences there are two axioms that are needed otherwise we can't have preference relations, these are:

- Transitivity. This axiom is necessary in order to rank choices in a preference scale, from the best to the worst and the indifferent ones. From a mathematical point of view, it is represented as $\forall a \in$ X (where $\forall$ means for "each one"), $\forall b \in$ X, and $\forall c \in$ X if $a \succcurlyeq b$ and $b \succcurlyeq c$ then $a \succcurlyeq c$, i.e. $(a, b) \in \succcurlyeq$ and $(b, c) \in \succcurlyeq \Rightarrow (a, c) \in \succcurlyeq$.

- Completeness. this axiom regards the certainty that an individual can always choose whether he preferers one choice to another. Its mathematical writing is $\forall a \in$ X and $\forall b \in$ X either $a \succcurlyeq b$ or $b \succcurlyeq a$, i.e. $(a, b) \in \succcurlyeq$ or $(b, a) \in \succcurlyeq$.

What stands underneath preferences is the utility function, which varies from individual to individual and will be discussed further in the next paragraph.

1.2 UTILITY FUNCTIONS

In economics we can evaluate preferences over a given set of choices thanks to the utility functions. The utility functions assigns a number also known as expected utility, or expected satisfaction, for a consumer from a given consumption or a certain choice, and formally is represented as U : X → R, or U ⊆ X × R, where R is a set of real numbers. A utility function U represents the preference relation $\succcurlyeq$

$$\forall a \in X, \forall b \in X, U(a) \geq U(b) \Leftrightarrow a \succcurlyeq b$$

Even though utility functions and preference relations can look quote alike, the utility function can actually give us more information, for example if we compare the consumptions of two goods (*a* and *b*) an individual can prefer to consume a certain good over the other and this will be shown in both the relations, but the utility function can help us also quantify how much consuming a certain amount of a good can be preferred (or indifferent) to the consumption of a certain amount of the other good.
At this point it becomes useful to introduce the two types of utility that reflect this difference:

- Cardinal Utility. This is a function that represents the intensity of the difference between U (*a*) and U (*b*). For example, we can say that for a certain individual the consumption of a certain good a is preferred three times as much that the consumption of a good b, and this will be formalized as $U(a) - U(b) = 3[U(a) - U(b)]$.

- Ordinal Utility. This is a function that reflects the preferences of an individual in a ranked (ordinal) scale. This means that there will be a qualitative evaluation rather than a quantitative one in comparing the utility deriving from two or more goods, which substantiate in the avoidance of the actual numbers behind the utility. In comparing $U(a)$ and $U(b)$ the only thing that matters is whether $U(a) \gtreqless U(b)$. A real example for the use of a cardinal utility function is when an individual has to decide whether to go take a certain flavor of ice-cream instead of another, there is no numerical calculation in utility behind that to say how much he prefers one over the other.

As for the preference functions we have to bear in mind that each individual has his own utility function which varies from the other individuals' ones, even though when studying a group of people we can treat it like a single entity with its own preference and utility functions that derive from the combination of the preferences and the utility functions of each member of the group.

1.3 GAME THEORY

Sometimes when we have to interact with others it may not be really easy to make decisions, especially in the economic field.

To obviate this problem, we generally try to simplify reality and here is where Game Theory comes into play. Game Theory studies mathematical models of interaction between individuals providing techniques or creating models in order to yield a better result in the decision making given a certain situation, called also game.

These "games", based upon mathematics, are created taking into account the desired variables, the number of actors involved in the decision process and the relative outcomes that can be simplified in numbers, generally called "payoffs". The payoffs can be a representation of anything, starting from a certain amount of utility to arrive to an amount of money, and are influenced by the number of variables accounted and most importantly by preferences; of course the more the variables the more representative of the reality the model will be, making it more difficult to evaluate and to find a solution.

The aim of the actors involved in the game is to make the best amongst the possible decisions, also referred as strategies, to maximize their final payoffs. This because one of the main assumptions of Game Theory is that individuals participating to a game are rational and so they play to maximize the outcome, but we will deepen this aspect in the next section.

To give a real example, let's suppose that there are two criminals that committed a crime and are taken by the police; the police officers don't know who committed the crime, but they have got enough evidences to charge them with minor offences, so they take the two prisoners into two separate rooms to interrogate them. To simplify let's suppose that each criminal has two possible actions, confess or deny the crime. If both deny the crime, they will be charged with minor offences and spend two years in jail each; if they both confess, they will both spend just one year in jail since the terms will be reduced due to their cooperation with the authorities, but if one confesses and the other denies the confessor will be given immunity, whereas the other will spend three years in jail.

		CRIMINAL 2	
		CONFESS	DENY
CRIMINAL 1	CONFESS	1,1	0,3
	DENY	3,0	2,2

The graph above represents mathematically the assumptions just proposed, also known as the "prisoners' dilemma". Given the preferences, the possibilities and the linked payoffs, each one of the criminals will try to maximize their payoffs, which translates into a problem of minimization of the jail years. For the criminal 1 the best outcome results when he cooperates and the other criminal denies, the second better outcome is when they both cooperate, the third is when they both deny the charges, and the worst outcome is when the criminal denies and the other cooperate with the authorities; for the criminal 2 the preferred outcomes are the same specular ones.

Since each prisoner will act to maximize his self-interest, both will tend to deny, ending up with two years of jail each.

We can further divide the games into two types, the “noncooperative games” and the “cooperative games”. The main difference between the two lies in the fact that while for the noncooperative games there will be a winner and a loser in terms of outcomes and there is no possibility to interact in order to find a final decision (such as in the proper sense of the word Game), in the cooperative games players can accord and negotiate to have a final desired outcomes (such as when two or more agents agree to make a deal and write down a contract). This distinction is really important and can lead to very different outcomes in most of the games, and to explain that let’s consider again the “prisoners’ dilemma”.

We can say that the prisoners’ dilemma is a noncooperative game since there’s no possibility for the two criminals to have an agreement on a strategy to follow and, as we saw, the equilibrium is reached when they both ended up with two years each of jail time. We can also transform this game into a cooperative one, we only need to let them mediate before taking their strategy; in this new prisoners’ dilemma the final outcome will be that both confess the crime and get one year each of jail time, reducing it of one year in respect to the noncooperative game.

At this point becomes crucial to introduce two fundamental principles at the basis of Game Theory known as Rationality and Common Knowledge, and a third concept that will be useful for the next chapters of this thesis, Prosociality.

1.4 RATIONALITY

Given that we don’t act in a vacuum, taking into account also others’ preferences become vital when we try to maximize our payoffs, since others’ preferences and actions may have effects on our final outcomes like in the above-mentioned prisoners’ dilemma example.

This is why the main assumption underneath all these models regarding Game Theory is that the players are rational. This means that each player will try to maximize his payoff.

To be more precise, we know that each players’ action or strategy will lead him to different outcomes, that will have the player strive to undertake the action or strategy that will result in the best expected payoff.

Rationality doesn’t necessarily mean that one player will seek to maximize monetary goals, but that he will be and will act coherently with his preferences and utility functions, which may be regarding anything. In fact, since rationality has its fundamentals on preferences and utility, it must respect all their axioms and rules such as cardinal and ordinal utility, making coherence in choices one of the most important rules to be satisfied; from a mathematical point of view that means rationality should always be complete and transitive for any choice ($\succcurlyeq_1, ..., \succcurlyeq_N$).

1.5 COMMON KNOWLEDGE

When we set up a game it is important to define what are the information each player has. Generally, it is required that each player knows not only about the possible strategies of game *per se* but also the information set the other players have, their utility function and the rules of the game. Moreover, an important information that each player is required to know is that each player playing the game is rational. All these requirements in terms of information are better known as Common Knowledge, which is required in a "game".

A particular fact F is said to be common knowledge between the players if each player knows F, each player knows that the others know F, each player knows that every other player knows that each player knows F, and so on (Watson, 2013).

It is important to say that Common Knowledge of the game doesn't necessarily imply that there won't be asymmetry in information.

1.6 BOUNDED RATIONALITY

Unfortunately, when it comes to reality there may be some cases where the actors deviate from the optimal strategy, like for example when there is the presence of "biases", or the calculations required to find out the best strategy are too humongous to be completed in the required time. For these reasons in the second half of the twentieth century the academy, especially in the figure of Herbert A. Simon, started to focus on the concept of "Bounded Rationality".

According to the theory of bounded rationality, during the decision-making process there are many factors that may influence or limit the rationality (in the game theory sense), those factors could be for example a scarce or limited information set, amount of time given to make a choice and evident limits given by the difficulty of the calculation required to find out the optimal

solution. In this framework an agent will not strive to find out the optimal solution, but he will try to find the most satisfying. The new model proposed by Simon is composed then by some exemplifications which are supposed to reflect more precisely the reality.
The first simplification proposed is about the Pay-off function. According to Simon we should consider only two possible combination of outcomes

- The two possible outcomes are (0,1); we consider the choice "unsatisfactory" when the outcome is equal to 0, whereas when the outcome is 1 the action is "satisfactory" for the individual;

- There are three possible outcomes (-1,0,1); depending on the circumstances we can see these outcomes as lose, tie, win.

An actual example for the first preposition given by Simon is when someone sells a house. The seller may individuate as a "satisfactory" price €100.000, hence each offer above this value should be accepted. Giving this, there may be several differences between satisfactory offers received, since for example an offer of €110.000 is more satisfactory than a €105.000 one albeit both acceptable, while from a "pure rationality" point of view the first offer is better than the second one.
In this view then, Simon sees the decision-making process as a continuous process where the level of satisfaction can change and be reassessed based upon "satisfactory" and "unsatisfactory" alternatives that come up during the process itself.
The objection to this idea may be answered in several different answer corresponding to a class of situations in which reasoning with this approach might be appropriate:

- First, the individual may not be confronted simultaneously with a number of buyers offering to purchase the house at different prices, but may receive a sequence of offers, and may have to decide to accept or reject each one before he receives the next. (Simon, 1955)

- Second, even if there was a more general pay-off function, W(*s*), capable of assuming more than two different values, the simplified outcome V(*s*) might be a satisfactory

> approximation to W(*s*). Suppose, for example, that there was some way of introducing a cardinal function, defined over the set S, say U(s). Suppose further that U(W) monotonic increasing function with a strongly negative second derivative (decreasing marginal utility). Then V(s) = V{W(s)} might be the approximation. (Simon, 1955)

A further step taken by Simon has been to introduce the information-gathering process. In fact, during real life, the decision-making process with the right set of information plays one of the most important roles to find the best options. According to Simon the information-gathering process must be costless, otherwise it will have an important impact of the decision-making process itself. To see why let's suppose that an agent has to buy a car and he is sure on the model he wants to buy. Each time he checks on websites or goes to a car retailer he will spend time and/or money to move, this will create a trade-off that will have to be considered in the decision making-process, and "then one element in the decision will be the determination of how far the mapping is to be refined" (Simon, 1955). When the agent has found a car with the price considered "satisfactory", according to the bounded rationality theory, he will be sure that there are no better options by acquiring additional information, then should stop to gather and buy that car.

1.7 RISK & BELIEFS

Another important factor that we have to consider when it comes to a decision-making process is the risk. Actors try to make the best decisions within the information set and the expectations on consequences that they have got. Since those expectations may sometimes reveal to be not accurate, we can define this uncertainty behind the results and outcomes of a certain choice as "Risk". In fact, "Risk" is sometimes used as a label or the residual variance in a theory of rational choice (March, 1994).

For what concerns the bounded rationality theory, its supporters are more interested in understanding what is the process that makes individuals or organization take risks.

When it comes to risks it is important to introduce the concept of beliefs and annexed probabilities, since those are the way we can refine and be more accurate with the outcomes of the various strategies. When the strategy a certain player will follow is not that clear it is vital to have beliefs, since that will help the other player finding out the best strategy. To be clearer let's recall again for example the example of the prisoners' dilemma game.

		CRIMINAL 2	
		CONFESS	DENY
CRIMINAL 1	CONFESS	1,1	0,3
	DENY	3,0	2,2

Criminal 1 can have some beliefs of the type "I am sure that Criminal 2 will Confess" or "I know that Criminal 2 will never Confess" where, in this case, the probability (*p*) correlated to each of the two events is equal to 1, and this will simplify the most possible the decision-making process. In most of the cases unfortunately, players will be likely to have ambiguous beliefs in the sense that the probability they correlate to a certain strategy played by the opponent variates from 0 to 1. The implications on the game are that a belief reduces the expected outcome by the probability related to it, so for example an outcome of 10 with p=1/2 will result in a new expected outcome of 5.

The importance of beliefs and assumptions can be seen also if we observe the changes in the decision-making proposed by James March. According to March (A Primer on Decision Making: How Decisions Happen, 1994), the modern decision-making process involves changes in the pure theory; those changes can be formalized in four points:

- Knowledge. What is assumed about the information decision makers have about the state of the world and about other actors?

- Actors. What is assumed about the number of decision makers?

- Preferences. What is assumed about the preferences by which consequences (and therefore alternatives) are evaluated?

- Decision rule. What is assumed to be the decision rule by which decision makers choose an alternative?

CHAPTER 2

2.1 EVIDENCE FROM LITTERATURE

Recently the academy has started to deepen the analysis on group thinking and the differences between group behavior and individual behavior. The main differences highlighted concern rationality and prosociality. Whereas the group behavior tends to be more rational with less space for a prosocial behavior, for individuals prosociality is more accentuated.

The first example of differences in terms of rational behavior between individuals and group has been documented by G. Bornstein, T. Kugler, and A. Ziegelmeyerb in (2004). To investigate this difference they had set up a game where the two competitors had to decide whether to accept an amount of money, which will increase each turn, or pass the other competitor that had to do the same (this type of games is better known also as centipede games); the end of the game will be reached when either a competitor decide to take the money or no one accepts at the end of the sixth turn, as shown in Fig. 1. It had been also assumed that each competitor would play in order to maximize his profit and that this is common knowledge.

This game has been submitted and played both by 2 individuals and by groups composed by 2 or 3 persons.

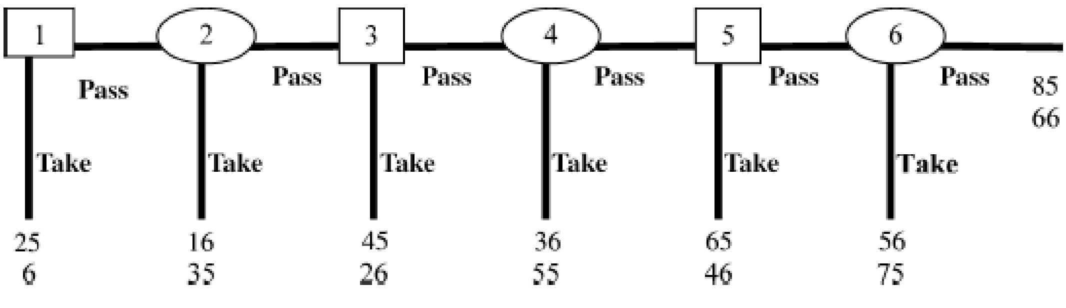


Fig. 1. The increasing-sum Centipede game. Player 1's decision nodes are denoted by squares, and Player 2's by circles. The upper payoff at each terminal node is the payoff of Player 1.

From a Game Theory point of view the best solution for this game can be easily found through a backward induction; since there will be a moment when a competitor will accept the money or pass ending up with a forced choice, but most importantly with a lowered outcome for the passing player, the best response to that for the other player would have been to have taken the money the turn before; this process will continue until reaching the first turn. This is why, rationally speaking, the best option for this game will always be to bail out and accept the money at the first turn available.

The sample for this study was composed by 144 undergraduate students at the Hebrew University of Jerusalem with no previous experience with the task. The participants were recruited by campus advertisements offering monetary rewards for participating in a decision task. Thirty-six subjects participated in the individual condition (in 6 cohorts of 6 participants) and 108 subjects participated in the group condition (in 6 cohorts of 18 participants) (Bornstein, Kugler, & Ziegelmeyerb, 2004). The game, shown in fig. 1, was played once and the payoffs are expressed in New Israel Shekels (where 1$ is equal to 4.5 NIS).

For what concerns the individuals' condition, they have been informed about the rules of the game, randomly divided into pairs without knowing who their opponent was and assured about the confidentiality of the final result. To further avoid other biases, the decision in the game has been made on a computer in separate rooms, and the final payment has been done outside the laboratory in a sealed envelope.

Regarding the groups' condition, each participant of each cohort has been randomly allocated into a group, that was seated in separated rooms from the other groups in order to have a free and private discussion. As for the individuals, each group didn't know the composition of the other groups and who they were competing against, since the decisions were received and taken from a computer. There was no specific instruction on how to make a decision about the game. The payoffs shown in Fig. 1 represents the payoffs for each member of the group.

Node	1	2	3	4	5	6	7
Individuals	0	0	1	1	12	1	3
Groups	0	0	3	7	5	3	0

Table 1

The result of the experiment, both for the individuals and the groups are shown in Table 1. As we can see the average terminal node at which the game ends for the individuals in 5.22, while the average terminal node at which the game ends for the groups is 4.44. The difference between the two distributions is statistically significant by a robust rank-order test ($\grave{U} = 2.36$, $p < .01$) (Bornstein, Kugler, & Ziegelmeyerb, 2004).

The authors gave a plausible solution for these results in the lack of prosociality of groups, also motivating this intuition saying that, after asking the players of the experiment, for the individuals the more the game was going on the more they felt like they had to reciprocate (as we can also see from the 3 individuals arrived at the seventh node). Moreover, since games lasted longer for individuals rather than for groups, the average join payoff for the individuals was higher.

Given that this experiment could not exclude that the plausible explanation for the final outcome of the games was that groups make less mistake than individuals, instead of being less prosocial, it has been conducted another experiment, where this time passing will not increase the payoff of the opponent, as shown in fig. 2.

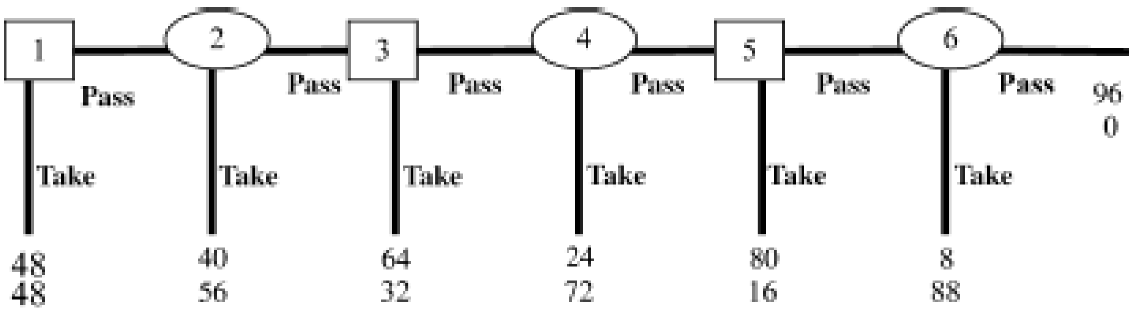


Fig. 2. Player 1's decision nodes are denoted by squares, and Player 2's by circles. The upper payoff at each terminal node is the payoff of Player 1.

The numerosity and the selection modalities have been the same of the first experiment applied to new people. The results have changed drastically from the first experiment, as shown in Table 2.

Node	1	2	3	4	5	6	7
Individuals	3	6	5	4	0	0	0
Groups	4	11	2	1	0	0	0

Table 2

As we can see the average terminal node for the individual dropped to 2.56, while for groups it drastically reduced to 2.00; this difference between the two is statistically significant by a robust rank-order test (Ù = 1.7, $p < .05$).

These new findings can let us conclude that groups tend to make fewer mistakes than individuals and, since the constant-sum game has been concluded earlier than the increasing-sum game, we can say that both groups and individuals act in a pro-social way, even though this prosociality is attenuated when decisions are made by groups.

Further studies have been conducted to demonstrate this positive difference between individuals and groups. An example of it is the paper written by Gary Charness and Matthias Sutter in (2012), where the authors have reported three important lessons:

- Groups are More Cognitively Sophisticated than individuals;

- Groups Can Help with Self-Control and Productivity Problems;

- Groups May Decrease Welfare Because of Stronger Self-interested Preferences.

To demonstrate the first point there have taken previous evidences from studies already conduced, as for example the "beauty-contest game". The beauty-contest game is a simultaneous move game where a certain number *n* of participants has to choose an integer number each from the interval [0,100], and the winner is the one who has chosen the number closer to the average multiplied by *p*, where *p* is a fraction less than 1 given before the game. Since *p* is not an integer number and less than 1, the rational strategy to play is to choose 0. To explain why let's suppose that $p = ½$, the average for a random choice from the interval [0,100] is 50, that would mean that if someone anticipates the random guessers the best strategy for him would be to choose 25; on the other hand if someone anticipates this strategy his best response would be to play either 12 or 13, and this would continue until we reach the equilibrium choice of zero. Through this game it has been demonstrated several times that groups choose each time lower numbers than individuals (as shown in Figure 3), thus suggesting that they are reasoning deeper about the strategy of the game and are expecting the other parties to reason deeper as well (Charness & Sutter, 2012).

Moreover, it has been found also that groups think one step ahead of individuals, as shown by the median in figure 3.

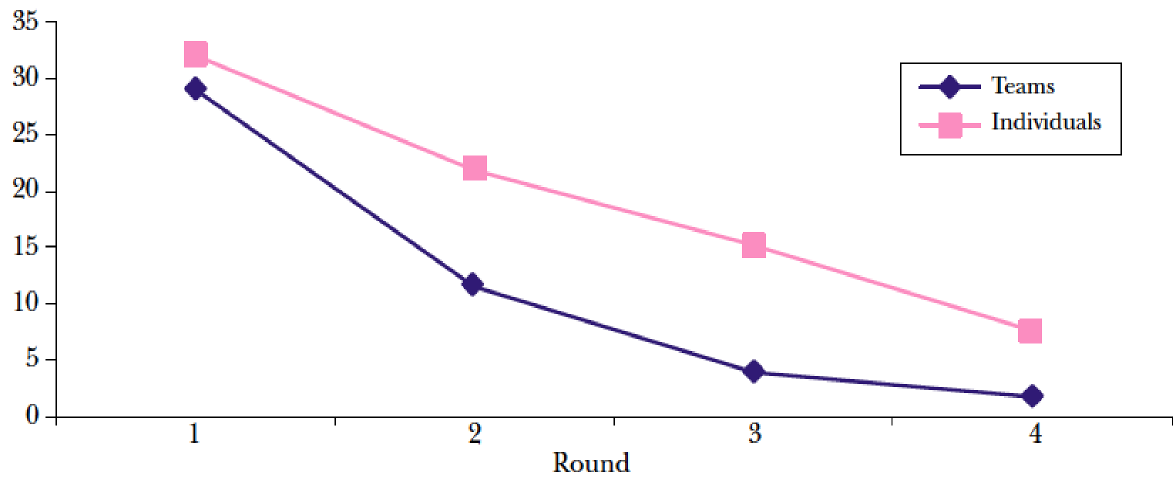


Figure 3; Source: (Kocher & Sutter, 2005).

The second point regarding self-control and productivity problems is quite more difficult to observe in a laboratory, since there are not many games that can be set up and not being biased. The solution to obviate this problem has been given by Falk and Ichino (2006), thanks to a field experiment. They had different subject perform a real-effort task such as putting mails into an envelope from a mass mailing. The experiment has been divided into two parts, during the first one there were individuals alone in rooms performing this task, and in the second part there were groups of two doing the same. The results have shown that groups were outperforming the individuals by 16% average in terms of productivity, meaning that being in a group had a positive impact.

The third point of this analysis made by Gary Charness and Matthias Sutter (2012) addresses the possibility for group-based decisions to lessen social welfare. The first game presented to verify this assumption is better known as "The trust game". This game has been first proposed in 1995 and consist of the presence of two players acting over multiple stages; the first player has to decide an amount x (where $x \geq 0$) and send it to the second player, before passing to the second player this amount will be tripled, the second player will have to decide how much of the amount $3x$ to pass back to the first player, when the second player decides how much to pass back the game ends. Rationally speaking since the first player knows the second player is not obliged to give back any amount, the best strategy is to have $x = 0$, which will not change the amount in the total welfare. As shown by the running of this experiment by Kugler, Bornstein, Kocher, and Sutter (2007) groups sent less amount of money when they were first-movers by a 20% compared to the individual players, and return less on average also as second-movers.

The table below reassumes the amounts sent by the senders of "The trust game" just described. The value in parenthesis is the standard deviation. Source: (Kugler, Bornstein, Kocher, & Sutter, 2007)

Sender	Responder	
	I(ndividual)	G(roup)
I(ndividual)	65.5 (36.4) $N = 32$	76.3 (31.2) $N = 25$
G(roup)	54.0 (41.6) $N = 25$	43.7 (42.4) $N = 27$

The table below reassumes the amounts sent back by the receivers of "The trust game" just described. The value in parenthesis is the standard deviation. Source: (Kugler, Bornstein, Kocher, & Sutter, 2007)

Sender	Respondent	
	I	G
I	25.1 (19.5)	25.1 (17.5)
G	23.3 (22.1)	16.7 (18.7)

Cases with $x = 0$ are excluded.

The implication of this experiment is that group-based decisions are more detrimental for the welfare rather than the individual-based ones. More support of this proposition comes from the battle-of-sex game and the prisoners' dilemma game where has been shown (Charness, Rigotti, & Rustichini, 2007) how group-based decisions lead to a lower level of welfare, thus concluding that groups may decrease welfare because of stronger self-interested preferences.

Another element on which we can further evaluate the differences between individuals and groups is how they make decision facing risks. This element is relevant when perhaps it comes to analysts in financial markets; knowing whether to work individually or set up workgroups may be relevant for the final outcome.

To evaluate the difference between individuals and groups when facing risk, we need a discriminant. Gary Charness, Edi Karni and Dan Levin (2006) in their study to explore these differences used the violation of the first-order stochastic dominance. There is a first-order stochastic dominance for a random variable X over another variable Y for each outcome a when X gives at least a as high probability as Y of obtaining a, and for some a X gives a higher probability than Y, formally:

$$P[X \geq a] \geq P[Y \geq a] \text{ for } \forall a, \textit{ and for some } a \; P[X \geq a] > P[Y \geq a]$$

When we receive a new information that should be included, and subsequently update our reasoning given that it has an impact on the final outcomes. However, if that information is not included on purpose, underestimated or overestimated then the subject has a *bias*.

The experimental study made by Charness et al. (2006) has been divided into two stages, one stage where the participants acquired the information, and a second stage where they have been required to choose between two lotteries. This division in two stages lets individuate the eventual *bias* due to the perception of the information acquired as "bad news" or "good news", violating the first-order stochastic dominance.

The two lotteries are two urns containing 6 black balls and 6 white balls each, and both urns are an equal-likely representation of the world, up and down. The back balls are positive payoff, since when you pick up a black ball you get paid, otherwise nothing changes if you pick up a white ball.

	Left Urn	Right Urn
Up ($p = .5$)	●●●●○○	●●●●●●
Down ($p = .5$)	●●○○○○	○○○○○○

As we can see in the "up" state the left urn has 4 black balls and the right one has 6, while in the "down" state the left urn has 2 while the right urn has none.

The design proposed by Charness et al. (2006) is composed by four treatment. The first treatment has been called "ABCD", where "A" stands for the presence of an emotional responses (*Affect*), "B" represents a Bayesian Updating needed (Update the probability of a certain event based upon prior knowledge), "C" is the presence of the Compounding effect, and "D" represents the first-order stochastic dominance. Players are required to pick up sequentially two balls from any of the two urns, and they will be paid depending on how many black balls they get. In this stage the order of the balls in the urns in terms of "up" and "down" will not change. To ensure information from both lotteries, the first 20 participants had the first-pick fixed. If the player picks a black ball from the Left urn, for first-order stochastic dominance that player should draw the second ball from the Right urn; vice versa, if the decision-maker picks up a white ball from the Left urn he should stick with that urn also for the second draw. The affect here can be isolated by seeing the emotional response to the first-round pick, if he goes against the Bayesian and first-order stochastic dominance there we know that the decision-maker is biased.

The second treatment, named “BCD” (as for the “ABCD” but without the *affect*), was created to limit this emotional influence on decisions. In this experiment the payoff linked to the draw of a ball were told at the end of the first draft, so the players could not be influenced by the pick.

In the third treatment, called “CD”, Charness, Karni and Levin tried to eliminate also the need of a Bayesian updating. To do so the number of draws has been reduced to one, informing the player that the possibility of the state “Up” for both urns is p =2/3. This will create the same probability of picking up a black ball at the first turn, but the subjects should not be influenced by the pick.

In the fourth and last treatment was focalized to evaluate the *Dominance* factor. For this new experiment the urn-lotteries have been suppressed and substituted by a dice roll. The treatment still requires the players to choose between two lotteries. In the first lottery it will be rolled a nine-faced dice and when a number from 1 to 5 shows up then the player awards $2, in the second lottery the $2 will be awarded for any number showing up from 1 to 6. These probabilities are equal to the ones in the first treatment (“ABCD”) in the case of a winning draw from the left urn.

There have been four treatments to evaluate and test the assumption that all these elements influence the decision-making process under risk.

To further evaluate the differences between individuals and groups the treatments “BCD”, “CD” and “D” have been repeated for two-person and the treatments “BCD” and “D” for three-person groups. The groups were free to decide independently and interacting with each other member of the same group.

It has been paid $0.30 for each successful draw in the “ABCD”, “BDC” and “CD” treatments, the difference in the number of periods between individuals and groups is due to time restriction and groups took more time than solo individuals.

Treatment summary – Individuals

Treatment	1st draw restriction	Payment	Participants	Periods
ABCD	In first 20 periods	Black balls from both draws	57	60
BCD	Left draw only	2nd draw only, color TBD	57	80
CD	None	Black balls	48	80
D	None	Black balls	104	1

Treatment summary – Groups

Treatment	1st draw restriction	Payment	Participants	Periods
BCD - 2 person	Left draw only	2nd draw only, color TBD	30	30
BCD - 3 person	Left draw only	2nd draw only, color TBD	21	30
CD - 2 person	None	Black balls	36	30
D - 3 person	None	Black ball	66	1
D - 2 person	None	Black ball	21	1

Source: (Charness, Karni, & Levin, 2006).

For what concerns the results regarding the first-order stochastic dominance, it has been observed that the bigger the group, the smaller the error as we can see from the Table of results just right below.

Treatment	Following success	Following Failure	Aggregate error %
BCD - 1	18.8% (420/2234)	39.9% (908/2276)	29.4% (1328/4510)
BCD - 2	15.4% (67/434)	32.6% (152/466)	24.3% (219/900)
BCD - 3	7.5% (23/307)	10.8% (35/323)	9.2% (58/630)
CD - 1	30.2% (1158/3840)	-	-
CD - 2	23.0% (248/1080)	-	-
D - 1	8.7% (9/104)	-	-
D - 2	3.0% (2/66)	-	-
D - 3	0.0% (0/21)	-	-

	Coefficient	Std. Error	*t*-statistic	P > |*t*|
CD	.1054	.0266	2.96	0.003
D	-.1013	.0412	-3.50	0.001
Pair	-.0562	.0356	-2.11	0.035
Trio	-.1024	.0289	-2.49	0.013
Constant	.1895	.0260	7.30	0.000

Number of obs. = 383; Adjusted R^2 = 0.1210

We see that in relation to the BCD baseline, the error rate in the CD treatment is significantly higher and the error rate in the D treatment is significantly lower. The error rate with larger groups (compared to the individual baseline) is significantly lower, with the negative coefficient for the trio dummy substantially higher than the coefficient for the pair dummy (Charness, Karni, & Levin, 2006).

Since the sample has been randomly selected is safe to state that subjects make better decisions, or also make more frequently first-order stochastic dominating strategies, when consulting with others. Moreover, the presence of affect, the need to apply Bayesian updating and the complexity of the alternatives tend to contribute to erroneous judgments and poor decisions (Charness, Karni, & Levin, 2006).

To further explore the decrease in welfare when groups make decisions, the effects of intergroup competition on group coordination on behavior in the minimal-effort game have been studied by G. Bornstein, U. Gneezy and R. Nagel (2002). It is needed to stress and better analyze the welfare outcome in the case of groups decision making because, if the aim of the decision process may have prosocial intentions for example, it is important to know which way is the better one. The first step of the analysis is to go over the experiment of Van Huyck et al. (1990). Their experiment studies the behavior in a minimal-effort game, a game where there are many Pareto-ranked equilibria. The rules of the game are really easy, each player has to choose simultaneously with the others an integer number from 1 to 7, and the player's payoff depends on that player's choice as well as the minimal number chosen by any of the other players in the group (Bornstein, Gneezy, & Nagel, 2002); the game has been played by a group of 14 to 16 players. The parameters of the game are so that players have an incentive to choose a high-minimum and players choosing a number higher than the group's minimum will be penalized. This game has seven equilibria, all of them reached when all players chose the same integer number. However, these equilibria can lead to very different levels of welfare and from a game-theory point of view players are incentivized to play the one with the lowest outcome.

The representation of payoffs of this game is in the table below

	Smallest number chosen by the participants in your group						
Your decision	7	6	5	4	3	2	1
7	130	110	90	70	50	30	10
6	–	120	100	80	60	40	20
5	–	–	110	90	70	50	30
4	–	–	–	100	80	60	40
3	–	–	–	–	90	70	50
2	–	–	–	–	–	80	60
1	–	–	–	–	–	–	70

Van Huyck et al. (1990) showed in their experiment that, when the game is played multiple times, and the outcomes are disclosed after playing each round, players converge on choosing the integer 1. Moreover, this game has been replicated many times and similar results have been showed also with the case of a two-person games by Cooper et al. (1990). These findings, however, are related to a single-group setting and this is why in their experiment Bornstein et al. (2002) reformulated the Van Huyck et al.'s (1990) game. In this reformulation there are two groups, A and B, each composed by *n* players. Each member of the groups independently chose an integer $e_{i \in A(B)}$ from 1 to 7. If the minimum chosen in team A, $\min_A = \min\{e_{i \in A}\}$, was larger than the one chosen in team B, $\min_B = \min\{e_{i \in B}\}$, team A won the competition and each of its members was paid according to the original payoff matrix. Members of team B were paid nothing. In case of a tie, $\min_A = \min_B$, each player in both teams was paid half the payoff in the original matrix (Bornstein, Gneezy, & Nagel, 2002).

The set of strict equilibria doesn't change within each group and, as in the previous game, the best response for each member is to play the minimum integer chosen in their group.

The participants of this game have been 210 undergraduate students, with no previous experience with the game, split in cohorts of 14.

At the beginning of the experiment, the subjects were randomly divided into two equal-sized groups, and then played 10 rounds of the game. The number of rounds to be played has been revealed in advance. In each round each subject had to choose an integer from 1 to 7. Following the completion of the round, each subject received feedback concerning: (a) the lowest number chosen by the members of her group in that round; (b) her earnings (in pesetas) in that round;

and (c) her cumulative earnings. In two of the treatments (as detailed below), subjects also received feedback concerning the lowest number chosen in the outgroup. Each session lasted about 30 min with an average payment of 1109 pesetas (approximately US $9) per subject (including the fixed show-up fee of 500 pesetas). Subjects were paid in private. Each 14-person cohort took part in one of the three following treatments (Bornstein, Gneezy, & Nagel, 2002).
To evaluate all the different behaves, there has been set up three different little variation of the game:

- Intergroup competition (*IC*). The group with the higher minimum won the competition and, as in the Huyck et al.'s (1990) game, the members of that group were paid according to the payoff matrix, whereas members of the losing groups received nothing. In case of a draw, the groups were paid the half of the original matrix. Six 14-person groups took this treatment and have been informed only of the minimum chosen in their group and the one they were competing against;

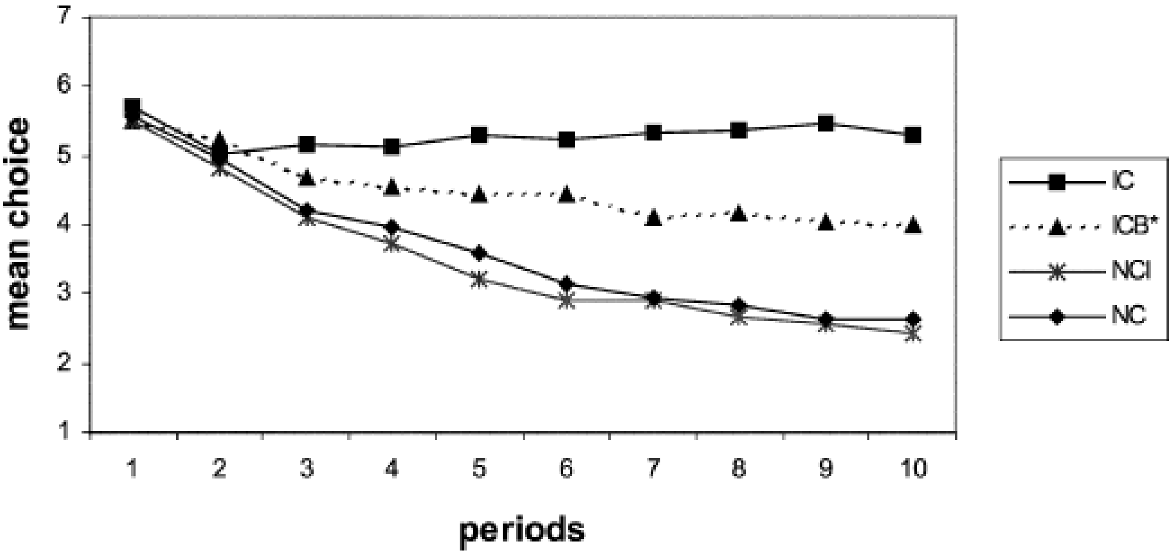
- No competition (*NC*). In this treatment, as in the Huyck et al.'s (1990) experiment, players were competing against each other and not against another group. This treatment has been made to have a control group. Players were informed of the minimum number chosen in their own group and three 14-person groups took this treatment;

- No competition with information (*NCI*). This treatment is pretty much the same as the one above, there is no competition between groups, the competition is within the same group, and there is no interdependence for the outcomes. However, the members of each group were informed about the minimum of the outgroup in addition to their own group's one. Six 14-person groups participated in this treatment.

The mean effort level was 5.3 in the *IC* treatment, 3.6 in the *NC* treatment, and 3.5 in the *NCI* treatment. The difference between the *IC* and the *NC* treatment is statistically significant by a Wilcoxon rank test ($z = 2.830$, $p < 0.0047$), whereas the difference between the *NC* and *NCI* control treatments is not significant (Bornstein, Gneezy, & Nagel, 2002).
The mean effort levels in Period 1 amongst the groups has not been significantly different from one another, and those have been 5.7, 5.6, and 5.5 for respectively treatments *IC*, *NC*, and *NCI*.

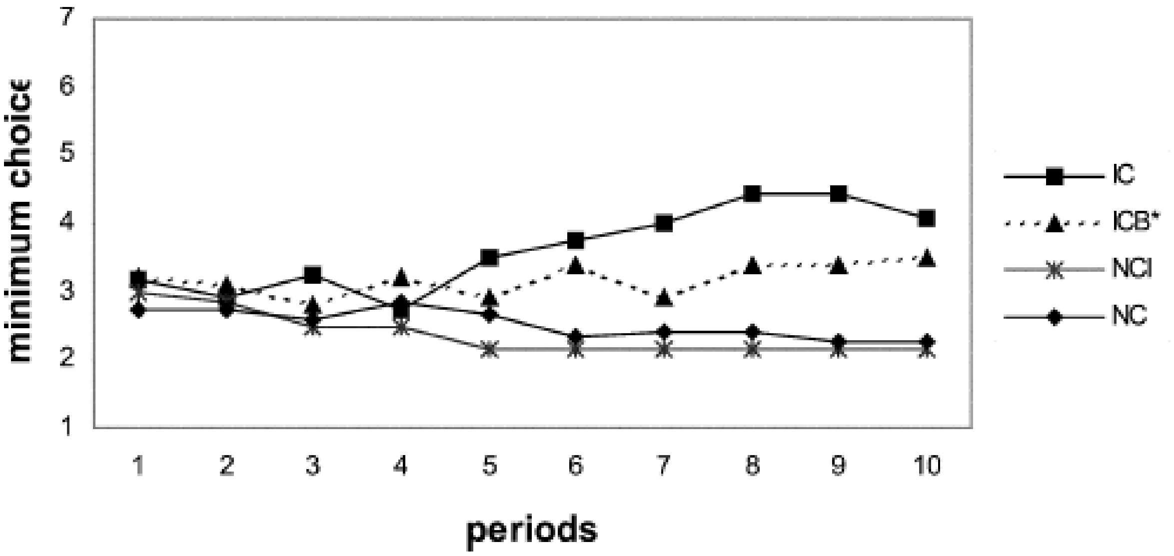
However, after the first round the subjects in the *NC* and *NCI* control treatments began to reduce their overall effort level in respect to those in the *IC* group, whereas the effort level in the *IC* group remained quite stable during the periods.

The table below shows the results for each treatment group regarding convergence in terms of outcomes. Source: (Bornstein, Gneezy, & Nagel, 2002)



With this data we can observe that the convergence to 1 is present when there is an intergroup competition. For what concerns the minimum effort levels the treatments showed that for the first half of the periods the levels have been quite similar across the three treatments. However, from the period 5, there was an increase in the average minimum in the IC treatment, as compared with a slight decrease in the minimum choice in the other two treatments. The mean minimums in Period 5 are 3.5, 2.2, and 2.3 in treatments IC, NC, and NCI, respectively. The differences between the IC experimental treatment and the NC and NCI control treatments are statistically significant ($z = -1.93$, $p < 0.02$; $z=-1.55$, $p <0.06$, respectively). The mean minimum in the last period was 4.1 in treatment IC as compared with 2.2 and 2.25 in treatments NC and NCI, respectively. The differences between the IC treatment and the NC and NCI treatments are statistically significant ($z = -0.203$, $p < 0.02$, and $z = -1.85$, $p <0.03$, respectively). The difference between the two control treatments is not significant (Bornstein, Gneezy, & Nagel, 2002).

The table below shows the results for each treatment group regarding minimum effort levels. Source: (Bornstein, Gneezy, & Nagel, 2002)



These results reflect in a certain way the findings about the convergence in terms of outcomes. In fact, since the outcome for the *IC* treatment is significantly higher than the outcomes for *NCI* and *NC* treatments, we expect the minimum level of effort to be significantly higher as well. The consistency of the minimum effort level for the treatments *NCI* and *NC* can be explained if a player assumes that the behavior of the other components of the group remains constant, in fact, if the effort level remains constant, the best response in the period $t + 1$, is to behave in the same way as the period t. This can be applied also for the treatment *IC*. However, in the treatment *IC*, after a loss, players can choose a higher effort level without decreasing their outcome, whereas this doesn't apply to the other two treatments.

The table below shows the changes of minimum effort for each treatment (underlined numbers indicate highest frequency of change within a specific case). Source: (Bornstein, Gneezy, & Nagel, 2002)

Min ingroup >, =, < Min outgroup	Change of min from previous period	IC	NCI	NC
>	Decrease	0.23	0.10	
>	Unchanged	0.54	0.88	n.a.
>	Increase	0.23	0.02	
=	Decrease	0.20	0.15	
=	Unchanged	0.53	0.85	n.a.
=	Increase	0.27	0.00	
<	Decrease	0.23	0.15	
<	Unchanged	0.26	0.66	n.a.
<	Increase	0.51	0.20	
Overall	Decrease	0.22	0.13	0.15
Overall	Unchanged	0.44	0.78	0.79
Overall	Increase	0.34	0.08	0.06

After these findings, Bornstein et al. opted for another treatment in order to separate the effect of intergroup competition from the security consideration that players may have playing a number close to 1 (When a player plays the number 1, he is sure to win, even if the sum is quite smaller if compared to the potential maximum win). This treatment is called "Intergroup competition for a bonus" (*ICB*). Here the members of both teams were paid according to the original game matrix regardless of the outcome of the competition and, in addition, the members of the winning group (i.e., the one with the higher minimum) were paid a (fixed) bonus (or half of it for a tie) (Bornstein, Gneezy, & Nagel, 2002).

The overall effort level for this treatment across the 10 periods was 4.5, which is not significantly different by a Wilcoxson rank-test from that observed in the *IC* treatment (z = -0.82, $p < 0.138$), but is significantly higher than the effort level in the *NCI* treatment (z = -1.10, $p < 0.06$). As shown by the Table of results above, the mean effort level in Period 1 for this treatment has been 5.5, whereas in the last period the mean has been 4. This difference is not significantly different from the one in the treatment *IC* (z = -1.10, $p < 0.138$), but significantly higher than the last-period mean in the *NCI* treatment (z = 1.55, $p < 0.06$).

For what concerned the minimal effort levels, in the *ICB* treatment we can see a change from 3.2 in the first period to a 3.5 mean in the last period, this is however not statistically different from the results observed in the *NCI* treatment.

What is curious to observe is that following a win or tie in the *ICB* treatment, the minimum remained the same in 81% of the cases, decreased in about 12% of the cases, and increased in the remaining 7% of the cases; whereas after a loss the minimum increased in 41% of the cases (Bornstein, Gneezy, & Nagel, 2002), showing that the participants in the *ICB* treatment acted really similar to those in the *NCI* treatment in case of a victory, and similar to the ones in the *IC* treatment in case of a loss.

The group performance as we have just seen, tends not to exploit all the potentialities and drop the group production. This loss in production is generally attributed to two factors:

- Free riding; this happens when members of a group are not held responsible for the outcomes singularly, but as a group, so components may not be incentivized to perform as they should or as they would as they would have been alone. In fact, in case of victory or rewards, all the group is rewarded independently from the singular effort;

- Coordination loss; this happens when the members of the group do not perform in a coordinate way, resulting in a decrease in the group outcome.

As this study shows, intergroup competition enhances productivity by improving intragroup coordination (Bornstein, Gneezy, & Nagel, 2002).

The comparison with the individuals cannot be made for what concerns free riding and coordination loss since it would not make any sense. However, these data can be useful to better investigate the differences between groups and individuals when reciprocity takes place.

To better understand when groups will be more or less accurate than individuals, we can report the study and the algorithm proposed by A. Koriat (2012). In his algorithm, called maximum-confidence slating (*MCS*), for each trial of the participants grouped in virtual couples, the decision that is made with higher confidence by one member of a virtual dyad is selected, circumventing dyadic interaction altogether (Koriat, 2012). According to the self-consistency model (*SCM*) of subjective confidence, if we let the participants interact with each other and combine their knowledge before making a decision, the maximum-confidence slating algorithm is supposed to generate a two-heads-better-than-one effect. However, the self-consistency model presupposes some conditions under which one head is better than two.

This model (*SCM*) can help understand when confidence judgements are indication of accuracy and the reason why. In many two-alternative forced-choice (2AFC) tasks, a relatively high within-person confidence-accuracy (C/A) correlation is typically observed across items, suggesting that participants can monitor the accuracy of their choices (Koriat, 2012).

In his experiment then the subjects were required to perform in a visual task. To get data for both individuals and groups participants, after being split into dyads, participants were required to give the answer individually first and then they could discuss with their partner and change their opinion if they felt like they wanted to. In each dyad the member with higher percent of correct answers has been labeled as high performing (*HP*), whereas the others have been referred as low performing (*LP*). Moreover, also two dummies have been introduced using the responses, the dummy high-confidence (*D-HC*), using the answers of the high performing players, and the dummy low-confidence (*D-LC*), using the answers of the other players. The treatment has been composed by five studies:

- The study number 1 was based on visual tasks where participants performed the experiment on their own, autocratically rating also their confidence in each answer in a 50-to100% scale. The same pairs of images have been presented to all participants in the same random order. The participants of the experiment were 38 and have been paired based upon their percentage of correct answers in order to compose 19 dyads. The results for this stage showed that accuracy was higher for *D-HC* than for *D-LC* [t (18) = 7.52, P < 0.0001], indicating a positive confidence-accuracy correlation in a between-individual comparison. Second, performance was more accurate for *D-HC* than for *HP* [t (18) = 6.69, P < 0.0001], supporting the two-heads-better-than-one effect. The superiority of *D-HC* performance was observed for 18 out of the 19 dyads (P < 0.0001) by a binomial test. Last, percent correct was significantly lower for *D-LC* than for *LP* [t (18) = 6.69, P < 0.0001] as showed in the summary table below (Koriat, 2012).

	HP (%)	*LP* (%)	*D-HC* (%)	*D-LC* (%)
		Study 1		
Oddball target	67.82	66.98	69.88	64.93

The same game has been played also to see whether there would have been differences between dyads and triplet, showing that the triplets performed better than the dyads;

- The second study was quite similar to the study number 1, just with the variation of using questions on general-knowledge on European countries (such as which one had larger area or more population) questions instead of visual questions. The results indicated that performance was more accurate for *D-HC* than for *HP* for both the area task and the population task; for both tasks, the results also indicated that three heads were better than two (Koriat, 2012);

- The third study has been based on a reanalysis of the data collected from the first two studies that tested the prediction of the self-consistency model. The focus has been on differences between *D-HC* and *HP* for *CC* and *CW* items. The findings showed that for the *CC* items, performance was better for *D-HC* (85.87%) than for *HP* (82.48%) [t (38) = 2.82, $P < 0.01$]; in contrast, for the *CW* items percent correct was lower for *D-HC* (19.87%) than for *HP* (26.60%) [t (38) = 2.86, $P < 0.01$]; for the *CW* items, *D-HC* performance was worse than that of the worst participant (*LP*) (25.16%) [t (38) = 2.20, $P < 0.05$]; for these items, the best accuracy was achieved by *D-LC* (31.89%), so that a two-heads-better-than-one effect can be obtained if the responses of the participant with lower confidence are selected (Koriat, 2012). The results of the third study are summarized in the table below.

		HP (%)	*LP* (%)	*D-HC* (%)	*D-LC* (%)
		Study 3			
Lines	CC	81.58	80.59	85.03	77.14
	CW	25.00	26.31	17.10	34.21
Shapes	CC	83.33	84.58	86.67	81.25
	CW	28.13	24.06	22.50	29.69

- The fourth study, as the third one, was based on a reanalysis which included *CC* and *CW* two-alternative forced-choice general-knowledge items. The *D-HC* has performed both the best performance for the *CC* items and the worst for the *CW* ones. For an external observer who cannot tell CW from CC items (cannot tell whether the consensual or high-confidence answer is right or wrong), applying the same heuristic across the board (for example, “take the high-confidence response”), or relying on dyadic decisions, can generally be beneficial. However, if the “crowd” is in error, reliance on confidence is bound to be misleading (Koriat, 2012);

- The fifth and last study has been used to extend the maximum-confidence slating algorithm also to a within-individual design. The tasks undertaken are the same of the study three, but here the participants performed them twice with an interval of one week. The hypothesis tested is that a compilation of the high-confidence choices across the two presentations should yield the same pattern of results as that observed for a between-person compilation; for the *CC* items, *D-HC* accuracy (82.54%) was higher than average performance accuracy (81.24%) [t (49) = 2.67, $P < 0.05$]. For the *CW* items, *D-HC* accuracy (24.06%) tended to be somewhat lower than average performance accuracy (25.00%) [t (49) = 1.05, $P < 0.31$]. In addition, across the *CC* items, confidence was higher when the correct choice was made than when the wrong choice was made, whereas the opposite was true for the *CW* items. (Koriat, 2012)

The results of the experiments showed that when correctness and consensuality between the two partners were separated due to the introduction of a load of different items to lure participants into the wrong answer, “confidence was correlated with the consensuality of the answer rather than its correctness (Koriat, 2012)”.

In the case of consensually correct (*CC*) items, the confidence-accuracy correlation has been showed to have a positive value, meanwhile it has been showed to be negative in the case of consensually wrong (*CW*) items.

The main implications of the two-heads-better-than-one are:

- In many situations the knowledge that is shared by all participants corresponds by and large to the truth, so that the maximum-confidence slating algorithm as well as social interaction are expected to yield decisions that are more accurate than those of each individual alone. Thus, maximum-confidence slating is expected to yield a two-heads-better-than-one effect for many perceptual and general-knowledge tasks in which the items are representative of their domain. Indeed, the wisdom-of-crowds phenomenon suggests that information that is aggregated across participants is generally closer to the truth than is the information provided by each individual participant (Koriat, 2012);

- If confidence is tuned to the “common knowledge” rather than to the truth, reliance on confidence can be misleading when the shared knowledge departs from the truth (Koriat, 2012).

It is peculiar to see how the benefits of the maximum-confidence slating in the case of consensually correct items are more beneficial for the cross-person slating (the one observed in the study three) than for the within-person confidence-based as observed in the fifth study. According to the author there are three main takeaways for this paper:

- The first one is regarding the two-heads-better-than-one, according to Koriat this effect should be observed when participants’ decisions are corrected by and large. This finding is coherent with all the evidence from the literature brought so far;

- In situations in which most participants tend to make the wrong decisions, the maximum-confidence slating algorithm, as well as social interaction, is expected to yield group decisions that are even less accurate than those of each individual alone (Koriat, 2012). This takeaway will have evidence also in the experiment of the beans-in-a-jar that will be proposed in the next chapters.

- The within-individual results (study number five) highlight a general perspective for the analysis of decision accuracy that goes beyond the effects of social interaction; This perspective, as captured by self-consistency model, involves the variations in confidence that occur both within individuals and between individuals when choice and confidence are based on the sampling of clues from a common database (Koriat, 2012)

When making decisions there are several factors that people have to take into account, and even if we always (or at least generally) we want to make the best decision that suits us, sometimes we can see actors imitate other actors. An explanation for this phenomenon is given by Marengo et al. (2016).

In their research they have studied the peer effects on investment choice manipulating three main dimensions of choice: time pressure, normative content of social imitation, and uncertainty of the investment. What is important to say is that in this study the social information given to the participants has been interpreted as non-instrumental: that is, a piece of knowledge about others' behavior that does not convey any clear informational advantage to a perfectly rational decision maker, who should simply maximize her expected utility given preferences and endowments, or in other words the rational updating of beliefs about probabilities are excluded (Marengo, Alexia, & Ploner, 2016).

The main reason to do so is that otherwise it is not possible to distingue their perspective from the traditional economic view of imitation as a form of rational social learning (Banjeree, 1993). In fact, as seen before, the acknowledgment of others' behavior can be perceived as an informative and then result in a Bayesian update of priors. Moreover, a boundedly-rational decision maker considers others' actions as relevant and may abandon private knowledge in favor of signals gathered from others (Marengo, Alexia, & Ploner, 2016). In the *ecologically-rational* approach of Gigerenzer, Todd, and ABC (1999), such a behavior simplifies the decision process obviating the need for further search and computation, especially when external (e.g. complexity) and internal (e.g. cognitive load) constraints are tight. The participants of the experiment have been asked to make investment decisions in a series of prospects, some with positive and other with negative expected value (Marengo, Alexia, & Ploner, 2016).

Moreover, the participants are provided with anonymous information about other participants' choices without the possibility to interact with each other, and the reward is based upon the individual decisions of each player.

There has been also manipulation for three dimensions:

- Uncertainty of the task at hand. Players were facing both risk lotteries where outcomes and probabilities are known, and uncertain lotteries, where probabilities of outcomes are not known;

- External cognitive constraints. This dimension regards time constraints; the participants have been given either shorter or longer time periods to make their choice;

- Normative content of the social signal. participants are either made aware of the choice of a single individual in their reference group or of the average choice of all individuals in the same group (Marengo, Alexia, & Ploner, 2016).

In order to make their investments, participants are provided with a set of E tokens, and each player can choose how many tokens they want to invest in an asset that has no return with probability *P* and 2.5 times return with probability 1-*P*

$$P = \begin{cases} E - x & p_L \\ E - x - 2.5x & 1 - p_L \end{cases}$$

Since E has been set equal to 200, if any player decides to bet 0, then he would get for sure the 200 tokens. The alternatives of investment proposed has been six, each of that changing the value of P_L. Five of those six prospects had the value of P_L disclosed, and have assumed the values: 2/8 (25%), 3/8 (37.5%), 4/8 (50%), 5/8 (62.5%), 6/8 (75%); whereas for the sixth one the value of P_L hasn't been disclosed in order to observe behavior in case of uncertainty.

The table below shows the results of the experiment in case of risk investments. The column EMP express the expected marginal profit of each token invested in these prospects, the column EV represent the expected value, and the column SD the standard deviation. Source: (Marengo, Alexia, & Ploner, 2016).

P	p_L	EMP	x = 0		x = 200	
			EV	SD	EV	SD
#1	0.250	+0.875	200.000	0.000	375.000	216.506
#2	0.375	+0.562	200.000	0.000	312.500	242.061
#3	0.500	+0.250	200.000	0.000	250.000	250.000
#4	0.625	−0.062	200.000	0.000	187.500	242.061
#5	0.750	−0.375	200.000	0.000	125.000	216.506

As we can see, investments with $P_L < 0.6$ have a negative expected marginal profit per token invested.

The participants have been randomly split into two groups, one group called *Targets* that played the game at an earlier date, and the group of *Observers*, who took part later than the *Targets*.

Participants of the *Targets* group have been asked to play the six prospects above in a randomized order. Of the six prospects only one is selected at the end to determine the actual reward for the player.

The *Observes* had their experiment split into two phases. In the first phase, the participants took the same experiment as the *Targets* group, playing the actual same six prospects. In the second phase the participants of this group had to take again the experiment, however this time, more information about what has been played already by the *Targets* group have been disclosed. The rewards for *Observes* participants are composed by the sum of two prospects payed, one from the phase 1 and one from the phase 2.

For those in the Observers group, two main dimensions are experimentally manipulated in a between-subjects fashion:

- Time Pressure and Information. Concerning Time Pressure, participants are either exposed to a condition of high time pressure (HIGH), in which choices in each round must be taken within 20 seconds, or to a condition of low time pressure (LOW), where choices must be taken within 40 seconds (Marengo, Alexia, & Ploner, 2016);

- Concerning Information, participants are either informed about the choice of a single individual randomly chosen among those in the Targets group (IND) or are informed about the average choice of all those in the Targets group (AGG) (Marengo, Alexia, & Ploner, 2016).

The table below shows the treatment labels. Source: (Marengo, Alexia, & Ploner, 2016).

		Time pressure	
		Low	High
Information	*Individual*	*IND:LOW*	*IND:HIGH*
	Aggregate	*AGG:LOW*	*AGG:HIGH*

Since in theory, individuals should have well defined and stable preferences, the amount invested among the phases for the participants of the *Observers* group should not change. The three hypotheses tested are:

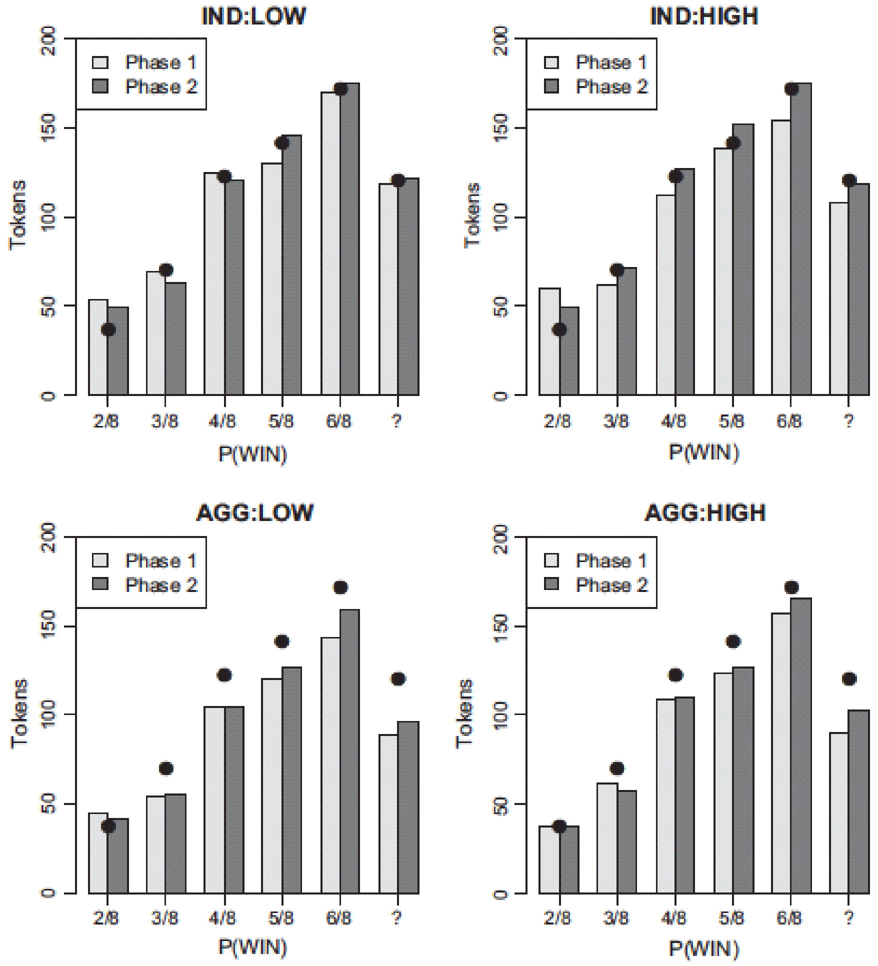
- Hypothesis 1. Stronger imitation of peer choices is observed in AGG than in IND;

- Hypothesis 2. Stronger imitation of peer choices is observed in HIGH than in LOW;

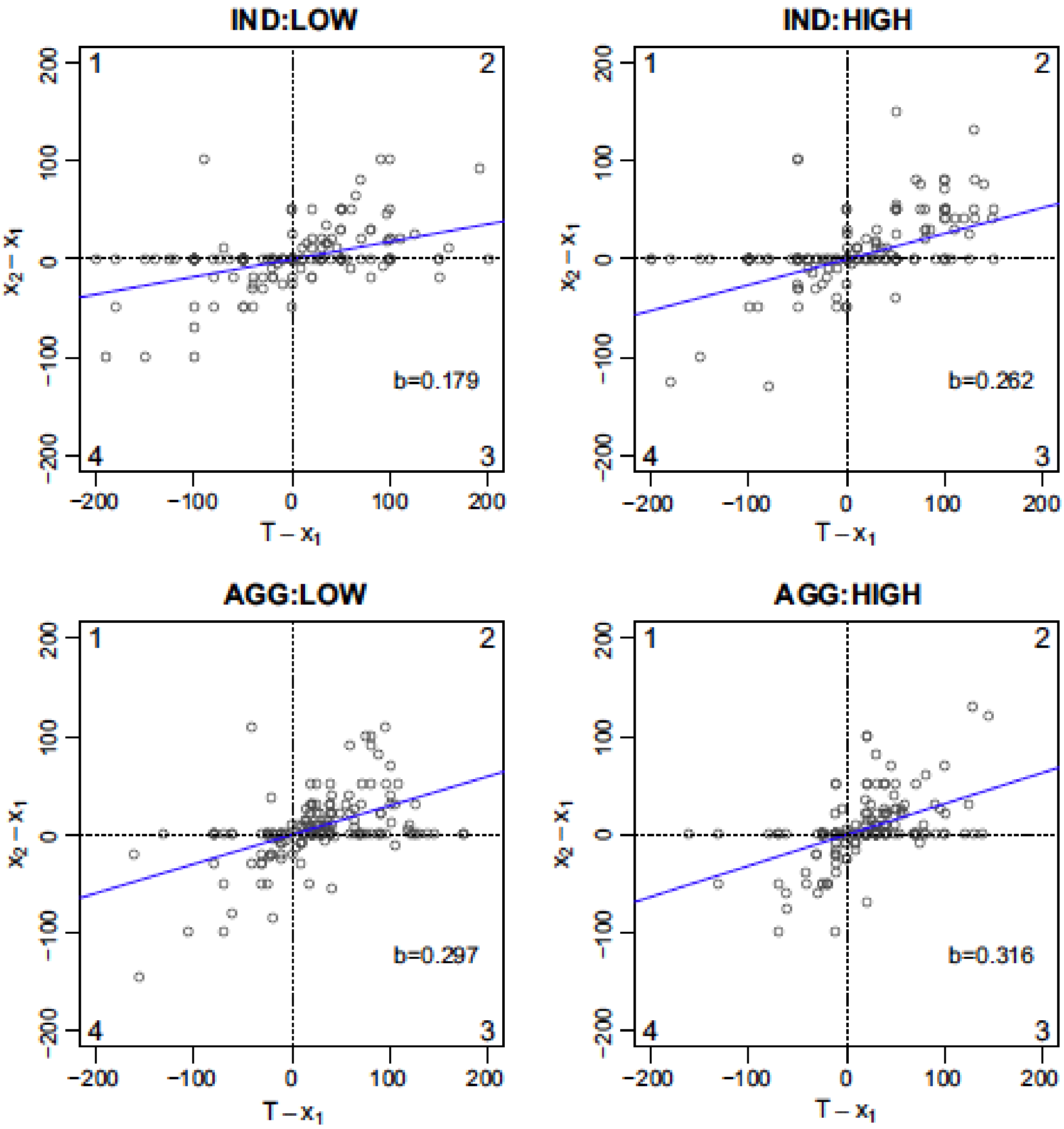
- Hypothesis 3. Stronger imitation of peer choices is observed for choices involving uncertain prospects than for choices involving risky prospects.

The results of the experiment show that participants tend to give more tokens to lotteries with higher expected values. On average, participants seem to grasp the features of the investment in terms of expected marginal profit (EMP). Concerning the uncertain prospect, it can be noticed that the median value of the support seems to represent a salient investment level when the situation is ambiguous (Marengo, Alexia, & Ploner, 2016).

The table below provides a representation of the average amount of token invested in the four experimental conditions. Source: (Marengo, Alexia, & Ploner, 2016).



For what concerns the peer effect, the results are summarized by the graph below. As we can see, choices are mainly located in quadrants 2 and 4, compatible with a tendency to imitate the target, and along the horizontal axis, revealing constancy of choices; the coefficients of the linear fit are positive in all four conditions, with the smallest value in IND:LOW and larger value in AGG:HIGH (Marengo, Alexia, & Ploner, 2016).



The difference between own choices in Phase 2 and in Phase 1 (x_2 - x_1) is reported on the vertical axis, while the difference between the Target and own choices in Phase 1 are reported on the horizontal axis (T - x_1). Source: (Marengo, Alexia, & Ploner, 2016).

The main results of the study are that:

- A large share of participants adopts peer choices as a reference for own choices, increasing (decreasing) their risk exposure when having a lower (higher) risk propensity than their peers (Marengo, Alexia, & Ploner, 2016);

- Larger differences between own choices and peer choices trigger larger imitative adjustments in risky investments. In the column 1 of the graph below, the coefficient of interaction term DIFF (it can be either given by the differential between the own choices in the two phases or the differential the *Targets* group choices and the choices during phase 1 of the *Observers* group) times UNC (a dummy that takes the value 0 when the lottery is risky and 1 when the lottery is uncertain) is negative and statistically significant. Thus, the positive impact of other choices observed for risky prospects is weakened by uncertainty (Marengo, Alexia, & Ploner, 2016);

- Imitative adjustments are weaker for uncertain prospects than for risky prospects. The interactions between our treatment dummies AGG and HIGH and the variable DIFF.*target* capture the impact of information and time pressure on choices, respectively. When taking into account the entire sample, only aggregate information fosters the impact of peer choices on own choices. However, when restricting attention to the sample of people revising their choices, both treatment variables foster the likelihood of imitating peer behavior (Marengo, Alexia, & Ploner, 2016);

- Among those who revise their choices, a more accurate description of average behavior in the reference group and less time to deliberate about the investment both foster imitative behavior in risky choices (Marengo, Alexia, & Ploner, 2016).

The table below represent the regression analysis of the study. In column 1 are shown all available, while in column 2 is composed by choices of those revising their investment from phase to phase. Source: (Marengo, Alexia, & Ploner, 2016).

DIFF.own	(1)	(2)
(Intercept)	−6.014 (9.725)	−47.998 (19.620)*
DIFF.target	0.220 (0.046)***	0.396 (0.069)***
ABOVE.target	−6.055 (3.280)°	−13.551 (5.938)*
UNC.lott	2.406 (3.060)	8.487 (9.603)
AGG	−3.191 (2.336)	−1.338 (4.253)
HIGH	3.955 (2.223)°	1.330 (3.925)
DIFF × UNC	−0.113 (0.051)*	−0.228 (0.120)°
DIFF × AGG	0.116 (0.052)*	0.178 (0.081)*
DIFF × HIGH	0.047 (0.046)	0.149 (0.068)*
Round	−0.482 (0.648)	0.794 (1.037)
Age	0.162 (0.407)	1.635 (0.841)°
Female	−7.005 (2.155)**	−9.685 (3.591)**
Tokens	0.022 (0.006)***	0.034 (0.012)**
R^2	0.284	0.540
Adj. R^2	0.271	0.522
Num. obs.	675	333
Sampling	All choices	$x_1 \neq x_2$

° $p < 0.1$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

As we have seen so far most of the individuals tend to be influenceable by external information. This is why in my experiment that will follow in the next chapters the resilience of groups will be tested to compare if they can be less influenceable than individuals.

2.2 STRONG RECIRPOCITY

With the term strong reciprocity, it is generally referred to a behavior that subjects may embrace. The peculiarity of this behavior is to act towards someone or something in a way to match (Reciprocate) the behavior those subjects have been receiving from that subject.

The essential feature of strong reciprocity is a willingness to sacrifice resources for rewarding fair and punishing unfair behavior even if this is costly and provides neither present nor future material rewards for the reciprocator (Fehr, Fischbacher, & Gächter, 2002).

In some cases, we can see that prosocial behavior and reciprocal behavior can actually overlap, but the main difference is that in a prosocial behavior the agent or the agents making decisions will have a certain altruistic behavior although the actions of the other players. In fact, a prosocial agent wants to increase the social outcome, whether it goes against him or not; meanwhile, a reciprocal agent can potentially also have a non-cooperative behavior if the other agent is not behaving according to his expectations.

To determine if an action is whether positive or negative, there are two dimensions that have to be considered:

- The potential consequences that the other actor's action will have on the subject;
- The intentions of the other actor.

For example, if an actor *x* behaves with the best intentions but the consequences turn out to be bad for the subject *y*, this subject *y* not necessarily will reciprocate.

To formalize it in a mathematical form (Falk & Fischbacher, 2000) we have to consider jointly kindness of the treatment of another person, which will be represented by the term φ, and the behavioral reaction to it, indicated with the symbol σ.

Since the outcome (Δ) can be defined as the consequences of an action, we can say that as long as $\Delta > 0$ the outcome is beneficial, while if $\Delta < 0$ it is detrimental for the individual. The intentions of the other actor are represented by the term ϑ, known as the intention factor; this variates from 0 to 1, where 1 is totally intentional and 0 is when no intentions are involved.

The combination of the outcome and the intention factor in a given node *n* yields the kindness term φ for that *n* node

$$\varphi(n) = \vartheta(n)\,\Delta(n)$$

After the kindness term, Falk and Fischbacher (2000) formalized the reciprocation term σ. In order to individuate it we first need to define $\nu(n, f)$ as the unique node that directly follows the node *n* on the path that leads to *f*, the end node. If *i* and *j* are the two players, and S_i and S_j the behave strategies of respectively *i* and *j*, we will have that

$$\sigma(n,f) = \pi_j\left(\nu(n,f), s_i'', s_i'\right) - \pi_j\left(n, s_i'', s_i'\right)$$

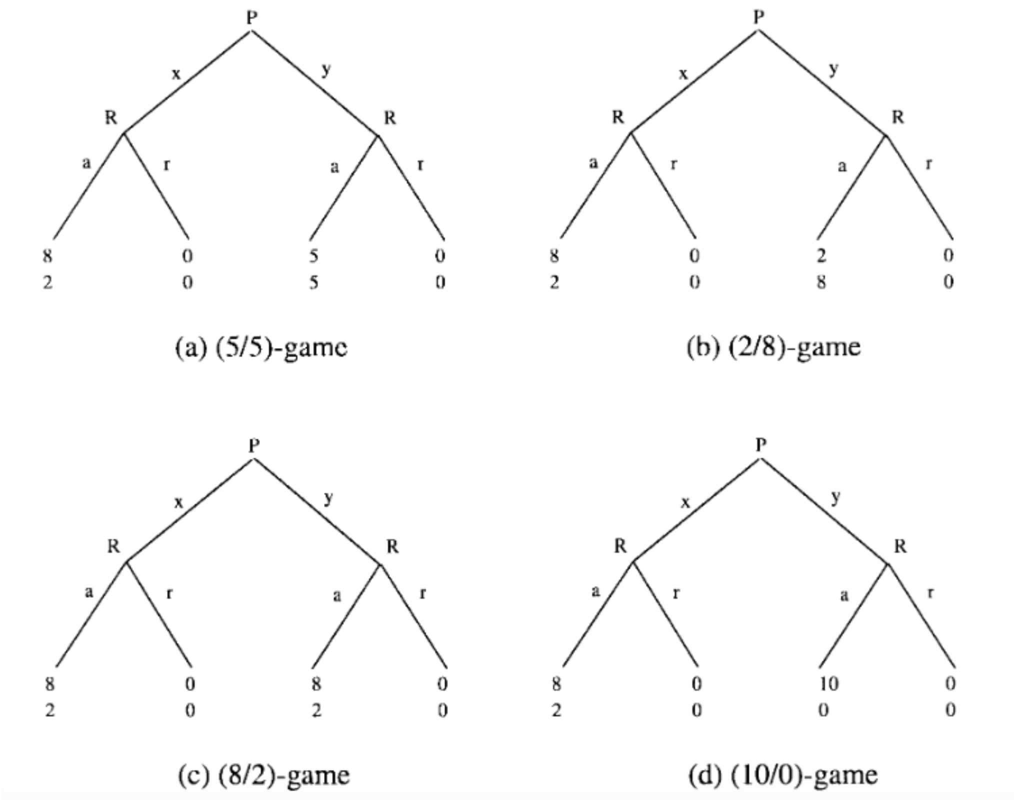
Where s_i' represents the *i*'s beliefs about the behavior strategy $s_j \in S_j$ that player *j* will adopt, better known also as first order belief. On the same false line, s_i'' represents the player *i*'s beliefs about the player *j*'s beliefs on what behavior strategy player *i* will adopt, and it is known as second order belief. Given this $\pi_j(\mathrm{n}, s_i'', s_i')$ will represent the expected outcome at the node *n* for player *j*, given the beliefs of player *i*; and $\pi_{\mathrm{j}}\ (\mathrm{v}\ (n, f), s_i'', s_i')$ will yield the expected outcome of player *j* at the following node, given the same information. The reciprocation term $\sigma(n, f)$ will then identify the reciprocal impact of changing the payoff from $\pi_j(\mathrm{n}, s_i'', s_i')$ to $\pi_{\mathrm{j}}\ (\mathrm{v}\ (n, f), s_i'', s_i')$, if this difference is positive then the player *i* will have a positive behavior with player *j*; otherwise, if this difference is negative, then player *i* is punishing the player *j*. According to Falk and Fischbacher (2000), the utility function is composed by two elements for a certain player *i* can be finally defined then as

$$U_i(f) = \pi_i(f) + \rho_i \sum_{\substack{n \to f \\ n \in N_i}} \varphi\ (n)\ \sigma\ (n, f)$$

The first element, $\pi_i(f)$, is the actual payoff that the player *i* will receive at the node *f*, while the second term, $\rho_i \sum_{\substack{n \to f \\ n \in N_i}} \varphi\ (n)\ \sigma\ (n, f)$, is defined reciprocity utility. In the reciprocity utility there is the introduction of the parameter ρ_i. This parameter is a constant that embeds the reciprocal preferences of player *i*. It variates from player to player, and the greater the more the certain player is inclined to reciprocate, given that the utility deriving from it will increase. In the case of $\rho_i = \rho_j = 0$ we have that the effect of reciprocity is none and the games will be equal to standard games.

2.3 EVIDENCE FROM LITERATURE ON RECIPROCITY

The first work on reciprocity that has to be reported is the one done by Fehr et al. (2003). In their work they demonstrated that the same offer in an ultimatum game can be rejected or accepted depending on the different strategies the proposer has, highlighting the presence of reciprocity. To do so they set up four different mini-ultimatum games (as shown in the figure below), with a sample of 90 subjects participating.



(a) (5/5)-game (b) (2/8)-game (c) (8/2)-game (d) (10/0)-game

In all the these four mini-ultimatum games the first-mover player (P) has two options *x* and *y*, while the respondent, the player R, has two options *a* (Accept) and *r* (Refuse). The upper numbers represent the outcomes of the player P, while the numbers below are the player R's payoffs. What is important to spot is that in the (c) (8/2)-game the player P has actually no choices, because either strategy have the same outcomes.

To add another level of analysis it has been asked to the responders, before playing the actual games, to theoretically answers what they would to at each possible node of these mini-ultimatum games without knowing the P's proposal. Each player has played all the four games, the opponents have been assigned randomly and were anonymous, the first-mover status has been assigned randomly, but most importantly the outcomes of the games have been told only after the ending of all four of them, not to influence the players any further.

Comparing the (8/2)-offer amongst all the games the results show that the highest rejection rate for this offer is in the (a) (5/5) game, with a rejection of 44.4%; the second is the (b) (2/8) game, with a 26.7% rate of rejection; the third is the (c) (8/2) game with a rejection rate of 18%; and last the (d) (10/0) game with a rate of 8.9%. The differences in rejection rate in those games are statistically significant ($p < .0001$).

From the 18% rejection rate of the (c) (8/2) game, the one where the player P has no actual choices, we can see that the consequences on the player R, or better in this case the outcome, have an impact on reciprocity.

Moreover, since this rejection rate grows up to 44.4% in the (a) (5/5) game, we can observe that also the intentions play an important role for what concerns reciprocal behavior, otherwise it would have kept the same level.

A further experiment to analyze the negative reciprocity is the ultimatum game (Falk, Fehr, & Fischbacher, 2003). In this sequential game, there are two players that have to split an amount of money; a first mover, known also as the proposer, that decides the division between his part and the other player's part, and a second-mover, known also as responder, which decides whether to accept the split proposed or to reject it, where if he rejects it both players receive nothing. The results of this game are quite consistent each time that it is played:

- Generally, the offers don't exceed the 50-50 split;
- The modal offer is between 50-50 and 60-40, where 60 is the proposer's part;
- Offers below 80-20 are extremely rare;
- If offers close to 50-50 are practically never rejected, the rejection rate for offers close to 80-20 have a high rejection rate.

The "standard" theory would predict diametral different outcomes. In fact, the optimal solution from the Game Theory's point of view would be a 99-1 split, since the proposer knows that 1 is better than 0, and that so the responder will then always accept given his rationality. This difference in the theoretical and actual outcome lies in the prosocial and reciprocal component. One of the major assumptions in economics is that agents (people) behave and act according to their own self-interests and behaviors that deviate from it are seen as irrational. However, there are many situations where being self-interested is worse off for the total welfare. The first experiment made to show it that will be presented is the one made by Berg et al. (1995). Their experiment is quite easy, and it is based upon an investment game.

The investment game as presented by those authors was composed by two stages. The players have been split into two separate rooms, A and B, in order to keep the anonymity, and each received a $10 show-up fee. In the first stage the player in the room A has to decide how much of their show-up fees, denoted with M_a, send to an anonymous counterpart in room B. This amount will be tripled before being given to the counterpart. In the second stage is the time of the counterpart in the room B to decide how much money to give back in return, denoted with $k_b(3M_a)$.

The strategy that may be implemented by player in room A are

$$\mathrm{M_a} = \{0, 1, 2, ..., 10\}$$

Whereas players in room B can choose the strategy

$$k_b = \{0, 3, 6, ..., 30\} \rightarrow \{0, 1, 2, ..., 30\}$$

Satisfying the constraint $0 \leq k_b(3M_a) \leq (3M_a)$.

The payoffs associated with these strategies are

$$P_a(M_a, k_b) = \$10 - M_a + k_b(3M_a)$$

And

$$P_b(M_a, k_b) = 3M_a - k_b(3M_a)$$

According to Game Theory the best strategy for this game would be to send nothing in both rounds. The reason again is really easy, since there is no assurance that the player in the room B will send back any amount of money, for the player in the room A sending money would be just a decrease in his welfare. Subjects use trust to facilitate exchange, if the following conditions are met (1995):

- Placing trust in the trustee (the player in the room B) puts the trustor at risk (the player in the room A);

- Relative to the set of possible actions, the trustee's decision benefits the trustor at a cost to the trustee;

- Both trustor and trustee are made better off from the transaction compared to the outcome which would have occurred if the trustor had not entrusted the trustee.

If there is the presence of trust and reciprocity, then that would mean at least one player in the room A has sent money to his counterpart, so the assumptions made by Berg et al. (1995) are:

$$A_0: M_a > 0$$

And

$$A_1: k_b(3M_a) > M_a$$

Since the simple strategy of not investing results in zero profit, trust cannot emerge as a norm unless it satisfies the positive profits hypothesis that, for at least some amounts M_a, the average net return is positive (Berg, Dickhaut, & McCabe, 1995), and then

$$A_2: \sum_b \frac{k_b(3M_a)}{N} > M_a$$

Rabin (1993) makes such an assumption by incorporating a "kindness" function into subjects' utility in such a way as to capture the following behavior: as one's counterpart increases his or her "kindness", the utility maximizing response is to be kinder in return. This suggests that:

$$A_3: M_a \; and \; k_b(3M_a) \; are \; positively \; correlated$$

Given all the results, it has been shown that the average sent has been \$5.16 for room A and the amount sent back \$4.66 for room B. The first assumption has actually little support, but however it cannot be rejected since 2 participants out of 32 in room A have sent zero to their counterparts. In room B we can see that, of the 28 participants that received an amount M_a greater than \$1, 11 returned more than their counterpart sent (Berg, Dickhaut, & McCabe, 1995), so this behavior support definitely the assumption 1. Also, the assumption 2 hold, since the average payback for \$5 investments and \$10 have been respectively \$7.17 and \$10.20. However, it has been shown that there is no actual correlation between the amount participants in room A have sent and the amount that the counterpart sends back (assumption 3 does not hold).

In order to further investigate the field of reciprocity Charness and Rabin (2002) have designed a wide range of simple experiments in order to find out what is that motivates individuals. Their data consists of 29 different games, that have been played by 467 participants, for a total of 1697 decisions.

The first step is to find out if individuals are motivated by social preferences (e.g. prosocial behaviors), and we consider two players A and B, we can outline the player B's preferences as:

$$U_B(\pi_A, \pi_B) = (\rho \cdot r + \sigma \cdot s + \theta \cdot q) \cdot \pi_A + (1 - \rho \cdot r - \sigma \cdot s - \theta \cdot q) \cdot \pi_B$$

Where:

π_A is the player A's monetary payoff

π_B is the player B's monetary payoff

$r = 1$ if $\pi_B > \pi_A$, and $r = 0$ otherwise;

$s = 1$ if $\pi_B < \pi_A$, and $s = 0$ otherwise;

$q = -1$ if A has misbehaved, and $q = 0$ otherwise.

This formulation says that B's utility is a weighted sum of her own material payoff and A's payoff, where the weight B places on A's payoff may depend on whether A is getting a higher or lower payoff than B and on whether A has behaved unfairly (Charness & Rabin, 2002). The parameters θ, σ, and ρ are respectively a mechanism for modeling reciprocity, and ranges of different "distributional preferences" based upon outcomes. When a player "misbehaves", then the value of the parameter θ is greater than 0, lowering the values of σ and ρ by an amount θ.

The experiments have been composed by fourteen sessions at the Universitat Pompeu Fabra in Barcelona and at the University of California at Berkley. The total treatments have been 32, the first 12 in Barcelona (Barc) and the last 20 at Berkley (Berk). In the games from 1 to 4 players played one game, from 5 to 12 players played two games, from 13 to 32 (all the experiments done at Berkley) each participant played four games.

The games played have been a two-person dictator game, three slightly different two-person response games, a three-person dictator game, and a three-person response game.

The tables that follow show the results for those games. The numbers in brackets after the label of the experiment represent the number of players that have undertaken the treatment, whereas the other numbers in parenthesis show (A,B) or (A,B,C) monetary payoffs. Source: (Charness & Rabin, 2002).

	Two-person dictator games			Left	Right
Berk29 (26)	B chooses (400,400) vs. (750,400)			.31	.69
Barc2 (48)	B chooses (400,400) vs. (750,375)			.52	.48
Berk17 (32)	B chooses (400,400) vs. (750,375)			.50	.50
Berk23 (36)	B chooses (800,200) vs. (0,0)			1.00	.00
Barc8 (36)	B chooses (300,600) vs. (700,500)			.67	.33
Berk15 (22)	B chooses (200,700) vs. (600,600)			.27	.73
Berk26 (32)	B chooses (0,800) vs. (400,400)			.78	.22
	Two-person response games— B's payoffs identical	Out	Enter	Left	Right
Barc7 (36)	A chooses (750,0) or lets B choose (400,400) vs. (750,400)	.47	.53	.06	.94
Barc5 (36)	A chooses (550,550) or lets B choose (400,400) vs. (750,400)	.39	.61	.33	.67
Berk28 (32)	A chooses (100,1000) or lets B choose (75,125) vs. (125,125)	.50	.50	.34	.66
Berk32 (26)	A chooses (450,900) or lets B choose (200,400) vs. (400,400)	.85	.15	.35	.65

Two-person response games— B's sacrifice helps A		Out	Enter	Left	Right
Barc3 (42)	A chooses (725,0) or lets B choose (400,400) vs. (750,375)	.74	.26	.62	.38
Barc4 (42)	A chooses (800,0) or lets B choose (400,400) vs. (750,375)	.83	.17	.62	.38
Berk21 (36)	A chooses (750,0) or lets B choose (400,400) vs. (750,375)	.47	.53	.61	.39
Barc6 (36)	A chooses (750,100) or lets B choose (300,600) vs. (700,500)	.92	.08	.75	.25
Barc9 (36)	A chooses (450,0) or lets B choose (350,450) vs. (450,350)	.69	.31	.94	.06
Berk25 (32)	A chooses (450,0) or lets B choose (350,450) vs. (450,350)	.62	.38	.81	.19
Berk19 (32)	A chooses (700,200) or lets B choose (200,700) vs. (600,600)	.56	.44	.22	.78
Berk14 (22)	A chooses (800,0) or lets B choose (0,800) vs. (400,400)	.68	.32	.45	.55
Barc1 (44)	A chooses (550,550) or lets B choose (400,400) vs. (750,375)	.96	.04	.93	.07
Berk13 (22)	A chooses (550,550) or lets B choose (400,400) vs. (750,375)	.86	.14	.82	.18
Berk18 (32)	A chooses (0,800) or lets B choose (0,800) vs. (400,400)	.00	1.00	.44	.56

Two-person response games— B's sacrifice hurts A		Out	Enter	Left	Right
Barc11 (35)	A chooses (375,1000) or lets B choose (400,400) vs. (350,350)	.54	.46	.89	.11
Berk22 (36)	A chooses (375,1000) or lets B choose (400,400) vs. (250,350)	.39	.61	.97	.03
Berk27 (32)	A chooses (500,500) or lets B choose (800,200) vs. (0,0)	.41	.59	.91	.09
Berk31 (26)	A chooses (750,750) or lets B choose (800,200) vs. (0,0)	.73	.27	.88	.12
Berk30 (26)	A chooses (400,1200) or lets B choose (400,200) vs. (0,0)	.77	.23	.88	.12

Three-person dictator games		Left	Right
Barc10 (24)	C chooses (400,400,x) vs. (750,375,x)	.46	.54
Barc12 (22)	C chooses (400,400,x) vs. (1200,0,x)	.82	.18
Berk24 (24)	C chooses (575,575,575) vs. (900,300,600)	.54	.46

Three-person response games		Out	In	Left	Right
Berk16 (15)	A chooses (800,800,800) or lets C choose (100,1200,400) or (1200,200,400)	.93	.07	.80	.20
Berk20 (21)	A chooses (800,800,800) or lets C choose (200,1200,400) or (1200,100,400)	.95	.05	.86	.14

In order to compare the power of self-interest and distributional models (competitive, difference-averse, and social welfare) Charness and Rabin (2002). in their experiment considered mostly how many observations in their games were consistent with the values of and permitted by the restrictions for each type of social preferences, when excluding reciprocity by imposing the restriction = 0 . Moreover, since reciprocity motivation are not considered yet, to analyze this difference, it is correct to take into account for the research just the seven dictator games.

The table below sums up all the results for the seven dictator games in the categories considered. Source: (Charness & Rabin, 2002)

	Total # observations	Narrow self-interest	Competitive	Difference aversion	Social welfare
B's behavior in the dictator games	232	158 (68%)	140 (60%)	175 (75%)	224 (97%)
B's behavior in the response games	671	532 (79%)	439 (65%)	510 (76%)	612 (91%)
B's behavior in all games	903	690 (76%)	579 (64%)	685 (76%)	836 (93%)
A's behavior, any predictions by A	671	636 (94%)	579 (86%)	671 (100%)	661 (99%)
A's behavior, correct predictions by A	671	466 (69%)	488 (73%)	603 (90%)	649 (97%)
All behavior, any predictions by A	1574	1326 (84%)	1158 (74%)	1356 (86%)	1497 (95%)
All behavior, correct predictions by A	1574	1156 (73%)	1067 (68%)	1288 (82%)	1485 (94%)

As we can see, the social-welfare preferences are much higher than all the others, that means they are better suited to explain behavior when there is no reciprocity; also difference aversion can actually be a suitable candidate, whereas narrow self-interest and competitiveness have performed worse.

We can also observe that a substantial number of players refused to receive less than another person when such refusal is costless (Charness & Rabin, 2002). Moreover, as we can see also from the table of the results within the two-person dictator games Berk8 and Berk 15, in both experiment player B is willing to help player A, and this help is consistent for both Different aversion and Social welfare.

The contrast in this behavior however, shows that the player B is less willing to sacrifice 100 to help A by 400 when by doing so he receives a lower payoff than A (Charness & Rabin, 2002).

When it comes to players A, we can see that 27% of them behaved in a way that intended an expected sacrifice from player B, meaning that the behaviors of players A is much more self-interested than players B's behavior. This difference has been shown by the authors to be significant at $p \approx .00$. Moreover, we can also note that in the games where players A could make a decision that could only lose their money but could also help B, the 33% of those players sacrificed. These factors can let us conclude that departures from self-interest are just as common for A's as for B's (Charness & Rabin, 2002).

The table below summarizes how each model performs in each class of choices among those choices where only one of the two choices is compatible with the model. Source: (Charness & Rabin, 2002).

Class of games	Narrow self-interest	Competitive	Difference aversion	Social welfare
B's behavior in the dictator games	132/206 (64%)	104/196 (53%)	49/106 (46%)	54/62 (87%)
B's behavior in the response games	346/479 (72%)	319/551 (58%)	350/517 (68%)	304/363 (84%)
B's behavior in all games	478/685 (70%)	423/747 (57%)	399/623 (64%)	358/425 (84%)
A's behavior, any predictions	172/226 (76%)	212/304 (70%)	32/32 (100%)	74/84 (88%)
A's behavior, correct predictions	466/671 (69%)	364/553 (66%)	181/249 (73%)	134/150 (89%)
All behavior, any predictions by A	650/911 (71%)	635/1051 (60%)	431/655 (66%)	432/509 (85%)
All behavior, correct predictions	944/1356 (70%)	787/1300 (61%)	580/872 (67%)	492/575 (86%)

The other results of the study shown also that players B are less likely to cause Pareto damage when this decreases inequality than when Pareto damage increases inequality (Charness & Rabin, 2002), as shown by the table below that summarizes the B's sacrifice rate by effect on inequality.

Class of games	Sacrifices/ chances	Probability of sacrifice
Games allowing Pareto-damage	59/357	17%
Decreases inequality	34/228	15%
No effect on inequality	4/35	11%
Increases inequality	21/94	22%
Games where sacrifice helps A	199/546	36%
Decreases inequality	99/212	47%
No effect on inequality	8/68	12%
Increases inequality	92/266	35%
All games	268/903	30%
Decreases inequality	133/440	30%
No effect on inequality	12/103	12%
Increases inequality	123/360	34%

Source: (Charness & Rabin, 2002)

The first three games set up by the authors that give us information on the reciprocal behavior ($\theta > 0$) are the three summarized in the table below

	Games with the choice between (400,400) and (750,400)	(400,400)	(750,400)
Berk29 (26)	B chooses (400,400) vs. (750,400)	.31	.69
Barc7 (36)	A chooses (750,0) or lets B choose (400,400) vs. (750,400)	.06	.94
Barc5 (36)	A chooses (550,550) or lets B choose (400,400) vs. (750,400)	.33	.67

As we can see when a player B has to decide whether to split (400,400) or (750,400), where the first number is the outcome of player A, the first strategy is played 31% of the times. However, if before making that decision there has been a "charitable" act from the player A, the percentage of that strategy drops to 6%. This is a behavior that we would expect given all the data collected so far. On the other hand, if we consider the experiment Berk5, we would expect the behavior of player B being punitive, since first of all it doesn't cost anything to player B, but the percentage of strategy (400,400) has increased by just a 2%.

If we consider also other seven games, where still B has to decide whether to sacrifice and help A, having the two viable strategies being (400,400) or (750,375), as shown in the table below,

Games with the choice between (400,400) and (750,375)		(400,400)	(750,375)
Barc2 (48)	B chooses (400,400) vs. (750,375)	.52	.48
Berk17 (32)	B chooses (400,400) vs. (750,375)	.50	.50
Barc3 (42)	A chooses (725,0) or lets B choose (400,400) vs. (750,375)	.62	.38
Barc4 (42)	A chooses (800,0) or lets B choose (400,400) vs. (750,375)	.62	.38
Berk21 (36)	A chooses (750,0) or lets B choose (400,400) vs. (750,375)	.61	.39
Barc1 (44)	A chooses (550,550) or lets B choose (400,400) vs. (750,375)	.93	.07
Berk13 (22)	A chooses (550,550) or lets B choose (400,400) vs. (750,375)	.82	.18

Source: (Charness & Rabin, 2002)

we can see that players B are willing to sacrifice to purse the social-welfare-maximizing allocation when they feel neutral towards A (Charness & Rabin, 2002).

However, if we compare the experiments Barc1 with Barc2, and Berk17 with Berk 13, we can observe that when the player A behaves in a selfish way, player B will reciprocate, being less willing to sacrifice to A, in fact, for the first two experiments we have a drop from 48% of strategy (400,400) to a 7%, and for the second two we have a drop from 50% to a 18%. Both these differences are significant at $p < .01$ (Charness & Rabin, 2002).

.

The table below summarizes the B's Response as a function of A's Help or Harm.

Source: (Charness & Rabin, 2002)

Class of games	Sacrifices/chances	Probability of sacrifice
All games allowing Pareto-damage	59/357	17%
A has helped B	2/36	6%
A has had no play	8/62	13%
A has hurt B	49/259	19%
All games where sacrifice by B helps A	199/546	36%
A helped B	100/278	36%
A had no play*	88/202	44%
A hurt B in violation of SWP	7/66	11%

When A hurts B, B is more likely to hurt A than otherwise and more likely to withdraw willingness to sacrifice to help A. The difference in Pareto-damaging B behavior when A helps B and when A hurts B is significant at $p \approx .02$; comparing behavior when A hurts B and when A either has no play or helps B is also significant at $p \approx .02$ (Charness & Rabin, 2002).

A comparison that we can make here between individuals and group is regarding the willingness to make a sacrifice, in fact, as shown by Charness and Sutter (2012), groups are more individualistic and then their propensity to make a sacrifice will be less than the individuals' one; his also has an impact on the welfare, since it will be decreased by the decisions made by groups.

Charness and Rabin (2002) made also three-person experiments also to observe what would happen with a "disinterested" player. Their results showed that people care about both the total surplus and the minimum payoff among others (as we can see from experiments Barc10 and Barc12), but also that people are willing to sacrifice some of their outcomes in order to equalize with the other players.

Moreover, Barc10 and Barc12 together show that social efficiency is not the only distributional driving force, as (1200,0,x) is more socially efficient than (750,375,x), but is chosen much less frequently ($p \approx .01$).

One of the most common problem faced by societies is the problem of public goods, a commodity or a service provided by a central entity that it is non-excludable (anyone that wants to use it can) and non-rival (if someone uses the public good that doesn't prevent anyone else to use it too). Public goods to be set up require a certain minimum amount of contribution, but once they are implemented anyone can use them. The problem with public goods is that the agents that have to contribute for their implementation, are incentivized to free-ride and not to contribute. To explain the problem, I will refer to the work of Fehr and Gätcher (2000).

In their experiment there were four group members, each provided with 20 tokens. The subjects were requested to decide simultaneously how many tokens of the 20 they wanted to invest for the public good, and how many to keep for themselves. Each token invested in the public good, whether the public good is implemented or not, has a private return of 0.4 per token, with a social return of 1.6, whereas each token kept has a private return of 1, but no social return. As we can see the maximum social return is present when all participants invest all their tokens, with each participant earning 32 tokens, however it is in the self-interest of each participant to free ride, contributing nothing, since this would guarantee an earning of 52 tokens, for a final total outcome of 72.

So far there is little space for reciprocity, since there is no opportunity for the players to punish each other. But negative reciprocity can play the role that if subjects expect that others free ride, and if they interpret that as a hostile act, then they can "punish" others by free riding, too. Fehr and Schimidt (1999) shown theoretically that even a minority of reciprocal subjects is capable of inducing a majority of selfish subjects to cooperate when there is the possibility to pay a price to act reciprocally.

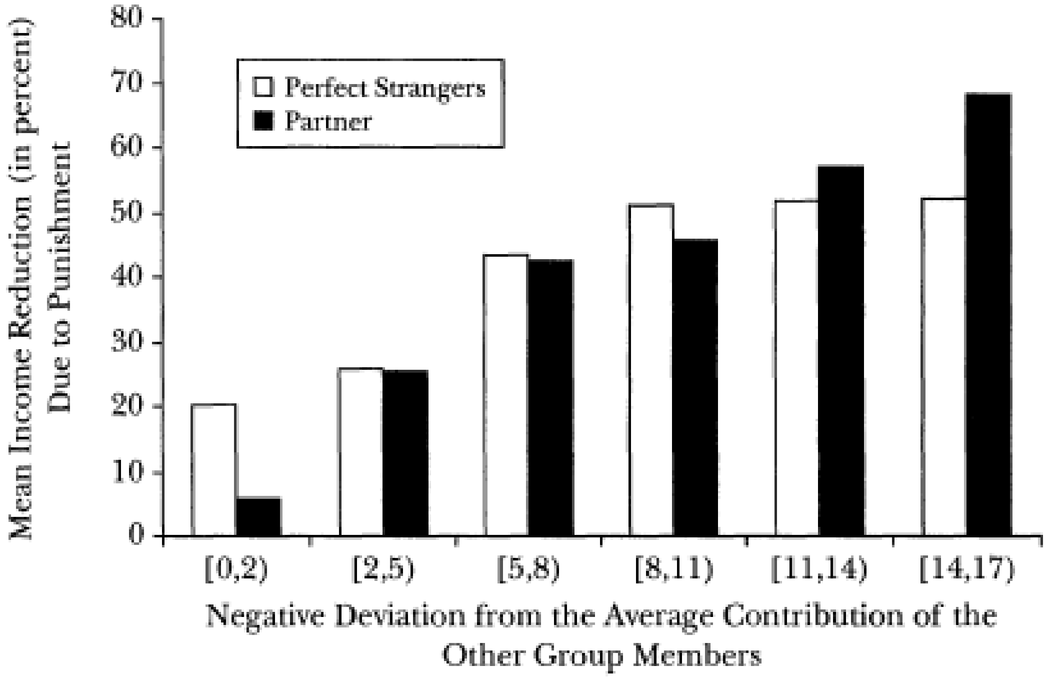
These are the reasons why Fehr and Gätcher (2000) divided their experiment into two version:

- The Perfect stranger version. It has been repeated six times and each time the groups were newly formed so that players were not playing again with each other. This ensures that there are no future rewards for an action in a period. Each player knew in advance how many rounds were left;

- The Partner version. This version has been repeated ten times and the participants were not shuffled, so they knew to play again with all the same members, giving the opportunity to implement some strategic behaviors.

The punishment in this experiment has been measured as the reduction in percentage terms of the punished subjects' outcome from round to round. It turns out that the negative deviation from others' average contributions to the public good is a strong determinant of punishment;

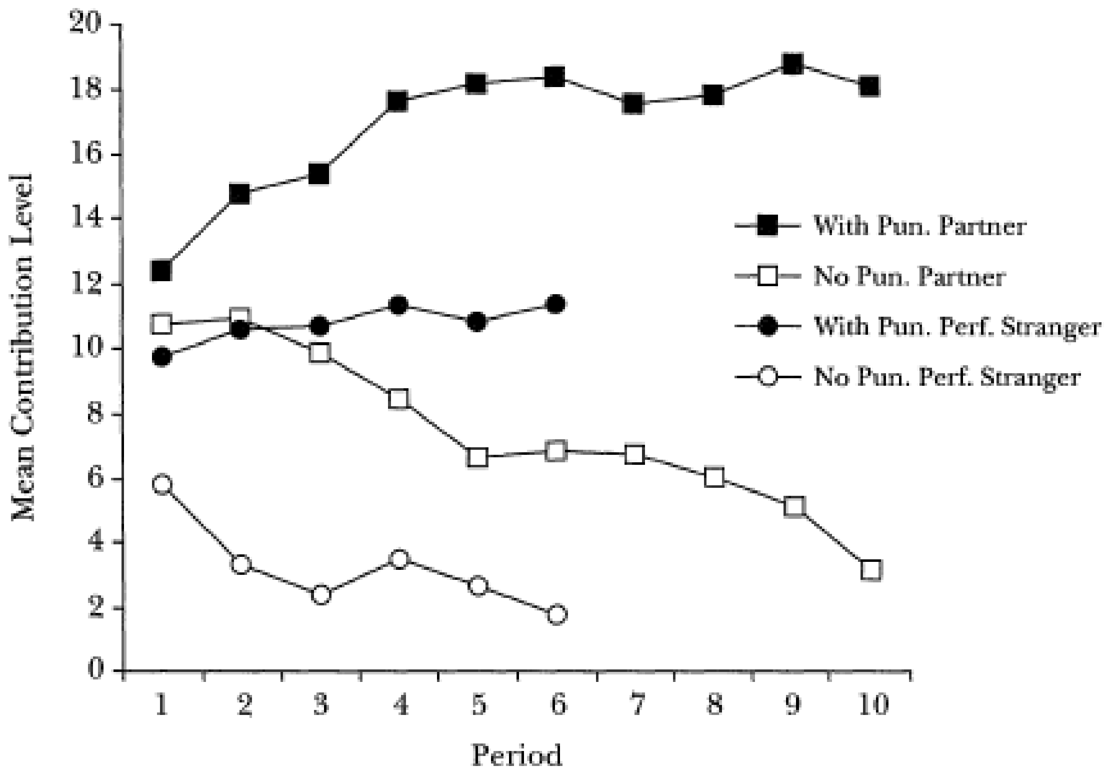
the more a subject free-rides relative to the others the more it gets punished (Fehr & Gätcher, 2000). The results of the fame are shown in the table below.



Source: (Fehr & Gätcher, 2000).

in period six of the Perfect Stranger version, 79% of the subjects free-rides completely and the rest contributes little; this high defection rate stands in sharp contrast to the contribution behavior in the games with a punishment opportunity: When subjects are perfect strangers, they can at least stabilize contributions at relatively high levels; in the Partner version they almost converge to the maximum level of contributions .

The table below represents the evolution of average contributions with and without the punishment option in the Partner and Perfect Stranger Condition. Source: (Fehr & Gätcher, 2000).



As we can see reciprocity in individuals plays an important role and can change consistently the outcome of the game.

2.4 EVIDENCES FOR ECONOMIC APPLICATIONS

All these studies have an important application in the economic field. In fact, knowing whether a certain contract is better than another one, or if it is better to delegate to a group or a single individual, is a matter on a day-to-day basis in most of the organizations.

Free riding effects and reciprocity can be present also within an organization, not just in public goods or experimental games, it is enough to think about work with high complexity during their production process, there is no way to know whether a worker is performing correctly or free-riding on the others.

For example, Fehr, Gächter and Kirchsteiger (1997) in order to verify if reciprocity plays a role in the effort level, made a laboratory experiment where experimental employers could offer a wage contract that stipulated a binding wage w and a desired effort level $\hat{e}$, and if an experimental worker accepted this offer, the work was free to choose the actual effort level e between a minimum (1) and a maximum level (10). Of course, higher level of effort was reflected in higher profit for the employer, but the salary for the employee didn't change for any value of *e*. The profit for the employer was expressed as

$$\pi = 10e - w$$

And the outcome of the worker as

$$u = w - c(e)$$

Where *c(e)* represents the cost of the effort for the employee, and *e* = 1 has a *c(e) = 0*.

In each session played there were more workers (8) than employers (6), and it was not allowed to hire more than 1 worker.

The results showed that many employers made quite generous offers, in fact on average, the offered contracts stipulate a desired effort of $\hat{e}$ = 7 and the offered wage implied that the worker receives 44 percent of the total surplus $u + \pi$; interestingly, many workers honor this generosity somewhat but not fully; the actual average effort is given by e = 4.4, which was substantially above the selfish choice of e = 1; however, only in 14 percent of all cases workers abide by the terms of the contract, while in 83 percent of all cases they shirk .

This means that on average, people may be willing to put extra effort in addition to what is desired in the contract if the job offer is generous.

However, when employers can punish employees if they don't perform as requested things changes as shown by Fehr, Gächter and Kirchsteiger (1997). Their experiment is equivalent to the one just described; however, the employers have the possibility to punish or rewards their employees after thy reveal their effort level.

For every token spent on rewards they could raise the worker's monetary income by 2.5 tokens and for every token spent on punishment, they could reduce the worker's income by 2.5 tokens, we would expect the employers neither punish nor reward the employees since it is costly.

The results have shown that if workers shirked in the experiments, employers punished in 68 percent of these cases; if there was overprovision, employers rewarded in 70 percent of these cases; and if workers exactly met the desired effort, employers still rewarded in 41 percent of the cases (Fehr, Gachter, & Kirchsteiger, 1997).

Comparing the two experiments, when there is the possibility of a reward or a punishment, employees have a higher effort level. In fact, the shrink rate has dropped from the 83% of the first experiment to a 26% when there's the possibility of a punishment/reward, but most importantly there has been also an increase in terms of monetary payoff.

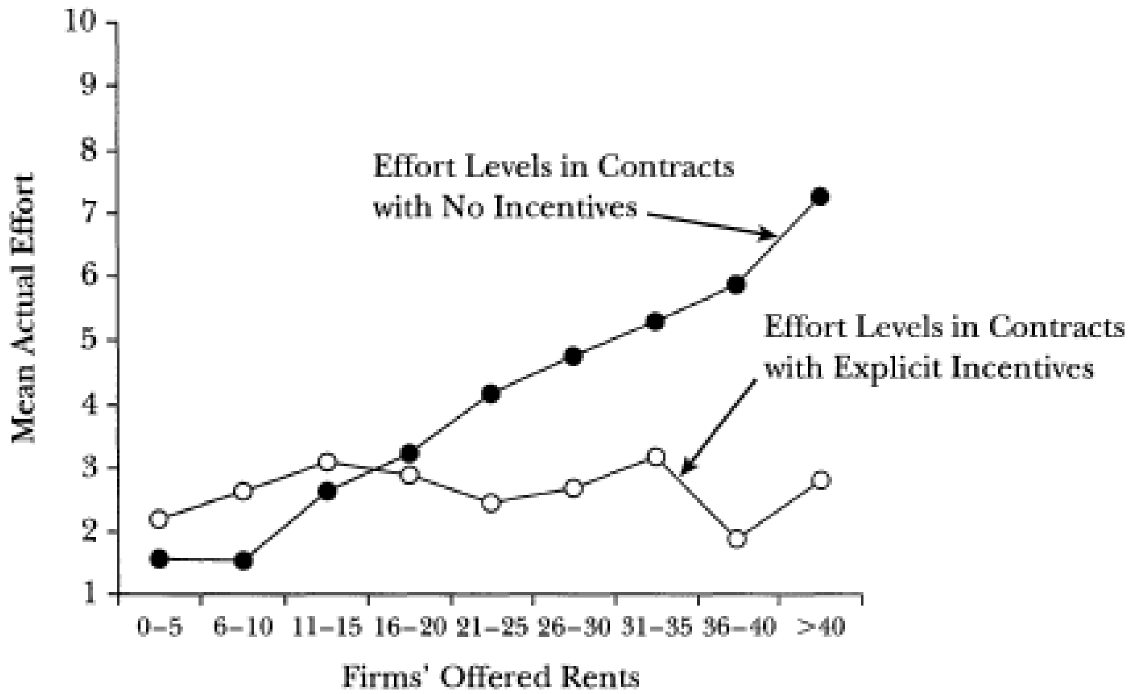
All this suggests that reciprocity actually contributes to the enforcements of contracts (Fehr, Gachter, & Kirchsteiger, 1997).

However, it is also possible that explicit incentives may cause a hostile atmosphere of threat and distrust, which reduces any reciprocity-based extra effort (Fehr, Gachter, & Kirchsteiger, 1997). In fact, as shown by Bewley (1995), managers stress that explicit "punishment should be rarely used as a way to obtain co-operation" because of the negative effects on work atmosphere.

This is why in another experiment Fehr and Gächter (2000), examined how worker reacts when there are no reward and punishment possibility, but there is an explicit performance incentive. This incentive consists of a fine that have to be paid by the worker to the employer if that worker is caught shrinking. The asset of this game is the same of the other two just described.

The maximal fine is fixed at a level such that a selfish risk-neutral worker will choose an effort level of 4 when faced with this fine, and has a .33 probability of verification (2000).

The differences are shown in the graphs below



	Mean contributions in all periods	
Groups	Without punishment opportunity	With punishment opportunity
1	7.0	17.5
	(6.3)	(4.3)
2	10.6	16.4
	(8.5)	(5.2)
3	6.7	18.4
	(7.8)	(3.6)
4	5.1	12.1
	(6.3)	(7.1)
5	6.4	14.3
	(7.2)	(7.0)
6	7.9	19.0
	(5.7)	(2.8)
7	7.4	19.0
	(7.1)	(3.4)
8	10.0	17.2
	(6.6)	(4.3)
9	3.9	17.0
	(5.9)	(5.0)
10	10.0	19.0
	(6.6)	(2.1)
Mean	7.5	17.0
	(6.8)	(4.5)

Source: (Fehr & Gätcher, 2000).

As we can see when there are no incentives the mean actual effort is higher than when there are explicit incentives, this means that explicit incentives may "crowd out" the effort level that employees may be willing to implement.

For the employers, the savings in wage costs more than offset the reductions in revenues that are caused by the lower effort in the incentive treatment; however, while the wage savings merely represent a transfer from the workers to the firms, the reduction in effort levels reduces the aggregate surplus; this shows that, in the presence of reciprocal types, efficiency questions and questions of distribution are inseparable (Fehr & Gätcher, 2000).

CHAPTER 3

3.1 THE DECREASING CENTIPEDE EXPERIMENT

To provide further evidence regarding differences in terms of rationality between individuals and groups and observe also differences in reciprocity levels, I have set up a sequential game, which is a revision of the classic centipede game. In this revised centipede there are two competitors who have to allocate an amount of €100 in a sequential game and I will test if

- *Hypothesis 1*: Groups are more rational than individuals.

- *Hypothesis 2*: Groups are less reciprocal than individuals.

The first-mover can propose to split the amount in any way he prefers, if the other player agrees then the game will end, otherwise the second player will make a counteroffer, this process will go on until a player accepts or until the fourth proposal. Each time that a player refuses the amount of money will be decreased by €20, and the last player to make an offer will be the one that was the respondent at the first stage. At the fourth proposal if the player 1 refuses both players will get 0 (Figure 4).

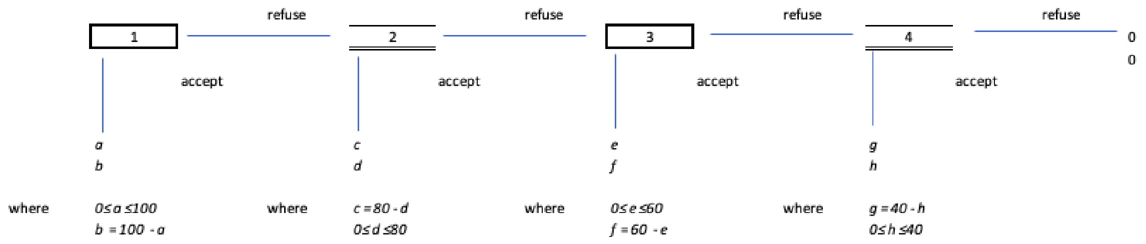


Figure 4

If we observe the formalization of this game, as shown in Figure 4, we see that the numbers with a thick box border represent the nodes where player 1 has to propose the split to player 2, whereas the numbers with a top and double bottom border represents the nodes where the player 2 has to make a proposal; in each node the outcomes on top represents the outcomes of player 1.

As for the other centipede games, this game can be easily solved through a backward induction reasoning, and the formalization without variables for the optimal solution will be as shown in Figure 5.

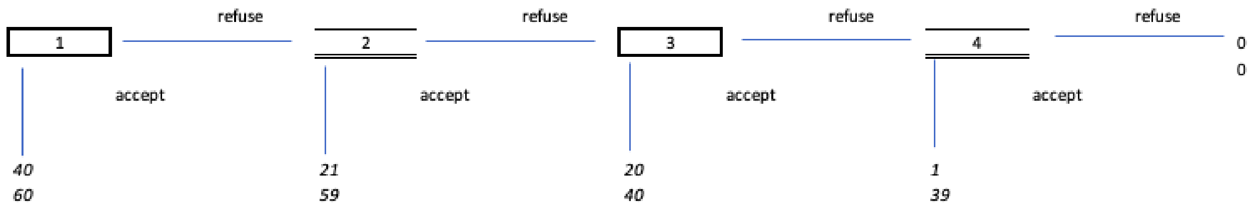


Figure 5

Since player 2 knows that if player 1 refuses his offer they will both get zero, he will offer him 1 and will take the rest. At the previous node then player 1 will be incentivized to offer player 2 an amount slightly higher than the one in the fourth node, then he will offer him 40 and propose to take the rest. This process will continue until the first node where the optimal strategy will be for player 1 to offer 60 to player 2 and take 40 for his own.

The sample for this experiment is composed by 80 persons, of which 20 will play as individuals and the others will be aggregate in groups of three randomly. To avoid spillover effects no one knew who his opponent was, the game has been proposed and given just in the parametrical form, and no time limits have been imposed. Offers and counteroffers have been taken and proposed by electronical devices not to influence the choices.

I decided to split the sample into three different macro groups, in the first group there were 7 individuals competing with other 7 individuals, the second was composed by 6 individuals playing against 6 groups, whilst the last group was composed by 7 groups against other 7 ones. The main difference with this experiment from the one conducted by G. Bornstein, T. Kugler, and A. Ziegelmeyerb (2004) is that through this centipede there may be the possibility to individuate the reciprocity effect. If expectations are violated in the split of the pile of money, we will expect the players to reciprocate, trying to drag down the final outcomes.

The results actually show that for many players, especially when it comes to individual players against each other, it is totally fine and acceptable to split into even parts the pile of money even if it is not the best strategy.

The table below shows the node where each macro-group stopped. In "individuals vs groups" are expressed the values with individuals as first proposers.

	1	2	3	4	no agreement
Individual vs individual	4	0	1	1	1
Individual vs group	2	1	0	0	2
group vs individual	1	1	0	0	0
group vs group	6	0	0	1	0

During the experiment the individuals has been allowed to decide the first split 12 times, and the 67% of those times they have proposed to split even the pile of money. The average of the amount proposed by the individuals in the first turn has been 54.92-45.08, but most importantly it has a gap from the best move of 14.92; while the groups has decided first the split for 9 times, and the average proposed has been 46.1-53.9, with a gap from the optimal of 6.1.

3.2 HYPOTHESIS 1

This first data is consistent with the literature for what concerns the individuals, as also showed by the work of Falk et al. (2003). The results however show an interesting result for what concerns the groups. To analyze this difference and test the first hypothesis I will re-adapt a volatility measure utilized by hedge-funds, called Tracking Error Volatility, to observe the standard deviation percentage difference of a portfolio in respect to a benchmark portfolio the fund is trying to imitate.

$$TE = \sqrt{\frac{\sum_{i=1}^{n}(R_i - R_B)^2}{N-1}} = \sigma(R_i - R_B)$$

Where TE is the Tracking Error, R_i is the fund's return of the *ith* portfolio, R_B is the return of the benchmark and N is the number of periods.

In this case the benchmark will be the game-theory right answer and as observations the proposed splits, N will be the number of players. The higher the value of the TE will be, the less accurate the answers sample observed. So, if the Hypothesis 1 is accurate we will expect individuals to have a higher value for he TE and groups to have lower ones. This analysis will be done just for the first round, since for the others the number of observations would be too small to capture satisfying data.

For the individuals proposing at the first round we have that the Tracking Error Volatility is

$$TE = \sqrt{\frac{\sum_{i=1}^{12}(R_i - 40)^2}{12 - 1}} = 21.49$$

Whereas for groups proposing at the first round, we have that

$$TE = \sqrt{\frac{\sum_{i=1}^{9}(R_i - 40)^2}{9 - 1}} = 8.45$$

As said before we were expecting the less rational macro-group to have a higher TE compared to the more rational. Since the individuals' TE is higher than the groups' one, hypothesis 1 holds.

For what concerns the second hypothesis that groups are less reciprocal than individuals, there are a little more problems to investigate it. As a starting point we may say that the acceptancy rate for individuals in the first round has been 55.6%, with an average proposed of 49, whereas the average rejected as the first round has been 37.75. For the groups in the first round the acceptancy rate has been 66.67%, with an average proposed of 54.25, on the other hand there has been an average of 47.5 rejected. From this first data we can observe that whereas the individuals have had a behavior predictable in a certain way for what concerns reciprocity, the fact that groups have rejected an average amount of 47.5 is a first clue to prove hypothesis 2, and a further proof for hypothesis 1. In fact, from a classical game of theory point of view rejecting an average amount of 47.5 is totally right, but from the studies done by Falk et al. (2003) we know that generally 50-50 offers, or close to them, are almost never rejected by reciprocal players.

Another datum interesting to observe is the one relative to the fourth node. While individuals have been 4 times in the spot to decide whether to accept the offer or end the game with no rewards, the groups have been in this spot just once. For what concerns the individuals as responders the three splits rejected have been the 1-39 splits, even if it is a suboptimal solution (the first number is the one that would be the reward for the respondent player); whereas the accepted offer has been 10-30. For what concerns the groups the only offer made at the fourth node has been a 1-39, accepted by the respondent.

From these data we can say that Hypothesis 2 holds, and then groups are less reciprocal than individuals. This may due to the fact that the groups in a certain way mitigate the reciprocal attitude of each component. However, it is important to say that even if this hypothesis holds, it may be due the small sample and then further studies with a bigger sample are required.

3.3 HYPOTHESIS 2

Since the data from the decreasing centipede experiment were insufficient to further test the hypothesis that groups are less reciprocal than individuals, an ultimatum game experiment will be conducted to analyze the behavior of groups.

Since the analysis is on reciprocity and not on rationality, the groups will not be required to propose a split of a 100€ amount, they will just have to decide whether to accept or not the offer.

The splits proposed will be plotted in order to analyze the behavior in each situation. Since presumably the splits above 50-50 will always be accepted and realistically will almost never be proposed, there will be 50 proposal to 50 groups of 3 people, starting from an even split, arriving to a 99-1 proposal, where 1 will be the reward of each member of the group.

To have more data, before discussing with the group whether to accept, each group member will be required to write on a piece of paper if they would accept the split, this will let acquire more data to compare groups' and individuals' behavior.

No instructions on how to give a final answer on the response have been given to groups, in order to leave the most possible freedom. Moreover, the participants didn't know the proposes were already plotted, and they have been told to play against other players not to false their behavior.

Since the rationality of the groups has already been demonstrated, we will expect now a smaller amount in terms of percentage of groups rejecting the proposals compared to the individuals. Realistically this difference can be due to the fact that groups, as it was for the beans-in-a-jar experiment, tend to smooth the individuals' choices, so more reciprocal individuals will be needed to have a reciprocal group.

The results show that, as expected, in the splits from 50-50 to 59-41 neither individuals nor groups rejected the offer. In the splits from 60-40 to 69-31 there has been the same rates for groups and individuals, equal to 80% acceptancy rate and 20% rejection rate. At this point there is the divergence between individuals and groups. Whereas the groups keep the same acceptancy-rejection rate constant and equal to 70%-30%, individuals have a growing rejection rate the more the stake proposed is lower. In fact, in the range from 70-30 to 79-21 the rejection rate has been the 33%, in the range from 80-20 to 89-11 the same rate has increased to the 40%, touching the 43% for the remaining offers from 90-10 to 99-1.

Treatment summary

	INDIVIDUALS		GROUPS	
RANGE	Yes	No	Yes	No
50-41	30	0	10	0
40-31	24	6	8	2
30-21	20	10	7	3
20-11	18	12	7	3
10-1	17	13	7	3

	INDIVIDUALS		GROUPS	
RANGE	Yes	No	Yes	No
50-41	100%	0	100%	0
40-31	80%	20%	80%	20%
30-21	67%	33%	70%	30%
20-11	60%	40%	70%	30%
10-1	57%	43%	70%	30%

This difference shows how the groups are able to rationalize and reduce the reciprocal approach that its components might have, giving the additional data needed to conclude that groups are less than individuals. Hypothesis 2 holds then.

3.4 WISDOM OF THE CROWD

Even if it has been shown so far that groups are better at decision-making than individuals, that doesn't necessarily mean individuals can't make better decisions that groups, this is just less frequent.

There may be cases where groups also make wrong decisions, and this is where "The wisdom of the Crowd" fits. According to the theory of "The wisdom of the Crowd" there are some fields where the collective opinion of a group of people (the "crowd") is generally better than the opinion of an individual expert. These fields are spatial reasoning, quantity estimation and general knowledge.

The "crowd" however doesn't necessarily have to interact with his members to deliberate, for example we can consider the answer of a crowd as of the average or the median of the answers given individually by its members. A first experiment we can quote to show the effectiveness of the crowd is the one done by Jack L. Treynor (1987). In his experiment it has been asked a class of students to guess the number of jelly beans in a jar in front of them. The results showed that, even if no one had previous knowledge to guess perfectly the number, the mean estimate of the answers (841) was really close to the right number (810). This experiment has been proposed another time, giving a little bit more information about the jar and the space between the top of the jar and the beans, still the mean of the answers (871) was really close to the right amount of beans in the jar (850). It is important to recall that in the first experiment two of the 46 participants guessed closer than the mean, and in the second experiment another participant of the 56 guessed closer than the mean of the group; this means, as said before, that there may be cases where individuals make better decisions than groups, but this cases are rare and for the most it is better off just being in groups.

On the one hand being in a crowd and deciding by averaging the answers can help to eliminate the accidental errors in terms of evaluation, on the other hand it doesn't solve the possible cognitive biases that can be accentuated by the crowd. An example for this is the speculative bubble; considering a particular market and simplifying many assumptions we can assume that there are two possible strategies for each player, "sell the share *x*" and "buy the share *x*", the players are the investors, the outcomes are represented by the price of the share. As long as the players are independent and just a small amount of them is biased, then the market works perfectly, but when the players are got by euphoria ("I buy because everyone buys") and then more and more become biased then it is created a bubble that is totally worse off for each player that has chosen the biased strategy.

In order to avoid this kind of inefficiencies we need to individuate what are the qualities or elements crowds need to have in order to be "wiser" than single individuals. James Surowiecki

(2004) defined what are the five elements that make wise crowds stand out from irrational crowds:

- Diversity of opinion. Each member of the crowd should have his own idea, as it was for the experiment of the beans in the jar;

- Independence. It is important that no one is influenced by the opinions of the other players, otherwise it may result in a situation like to one for the share market bubble;

- Decentralization. People specialize in what they like more, this will help for two reasons, first it will help with the diversity of opinion and then "it fosters, and in turn is fed by, specialization" (2004);

- Aggregation. There must be a way to collect individual preferences and decisions in order to individuate the collective element. An example for this is the election polls;

- Trust. Individuals should trust the crowd to be even in his judgement. An example for this is with the public goods. Public goods require a contribution to be set up, but when they are set up they are non-excludable (everyone can use them) and non-rival (the use of it by someone doesn't prevent someone else to use it); if someone freerides that may affect the decisions of the other contributors to pay as well, which will end up in not having that particular public good.

3.5 BEANS-IN-A-JAR EXPERIMENT

I will run again twice the beans-in-a-jar experiment with a little twist to observe also the differences between groups, individuals and crowds, and test 3 hypotheses:

- *Hypothesis 3*: Groups are more independent than individuals;

- *Hypothesis 4*: The outcome of groups decision is closer to the right amount than the individuals;

- *Hypothesis 5*: Lack of diversity of opinion and independence make the crowd unwise.

In the first experiment we will have 52 players picked up randomly at LUISS university. In the first-round individuals will be required to guess the right amount of beans in the jar, but the answers will have to be written on an enumerated piece of paper and then collected. In the second-round starting from the person with the assigned number 1, the players will have to say out loud their guess and write it again on a new piece of paper. In this round, players are allowed to change their guesses, however the amount of beans in the jar will not change.
If individuals are independent in the choice, they are supposed to stick with their previous guess (since no new information has been added), if not then the mean of the answers will change and consequently the guess of the crowd will be different. Moreover, if the concentration of the answers will be more accentuated in the second-round we could say that the diversity of opinion is reduced and as a result see the effect on the crowd result. The same experiment will be repeated again but this time there will be 45 people divided into groups of 3 making decisions, this way differences between individuals and groups can be pointed out; it is important to say that there have been no restrictions on how to give the final guess in both rounds.

3.6 RESULTS OF BEANS-IN-A-JAR EXPERIMENT

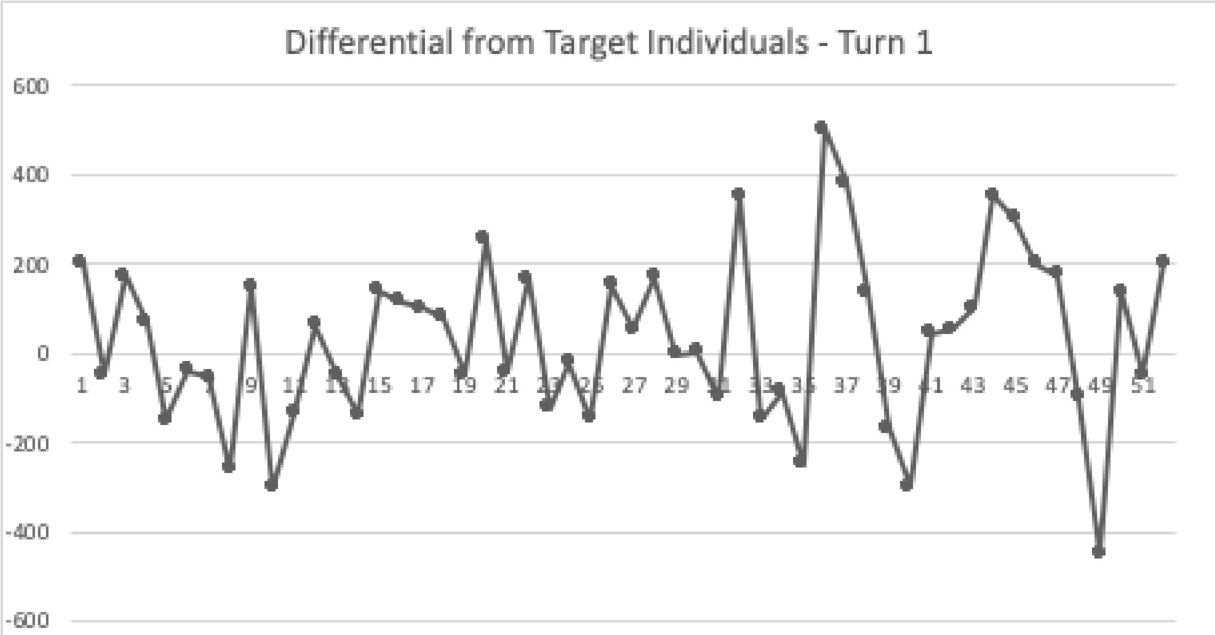
For what concerns the individuals the results show that in the first round the mean of the answers has been 820.73, with a gap to the target (the right amount of beans in the jar, which in this case is 853) of 32.27, while the median has been 803.5. In the second turn, the one where *n* participant have to tell out loud their guess and can change it, also knowing the answers of the *n*-1 players before him, we can observe a new mean of 798.65, a new median of 787, and an average cap from the target of 54.35. It is also interesting to observe that, while in the first round the minimum value of the guesses was 350 and the maximum was 1300, the new minimum value in the second round is 600 and the new maximum is 1000, reducing the widespread of the answers from 950 to 400.

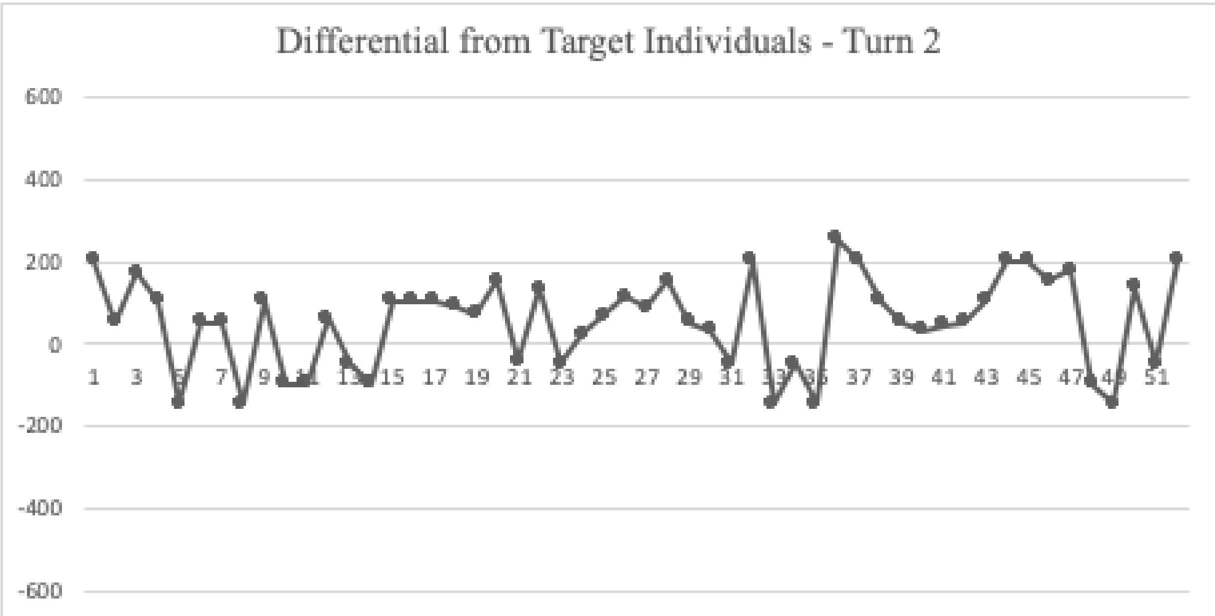
For what concerns the groups we can observe that at the first round the mean of the answers has been 825.13, with a median of 830 and an average discrepancy from the target (812) of 13.13. In the second round we can see that the mean hasn't had a huge variation, the new mean is in fact 825.8, with a new median of 825, and an average gap from the target of 13.8. The ranges in both turns are not large, in the first turn the minimum has been 730 and the maximum 920, with a widespread of 190, while in the second turn the minimum has been 750 and the maximum 910 with a new spread of 160. A first explanation for this can be that groups may have a certain intrinsic mechanism that smooths the out-range values, resulting in a more accurate prediction.

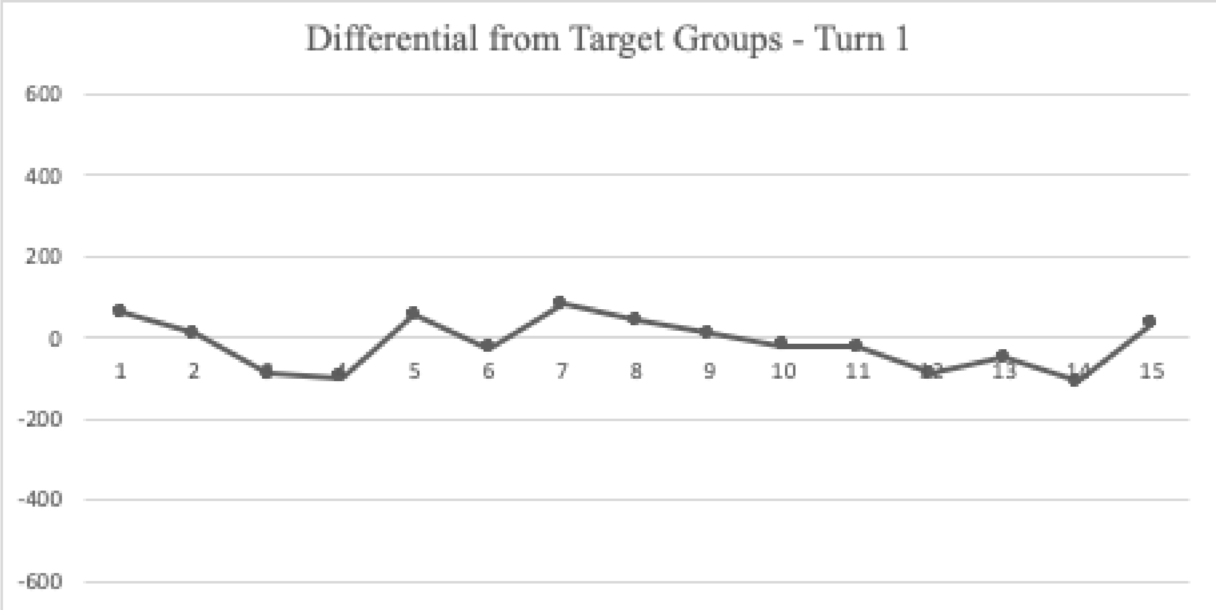
An interesting data for the analysis is the average change from the first and the second turn between individuals and groups; in fact, while individuals had an absolute change in their answers of 22.08, the groups saw a variation of 0.67. This can be a first indication of the possible difference in terms of influenceability between individuals and groups.

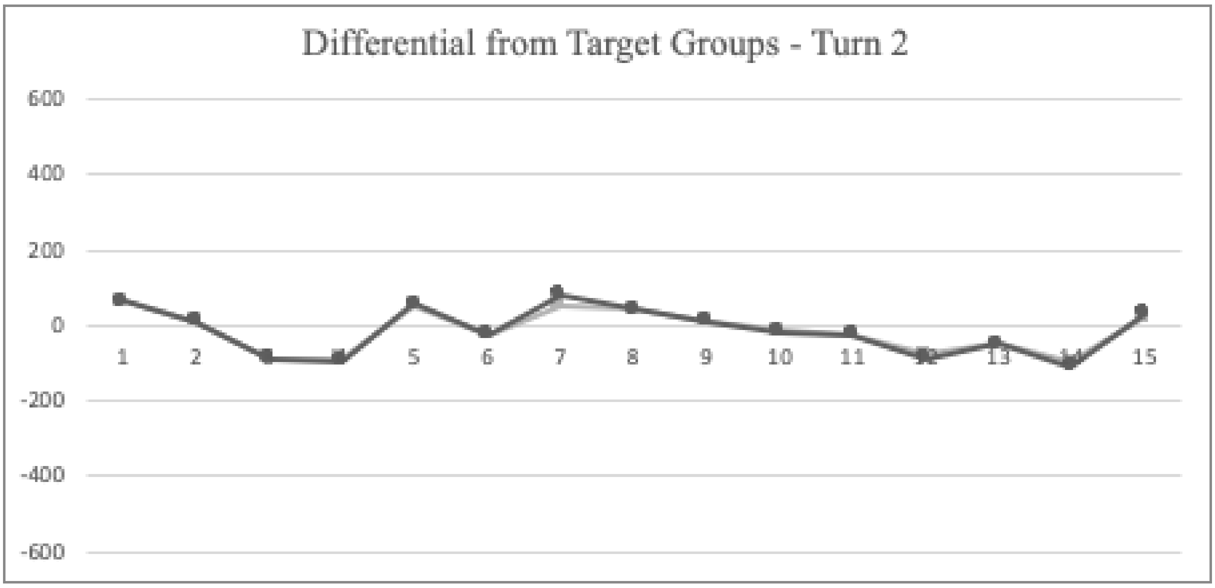
Examination Summary

		INDIVIDUALS	GROUPS
	Participants	52	45 (15 Groups)
	Target	853	812
Guess 1	Mean	820.73	825.13
	Median	803.5	830
	Gap to the target	32.27	13.13
	Minimum	350	730
	Maximum	1300	920
Guess 2	Mean	798.65	825.8
	Median	787	825
	Gap to the target	54.35	13.8
	Minimum	600	750
	Maximum	1000	910
	Change in guesses (Abs)	22.08	0.67


Differential from Target Individuals - Turn 1


Differential from Target Individuals - Turn 2





As we can see also graphically the differential from the target, which may be a good proxy both to see the both differences in terms of accuracy and influenceability, we expect a flatter curve, compared to the first round, in the second round for the macro-group less independent. In fact, we can see that the individuals' curve gets way flatter if compared to the first round, whereas for groups the curve almost doesn't change.

For what concerns diversity of opinion, the more it is present the more the values of the curve will be spread around. As we can see, where for the individuals there are many different values, for the groups the answers have a smaller range, meaning that the diversity of opinion is lacking.

3.7 HYPOTHESIS 3

With the third hypothesis I would like to verify if groups are more independent than individuals, where for more independent I intend that the answers in the second round will change less for groups than for individuals. It is important to recall that an individual *n* can be influenced just by the *n*-1 individuals that told their answer. In fact, both for the first individual and the first group of the list there hasn't been a change in the answer.

In case of influence, I expect that the dispersion in the second turn will be lower. A first indicator for this is the observation of change in the two range between the rounds. For the individuals we can see that, if at the first turn we observe an interval of 950, in the second turn the range is drastically reduced by a 57.89%, while for the groups we see that it has been reduced only by 15.89%, from 190 to 160.

Another examination that can be done is to compare the variance σ^2 between the two turns and then compare the rapport of the two between the two study macro-groups. Since the variance is an index of dispersion, the more the value the more the answers will be dispersed. If during the examination turns there has been no influence, I expect that the value of the variance will not be reduced, or at least the change will be slight.

At the first turn for the individuals we have:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{52}(x_i - 820.73)^2}{52} = 36189.22$$

While at the second we have

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{52}(x_i - 798.65)^2}{52} = 12231.88$$

We can observe that there is an important drop between the two variances, this signifies that there is a dependence between the set of information acquired between the two turns and the answers given.

For what concerns the groups at the first turn the variance is:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{52}(x_i - 825.13)^2}{52} = 3893.55$$

While at the second turn it is:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{52}(x_i - 825.8)^2}{52} = 3123.89$$

Here the reduction in terms of variance is slight, this can be a consequence of a major consistency of the groups with their guesses and the small size of the sample.

Comparing the two rapports of variance between the macro-groups of observation we expect that the rapport of variances will have a smaller number for the more influenced ones. So, if groups are more independent, we will expect a higher rapport of variances between the two turns.

If the inequality below is verified, then the hypothesis 3 holds.

$$\frac{\sigma_{i\ 2}^2}{\sigma_{i\ 1}^2} < \frac{\sigma_{g\ 2}^2}{\sigma_{g\ 1}^2}$$

Where "$\sigma_{i\ 2}^2$" is the individuals' variance in the second round; "$\sigma_{i\ 1}^2$" is the individuals' variance in the first round; "$\sigma_{g\ 2}^2$" is the variance of groups in the second round and "$\sigma_{g\ 1}^2$" is groups' variance in the first round.

$$\frac{12231.88}{36189.22} < \frac{3123.89}{3893.55}$$

$$33.80\% < 80.23\%$$

Since the rapport of variances for individuals is lower than the groups' one, this is another evidence for the hypothesis 3.

Given that we know that the players are rational, in the case of no further information about the game we expect them to stick with their first guess in the second round, and this should be the

same even if they have information about the other players given the fact that they are not competing with each other and then those information are useless for the final outcome.
As said before, we know that the player *n* can be influenced just by the information set of the *n*-1 players before him so it is likely that, if there is an influence, the player *n* will tend to the values of the *n*-1 players before, weighting their guesses. If groups are more independent, then we expect those to stick to their first guesses or at least to "converge" slower to the values of the individuals before.

3.8 HYPOTHESIS 4

With the fourth Hypothesis I would like to verify that the outcome of groups decision is closer to the right amount than the individuals. Since the amount of beans in the jar and the sample changes between groups and individuals I will use again the Tracking Error Volatility.
In this re-adaptation I will use as benchmark the number of beans in the jar and as observations the individuals guesses, and N will be the number of players, computing this for both the first and the second round. The higher the value of the TE will be, the less accurate the answers sample observed will be. So, if the Hypothesis 4 is accurate we will expect individuals to have a higher value for he TE and groups to have lower ones.
Calculating the TE for the individuals guessing in the first round we will have that

$$TE = \sqrt{\frac{\sum_{i=1}^{52}(R_i - 853)^2}{52 - 1}} = 193.01$$

And in the second individuals' round

$$TE = \sqrt{\frac{\sum_{i=1}^{52}(R_i - 853)^2}{52 - 1}} = 123.46$$

As we can see the value of the TE in both cases is quite high, and moreover, we can get evidence for the Hypothesis 3 about the influences between the two turns played, since the TE is reduced in amount when there is a reduction in the variance.
Whereas for what concerns the groups we will have that at the first round the TE is

$$TE = \sqrt{\frac{\sum_{i=1}^{52}(R_i - 812)^2}{15-1}} = 63.86$$

While at the second turn we have

$$TE = \sqrt{\frac{\sum_{i=1}^{52}(R_i - 812)^2}{15-1}} = 57.69$$

Observing the values of TEs we can say that for what concerns the groups their guesses are more accurate in both rounds in respect to the individuals' ones. This will make the Hypothesis that outcome of groups decision is closer to the right amount than the individuals hold.

3.9 HYPOTHESIS 5

With the fifth and last Hypothesis I would like to test if the lack of diversity of opinion and independence make the crowd unwise, where for unwise I intend that the evaluation of the crowd with diversity of opinion and independence will be more accurate.

From the previous hypothesis we know that fort both individuals and groups, even if for the second it is way less accentuated, that in the second round there is less independence and less diversity of opinion. If this hypothesis holds, we would expect the difference between the two turns to be significantly different in statistics terms, because we know that there is an influence between the two turns (otherwise the players would have kept their previous answers), but this difference may not be relevant.

I feel the need to stress the point that the even if the difference between the two turns it is not relevant that doesn't mean that there has been an influence. Since we assume that players are rational if they don't stick with their guesses between the two rounds, even if the difference is not relevant, that means in one of the two turns they had guesses randomly and then the players are not rational, since the information set about the beans in the jar is the same also the answer should be the same. Given that the rationality of the players is a pre-assumption necessary for all the experiment to have sense, that means that there has been an influence.

At a first glance, we can see that for what concerns the individuals there is an average gap to the target between the two turns of 22.08, but since

$$\frac{\bar{x}}{\sigma} = \frac{22.08}{105.43} = 0.21$$

this difference it is not significative, neither at a 10% interval of confidence. So, for what concerns individuals the hypothesis 5 does not hold.

For what concerns the groups, we would expect this effect to be even smaller, since the mean of the differences between the two turns is 0.67, then

$$\frac{\bar{x}}{\sigma} = \frac{0.67}{9.67} = 0.07$$

Even this rapport is not significative at any level of confidence.

We can conclude saying that, for both individuals and groups, the lack of diversity and independence is not significantly influencing the final outcome. However, these results can be due to the small size of the samples and then a bigger group could potentially lead to different results.

CONCLUSION AND OTHER CONSIDERATIONS

To review we can see that the differences between individuals and groups are quite marked and pretty consistent among the treatments. Groups tend to be more accurate, more independent, more rational and less reciprocal than individuals, and I think one explanation for it is an internal mechanism that let the group smooth and eliminate the outrange values or decision that each member have.

However, when decision making is done by individuals there are some benefits, the first and most important is that decision made by individuals tend to create a higher welfare; moreover, individuals are more other caring, and work towards reduction of differences.

However, since individuals are more reciprocal than groups, the positive effects that we would have without a "misbehave" of other players in terms of total welfare may receive a setback, since the final aim of the individual will be to punish the misbehaving actor, no matter if the individual will have a loss too. The direct consequence is that it becomes quite easy to decide whether to implement an individual-based or a group-based decision process, since the two process can have various and different outcome but still quite consistent.

Further studies to go deeper into the differences between individuals and groups are required, since the answer to the question "which is the better one between the two?" is not so easy to give and, most of the times, as shown here, it depends on the context and various external factors. And also, the applications for this study in the economic field are many and always actual, considering the fact that decision making is a day-to-day element.

Bibliography

Banjeree, A. V. (1993). A simple model of herd behavior. *Quarterly Journal of Economics*, 797-817.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 122-142.

Bewley, T. (1995). A Depressed Labor Market s Explained by Participants. *American Economic Review*, 250-254.

Bornstein, G., Gneezy, U., & Nagel, R. (2002, May 4). The effect of intergroup competition on group coordination: an experimental study. *Games and Economic Behavior*, 1-25. Retrieved from Academic Press: www.academicpress.com

Bornstein, G., Kugler, T., & Ziegelmeyerb, A. (2004). Individual and group decisions in the centipede game: Are groups more ‘‘rational’’ players? *Journal of Experimental Social Psychology*, 599-605.

Charness, G., & Rabin, M. (2002). Understanding Social Preferences with simple tests. *The Quarterly Journal of Economics*, 817-869.

Charness, G., & Sutter, M. (2012). Groups Make Better Self-Interested Decisions. *Journal of Economic Perspectives*, 157–176.

Charness, G., Karni, E., & Levin, D. (2006). *Individual and Group Decision Making Under Risk: An Experimental Study of Bayesian Updating and Violations of First-order Stochastic Dominance.*

Charness, G., Rigotti, L., & Rustichini, A. (2007). Individual Behavior and Group Membership. *American Economic Review*, 97(4): 1340–52.

Cooper, R., DeJong, D., Forsythe, R., & Ross, T. (1990). Selection criteria in coordination games: Some experimental results. *American Economic Review*, 218-233.

Falk, A., & Fischbacher, U. (2000). A theory of reciprocity. *A theory of reciprocity*. Zürich, Switzerland: University of Zurich.

Falk, A., & Ichino, A. (2006). Clean Evidence on Peer Effects. *Journal of Labor Economics*, 24(1): 39–57.

Falk, A., Fehr, E., & Fischbacher, U. (2003). ON THE NATURE OF FAIR BEHAVIOR. *Economic Inquiry*, 20-26.

Fehr, E., & Gächter, S. (2000). *Do Incentive Contracts Crowd-Out Voluntaiy Cooperation?* Zurich: Institute for Empirical Research in Economics.

Fehr, E., & Gätcher, S. (2000). Cooperation and Punishment in Public Goods Experiments. *THE AMERICAN ECONOMIC REVIEW*, 980-994.

Fehr, E., & Schmidt, K. (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 817-868.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). STRONG RECIPROCITY, HUMAN COOPERATION, AND THE ENFORCEMENT OF SOCIAL NORMS. *Human Nature*, Vol. 13, No. 1, pp. 1-25. .

Fehr, E., Gachter, S., & Kirchsteiger, G. (1997). Reciprocity as a Conctract Enforcement Device. *Econometrica*, 833-860.

Gigerenzer, G., Todd, P. M., & ABC, G. (1999). Simple heuristics that make us smart. *Oxford University Press*, 97-118.

Kocher, M., & Sutter, M. (2005). The Decision Maker Matters: Individual versus Group Behavior in Experimental Beauty-Contest Games. *Economic Journal*, 200–223.

Koriat, A. (2012). When Are Two Heads Better than One and Why? *Science*, 360-360.

Kugler, T., Bornstein, G., Kocher, M., & Sutter, M. (2007). Trust between Individuals and Groups: Groups are Less Trusting Than Individuals But Just as Trustworthy. *Journal of Economic Psychology*, 28(6): 646–57.

March, J. (1994). *A Primer on Decision Making: How Decisions Happen.* New York: The Free Press.

Marengo, L., Alexia, D., & Ploner, M. (2016). I did it your way. An experimental investigation of peer effects in investment choices. *Journal of Economic Psychology*, 113-123.

McKinsey. (2009). *How companies make good decisions: McKinsey Global Survey Results.* McKinsey & Company.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 1281-1302.

Simon, H. A. (1955). A Behavioral model of rational choice. *The Quarterly Journal of Economics*, Vol. 69, No. 1 (Feb., 1955), pp. 99-118.

Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations.* United States: Doubleday; Anchor .

Treynor, J. (1987). Market Efficiency and the Bean Jar Experiment. *Financial Analysts Journal*, 50-53.

Van Huyck, J., Battalio, R., & Beil, R. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, 234-248.

Watson, J. (2013). *Strategy.* W. W. Norton & Company.

Watson, J. (2013). *Strategy, third edition.* W. W. Norton & Company, Inc.

LIBERA UNIVERSITA' INTERNAZIONALE DEGLI STUDI SOCIALI

Department of Business and Management

Chair of Managerial Decision Making

SUMMARY

Differences in the decision-making process between Individuals and Groups

SUPERVISOR

Prof. Luigi Marengo

CANDIDATE

Lorenzo Simoni

ID Number: 697451

CO-SUPERVISOR

Prof. Luca Panaccione

Table of Contents

INTRODUCTION

In this thesis I will analyze the differences in terms of behavior between individuals and groups when it comes to decision making. To do so I will use Game theory concepts and axioms, since with game-theoretical decisions there is generally no space for individual interpretation of the right answer, it is either right or wrong, making the analytic process the most impartial possible. The first part then will be focused on explaining and assembling the Game-theory postulates, such as the definition of rationality and the explanation of utility functions, in order to have a full understanding of the treatments in the following chapters.

The areas of analysis will be on differences in:

- Rational behavior;
- Reciprocal behavior;
- Independence in decision making;
- Outcomes of decision making.

Moreover, at the beginning of each demonstration there will be a brief introduction to summarize and show the previous experiments and conclusions already made by the academy in the same or similar fields. For what concerns the groups it has to be said that the academy has started to investigate their behavior just recently and that's the reason why many things that might seem obvious had to be tested in order to make the right comparisons and get data for the comparison.

The experiments done and proposed are three, the first is a revised centipede, the second is a classic ultimatum game (those will be really relevant to show differences in rationality and reciprocal behaviors), and the third one is a guessing game, where participants have to guess the right amount of beans in a jar; this will show the differences in terms of independency (intended as the capability to stick to own ideas) in decision making and the accuracy of those decisions.

All the results show with a certain consistency that the groups are less influenceable and more precise than individuals. The explanation for this may be the fact that the groups have an intrinsic mechanism that smooth the components ideas and then it makes the final decision more accurate, whereas the individuals have no one to compare their ideas with. However, making decisions in groups has its own downsides. In fact, it will be shown that, even if the outcome of the groups is consistently bigger that the individual's outcome, the overall welfare is decreased.

THE DECREASING CENTIPEDE EXPERIMENT

To provide further evidence regarding differences in terms of rationality between individuals and groups and observe also differences in reciprocity levels, I have set up a sequential game, which is a revision of the classic centipede game. In this revised centipede there are two competitors who have to allocate an amount of €100 in a sequential game and I will test if

- *Hypothesis 1*: Groups are more rational than individuals.

- *Hypothesis 2*: Groups are less reciprocal than individuals.

The first-mover can propose to split the amount in any way he prefers, if the other player agrees then the game will end, otherwise the second player will make a counteroffer, this process will go on until a player accepts or until the fourth proposal. Each time that a player refuses the amount of money will be decreased by €20, and the last player to make an offer will be the one that was the respondent at the first stage. At the fourth proposal if the player 1 refuses both players will get 0 (Figure 4).

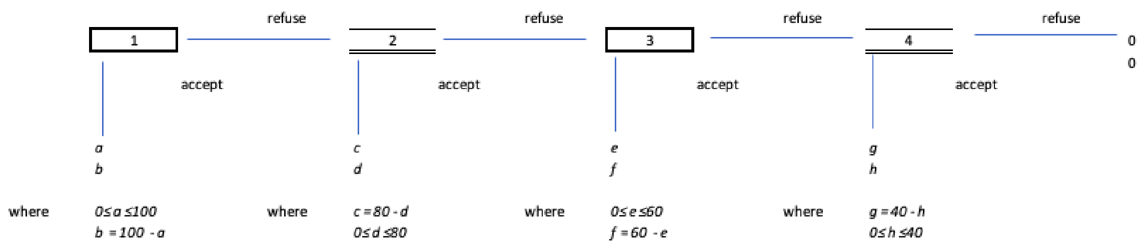


Figure 4

If we observe the formalization of this game, as shown in Figure 4, we see that the numbers with a thick box border represent the nodes where player 1 has to propose the split to player 2, whereas the numbers with a top and double bottom border represents the nodes where the player 2 has to make a proposal; in each node the outcomes on top represents the outcomes of player 1.

As for the other centipede games, this game can be easily solved through a backward induction reasoning, and the formalization without variables for the optimal solution will be as shown in Figure 5.

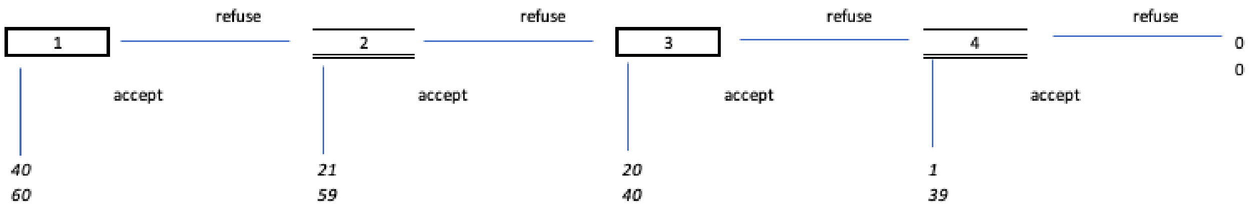


Figure 5

Since player 2 knows that if player 1 refuses his offer they will both get zero, he will offer him 1 and will take the rest. At the previous node then player 1 will be incentivized to offer player 2 an amount slightly higher than the one in the fourth node, then he will offer him 40 and propose to take the rest. This process will continue until the first node where the optimal strategy will be for player 1 to offer 60 to player 2 and take 40 for his own.

The sample for this experiment is composed by 80 persons, of which 20 will play as individuals and the others will be aggregate in groups of three randomly. To avoid spillover effects no one knew who his opponent was, the game has been proposed and given just in the parametrical form, and no time limits have been imposed. Offers and counteroffers have been taken and proposed by electronical devices not to influence the choices.

I decided to split the sample into three different macro groups, in the first group there were 7 individuals competing with other 7 individuals, the second was composed by 6 individuals playing against 6 groups, whilst the last group was composed by 7 groups against other 7 ones. The main difference with this experiment from the one conducted by G. Bornstein, T. Kugler, and A. Ziegelmeyerb (2004) is that through this centipede there may be the possibility to individuate the reciprocity effect. If expectations are violated in the split of the pile of money, we will expect the players to reciprocate, trying to drag down the final outcomes.

The results actually show that for many players, especially when it comes to individual players against each other, it is totally fine and acceptable to split into even parts the pile of money even if it is not the best strategy.

The table below shows the node where each macro-group stopped. In "individuals vs groups" are expressed the values with individuals as first proposers.

	1	2	3	4	no agreement
Individual vs individual	4	0	1	1	1
Individual vs group	2	1	0	0	2
group vs individual	1	1	0	0	0
group vs group	6	0	0	1	0

During the experiment the individuals has been allowed to decide the first split 12 times, and the 67% of those times they have proposed to split even the pile of money. The average of the amount proposed by the individuals in the first turn has been 54.92-45.08, but most importantly it has a gap from the best move of 14.92; while the groups has decided first the split for 9 times, and the average proposed has been 46.1-53.9, with a gap from the optimal of 6.1.

HYPOTHESIS 1

This first data is consistent with the literature for what concerns the individuals, as also showed by the work of Falk et al. (2003). The results however show an interesting result for what concerns the groups. To analyze this difference and test the first hypothesis I will re-adapt a volatility measure utilized by hedge-funds, called Tracking Error Volatility, to observe the standard deviation percentage difference of a portfolio in respect to a benchmark portfolio the fund is trying to imitate.

$$TE = \sqrt{\frac{\sum_{i=1}^{n}(R_i - R_B)^2}{N-1}} = \sigma(R_i - R_B)$$

Where TE is the Tracking Error, R_i is the fund's return of the i^{th} portfolio, R_B is the return of the benchmark and N is the number of periods.

In this case the benchmark will be the game-theory right answer and as observations the proposed splits, N will be the number of players. The higher the value of the TE will be, the less accurate the answers sample observed. So, if the Hypothesis 1 is accurate we will expect individuals to have a higher value for he TE and groups to have lower ones. This analysis will be done just for the first round, since for the others the number of observations would be too small to capture satisfying data.

For the individuals proposing at the first round we have that the Tracking Error Volatility is

$$TE = \sqrt{\frac{\sum_{i=1}^{12}(R_i - 40)^2}{12 - 1}} = 21.49$$

Whereas for groups proposing at the first round, we have that

$$TE = \sqrt{\frac{\sum_{i=1}^{9}(R_i - 40)^2}{9 - 1}} = 8.45$$

As said before we were expecting the less rational macro-group to have a higher TE compared to the more rational. Since the individuals' TE is higher than the groups' one, hypothesis 1 holds.

For what concerns the second hypothesis that groups are less reciprocal than individuals, there are a little more problems to investigate it. As a starting point we may say that the acceptancy rate for individuals in the first round has been 55.6%, with an average proposed of 49, whereas the average rejected as the first round has been 37.75. For the groups in the first round the acceptancy rate has been 66.67%, with an average proposed of 54.25, on the other hand there has been an average of 47.5 rejected. From this first data we can observe that whereas the individuals have had a behavior predictable in a certain way for what concerns reciprocity, the fact that groups have rejected an average amount of 47.5 is a first clue to prove hypothesis 2, and a further proof for hypothesis 1. In fact, from a classical game of theory point of view rejecting an average amount of 47.5 is totally right, but from the studies done by Falk et al. (2003) we know that generally 50-50 offers, or close to them, are almost never rejected by reciprocal players.

Another datum interesting to observe is the one relative to the fourth node. While individuals have been 4 times in the spot to decide whether to accept the offer or end the game with no rewards, the groups have been in this spot just once. For what concerns the individuals as responders the three splits rejected have been the 1-39 splits, even if it is a suboptimal solution (the first number is the one that would be the reward for the respondent player); whereas the accepted offer has been 10-30. For what concerns the groups the only offer made at the fourth node has been a 1-39, accepted by the respondent.

From these data we can say that Hypothesis 2 holds, and then groups are less reciprocal than individuals. This may due to the fact that the groups in a certain way mitigate the reciprocal attitude of each component. However, it is important to say that even if this hypothesis holds, it may be due the small sample and then further studies with a bigger sample are required.

HYPOTHESIS 2

Since the data from the decreasing centipede experiment were insufficient to further test the hypothesis that groups are less reciprocal than individuals, an ultimatum game experiment will be conducted to analyze the behavior of groups.

Since the analysis is on reciprocity and not on rationality, the groups will not be required to propose a split of a 100€ amount, they will just have to decide whether to accept or not the offer.

The splits proposed will be plotted in order to analyze the behavior in each situation. Since presumably the splits above 50-50 will always be accepted and realistically will almost never be proposed, there will be 50 proposal to 50 groups of 3 people, starting from an even split, arriving to a 99-1 proposal, where 1 will be the reward of each member of the group.

To have more data, before discussing with the group whether to accept, each group member will be required to write on a piece of paper if they would accept the split, this will let acquire more data to compare groups' and individuals' behavior.

No instructions on how to give a final answer on the response have been given to groups, in order to leave the most possible freedom. Moreover, the participants didn't know the proposes were already plotted, and they have been told to play against other players not to false their behavior.

Since the rationality of the groups has already been demonstrated, we will expect now a smaller amount in terms of percentage of groups rejecting the proposals compared to the individuals. Realistically this difference can be due to the fact that groups, as it was for the beans-in-a-jar experiment, tend to smooth the individuals' choices, so more reciprocal individuals will be needed to have a reciprocal group.

The results show that, as expected, in the splits from 50-50 to 59-41 neither individuals nor groups rejected the offer. In the splits from 60-40 to 69-31 there has been the same rates for groups and individuals, equal to 80% acceptancy rate and 20% rejection rate. At this point there is the divergence between individuals and groups. Whereas the groups keep the same acceptancy-rejection rate constant and equal to 70%-30%, individuals have a growing rejection rate the more the stake proposed is lower. In fact, in the range from 70-30 to 79-21 the rejection rate has been the 33%, in the range from 80-20 to 89-11 the same rate has increased to the 40%, touching the 43% for the remaining offers from 90-10 to 99-1.

Treatment summary

	INDIVIDUALS		GROUPS	
RANGE	Yes	No	Yes	No
50-41	30	0	10	0
40-31	24	6	8	2
30-21	20	10	7	3
20-11	18	12	7	3
10-1	17	13	7	3

	INDIVIDUALS		GROUPS	
RANGE	Yes	No	Yes	No
50-41	100%	0	100%	0
40-31	80%	20%	80%	20%
30-21	67%	33%	70%	30%
20-11	60%	40%	70%	30%
10-1	57%	43%	70%	30%

This difference shows how the groups are able to rationalize and reduce the reciprocal approach that its components might have, giving the additional data needed to conclude that groups are less than individuals. Hypothesis 2 holds then.

BEANS-IN-A-JAR EXPERIMENT

I will run again twice the beans-in-a-jar experiment with a little twist to observe also the differences between groups, individuals and crowds, and test 3 hypotheses:

- *Hypothesis 3*: Groups are more independent than individuals;

- *Hypothesis 4*: The outcome of groups decision is closer to the right amount than the individuals;

- *Hypothesis 5*: Lack of diversity of opinion and independence make the crowd unwise.

In the first experiment we will have 52 players picked up randomly at LUISS university. In the first-round individuals will be required to guess the right amount of beans in the jar, but the answers will have to be written on an enumerated piece of paper and then collected. In the second-round starting from the person with the assigned number 1, the players will have to say out loud their guess and write it again on a new piece of paper. In this round, players are allowed to change their guesses, however the amount of beans in the jar will not change.
If individuals are independent in the choice, they are supposed to stick with their previous guess (since no new information has been added), if not then the mean of the answers will change and consequently the guess of the crowd will be different. Moreover, if the concentration of the answers will be more accentuated in the second-round we could say that the diversity of opinion is reduced and as a result see the effect on the crowd result. The same experiment will be repeated again but this time there will be 45 people divided into groups of 3 making decisions, this way differences between individuals and groups can be pointed out; it is important to say that there have been no restrictions on how to give the final guess in both rounds.

RESULTS OF BEANS-IN-A-JAR EXPERIMENT

For what concerns the individuals the results show that in the first round the mean of the answers has been 820.73, with a gap to the target (the right amount of beans in the jar, which in this case is 853) of 32.27, while the median has been 803.5. In the second turn, the one where *n* participant have to tell out loud their guess and can change it, also knowing the answers of the *n*-1 players before him, we can observe a new mean of 798.65, a new median of 787, and an average cap from the target of 54.35. It is also interesting to observe that, while in the first round the minimum value of the guesses was 350 and the maximum was 1300, the new minimum value in the second round is 600 and the new maximum is 1000, reducing the widespread of the answers from 950 to 400.

For what concerns the groups we can observe that at the first round the mean of the answers has been 825.13, with a median of 830 and an average discrepancy from the target (812) of 13.13. In the second round we can see that the mean hasn't had a huge variation, the new mean is in fact 825.8, with a new median of 825, and an average gap from the target of 13.8. The ranges in both turns are not large, in the first turn the minimum has been 730 and the maximum 920, with a widespread of 190, while in the second turn the minimum has been 750 and the maximum 910 with a new spread of 160. A first explanation for this can be that groups may have a certain intrinsic mechanism that smooths the out-range values, resulting in a more accurate prediction.

An interesting data for the analysis is the average change from the first and the second turn between individuals and groups; in fact, while individuals had an absolute change in their answers of 22.08, the groups saw a variation of 0.67. This can be a first indication of the possible difference in terms of influenceability between individuals and groups.

For what concerns diversity of opinion, the more it is present the more the values of the curve will be spread around. In fact, where for the individuals there are many different values, for the groups the answers have a smaller range, meaning that the diversity of opinion is lacking.

HYPOTHESIS 3

With the third hypothesis I would like to verify if groups are more independent than individuals, where for more independent I intend that the answers in the second round will change less for groups than for individuals. It is important to recall that an individual *n* can be influenced just by the *n*-1 individuals that told their answer. In fact, both for the first individual and the first group of the list there hasn't been a change in the answer.

In case of influence, I expect that the dispersion in the second turn will be lower. A first indicator for this is the observation of change in the two range between the rounds. For the individuals we can see that, if at the first turn we observe an interval of 950, in the second turn the range is drastically reduced by a 57.89%, while for the groups we see that it has been reduced only by 15.89%, from 190 to 160.

Another examination that can be done is to compare the variance σ^2 between the two turns and then compare the rapport of the two between the two study macro-groups. Since the variance is an index of dispersion, the more the value the more the answers will be dispersed. If during the examination turns there has been no influence, I expect that the value of the variance will not be reduced, or at least the change will be slight.

At the first turn for the individuals we have:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{52}(x_i - 820.73)^2}{52} = 36189.22$$

While at the second we have

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{52}(x_i - 798.65)^2}{52} = 12231.88$$

We can observe that there is an important drop between the two variances, this signifies that there is a dependence between the set of information acquired between the two turns and the answers given.

For what concerns the groups at the first turn the variance is:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{52}(x_i - 825.13)^2}{52} = 3893.55$$

While at the second turn it is:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{52}(x_i - 825.8)^2}{52} = 3123.89$$

Here the reduction in terms of variance is slight, this can be a consequence of a major consistency of the groups with their guesses and the small size of the sample.

Comparing the two rapports of variance between the macro-groups of observation we expect that the rapport of variances will have a smaller number for the more influenced ones. So, if groups are more independent, we will expect a higher rapport of variances between the two turns.

If the inequality below is verified, then the hypothesis 3 holds.

$$\frac{\sigma_{i\,2}^2}{\sigma_{i\,1}^2} < \frac{\sigma_{g\,2}^2}{\sigma_{g\,1}^2}$$

Where "$\sigma_{i\,2}^2$" is the individuals' variance in the second round; "$\sigma_{i\,1}^2$" is the individuals' variance in the first round; "$\sigma_{g\,2}^2$" is the variance of groups in the second round and "$\sigma_{g\,1}^2$" is groups' variance in the first round.

$$\frac{12231.88}{36189.22} < \frac{3123.89}{3893.55}$$

$$33.80\% < 80.23\%$$

Since the rapport of variances for individuals is lower than the groups' one, this is another evidence for the hypothesis 3.

Given that we know that the players are rational, in the case of no further information about the game we expect them to stick with their first guess in the second round, and this should be the same even if they have information about the other players given the fact that they are not competing with each other and then those information are useless for the final outcome.

As said before, we know that the player *n* can be influenced just by the information set of the *n*-1 players before him so it is likely that, if there is an influence, the player *n* will tend to the values of the *n*-1 players before, weighting their guesses. If groups are more independent, then we expect those to stick to their first guesses or at least to "converge" slower to the values of the individuals before.

HYPOTHESIS 4

With the fourth Hypothesis I would like to verify that the outcome of groups decision is closer to the right amount than the individuals. Since the amount of beans in the jar and the sample changes between groups and individuals I will use again the Tracking Error Volatility.

In this re-adaptation I will use as benchmark the number of beans in the jar and as observations the individuals guesses, and N will be the number of players, computing this for both the first and the second round. The higher the value of the TE will be, the less accurate the answers sample observed will be. So, if the Hypothesis 4 is accurate we will expect individuals to have a higher value for he TE and groups to have lower ones.

Calculating the TE for the individuals guessing in the first round we will have that

$$TE = \sqrt{\frac{\sum_{i=1}^{52}(R_i - 853)^2}{52 - 1}} = 193.01$$

And in the second individuals' round

$$TE = \sqrt{\frac{\sum_{i=1}^{52}(R_i - 853)^2}{52 - 1}} = 123.46$$

As we can see the value of the TE in both cases is quite high, and moreover, we can get evidence for the Hypothesis 3 about the influences between the two turns played, since the TE is reduced in amount when there is a reduction in the variance.

Whereas for what concerns the groups we will have that at the first round the TE is

$$TE = \sqrt{\frac{\sum_{i=1}^{52}(R_i - 812)^2}{15 - 1}} = 63.86$$

While at the second turn we have

$$TE = \sqrt{\frac{\sum_{i=1}^{52}(R_i - 812)^2}{15 - 1}} = 57.69$$

Observing the values of TEs we can say that for what concerns the groups their guesses are more accurate in both rounds in respect to the individuals' ones. This will make the Hypothesis that outcome of groups decision is closer to the right amount than the individuals hold.

HYPOTHESIS 5

With the fifth and last Hypothesis I would like to test if the lack of diversity of opinion and independence make the crowd unwise, where for unwise I intend that the evaluation of the crowd with diversity of opinion and independence will be more accurate.

From the previous hypothesis we know that fort both individuals and groups, even if for the second it is way less accentuated, that in the second round there is less independence and less diversity of opinion. If this hypothesis holds, we would expect the difference between the two turns to be significantly different in statistics terms, because we know that there is an influence between the two turns (otherwise the players would have kept their previous answers), but this difference may not be relevant.

I feel the need to stress the point that the even if the difference between the two turns it is not relevant that doesn't mean that there has been an influence. Since we assume that players are rational if they don't stick with their guesses between the two rounds, even if the difference is not relevant, that means in one of the two turns they had guesses randomly and then the players are not rational, since the information set about the beans in the jar is the same also the answer should be the same. Given that the rationality of the players is a pre-assumption necessary for all the experiment to have sense, that means that there has been an influence.

At a first glance, we can see that for what concerns the individuals there is an average gap to the target between the two turns of 22.08, but since

$$\frac{\bar{x}}{\sigma} = \frac{22.08}{105.43} = 0.21$$

this difference it is not significative, neither at a 10% interval of confidence. So, for what concerns individuals the hypothesis 5 does not hold.

For what concerns the groups, we would expect this effect to be even smaller, since the mean of the differences between the two turns is 0.67, then

$$\frac{\bar{x}}{\sigma} = \frac{0.67}{9.67} = 0.07$$

Even this rapport is not significative at any level of confidence.

We can conclude saying that, for both individuals and groups, the lack of diversity and independence is not significantly influencing the final outcome. However, these results can be due to the small size of the samples and then a bigger group could potentially lead to different results.

CONCLUSION AND OTHER CONSIDERATIONS

To review we can see that the differences between individuals and groups are quite marked and pretty consistent among the treatments. Groups tend to be more accurate, more independent, more rational and less reciprocal than individuals, and I think one explanation for it is an internal mechanism that let the group smooth and eliminate the outrange values or decision that each member have.

However, when decision making is done by individuals there are some benefits, the first and most important is that decision made by individuals tend to create a higher welfare; moreover, individuals are more other caring, and work towards reduction of differences.

However, since individuals are more reciprocal than groups, the positive effects that we would have without a "misbehave" of other players in terms of total welfare may receive a setback, since the final aim of the individual will be to punish the misbehaving actor, no matter if the individual will have a loss too. The direct consequence is that it becomes quite easy to decide whether to implement an individual-based or a group-based decision process, since the two process can have various and different outcome but still quite consistent.

Further studies to go deeper into the differences between individuals and groups are required, since the answer to the question "which is the better one between the two?" is not so easy to give and, most of the times, as shown here, it depends on the context and various external factors. And also, the applications for this study in the economic field are many and always actual, considering the fact that decision making is a day-to-day element.

Bibliography

Banjeree, A. V. (1993). A simple model of herd behavior. *Quarterly Journal of Economics*, 797-817.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 122-142.

Bewley, T. (1995). A Depressed Labor Market s Explained by Participants. *American Economic Review*, 250-254.

Bornstein, G., Gneezy, U., & Nagel, R. (2002, May 4). The effect of intergroup competition on group coordination: an experimental study. *Games and Economic Behavior*, 1-25. Retrieved from Academic Press: www.academicpress.com

Bornstein, G., Kugler, T., & Ziegelmeyerb, A. (2004). Individual and group decisions in the centipede game: Are groups more ‘‘rational’’ players? *Journal of Experimental Social Psychology*, 599-605.

Charness, G., & Rabin, M. (2002). Understanding Social Preferences with simple tests. *The Quarterly Journal of Economics*, 817-869.

Charness, G., & Sutter, M. (2012). Groups Make Better Self-Interested Decisions. *Journal of Economic Perspectives*, 157–176.

Charness, G., Karni, E., & Levin, D. (2006). *Individual and Group Decision Making Under Risk: An Experimental Study of Bayesian Updating and Violations of First-order Stochastic Dominance.*

Charness, G., Rigotti, L., & Rustichini, A. (2007). Individual Behavior and Group Membership. *American Economic Review*, 97(4): 1340–52.

Cooper, R., DeJong, D., Forsythe, R., & Ross, T. (1990). Selection criteria in coordination games: Some experimental results. *American Economic Review*, 218-233.

Falk, A., & Fischbacher, U. (2000). A theory of reciprocity. *A theory of reciprocity*. Zürich, Switzerland: University of Zurich.

Falk, A., & Ichino, A. (2006). Clean Evidence on Peer Effects. *Journal of Labor Economics*, 24(1): 39–57.

Falk, A., Fehr, E., & Fischbacher, U. (2003). ON THE NATURE OF FAIR BEHAVIOR. *Economic Inquiry*, 20-26.

Fehr, E., & Gächter, S. (2000). *Do Incentive Contracts Crowd-Out Voluntaiy Cooperation?* Zurich: Institute for Empirical Research in Economics.

Fehr, E., & Gätcher, S. (2000). Cooperation and Punishment in Public Goods Experiments. *THE AMERICAN ECONOMIC REVIEW*, 980-994.

Fehr, E., & Schmidt, K. (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 817-868.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). STRONG RECIPROCITY, HUMAN COOPERATION, AND THE ENFORCEMENT OF SOCIAL NORMS. *Human Nature*, Vol. 13, No. 1, pp. 1-25. .

Fehr, E., Gachter, S., & Kirchsteiger, G. (1997). Reciprocity as a Conctract Enforcement Device. *Econometrica*, 833-860.

Gigerenzer, G., Todd, P. M., & ABC, G. (1999). Simple heuristics that make us smart. *Oxford University Press*, 97-118.

Kocher, M., & Sutter, M. (2005). The Decision Maker Matters: Individual versus Group Behavior in Experimental Beauty-Contest Games. *Economic Journal*, 200–223.

Koriat, A. (2012). When Are Two Heads Better than One and Why? *Science*, 360-360.

Kugler, T., Bornstein, G., Kocher, M., & Sutter, M. (2007). Trust between Individuals and Groups: Groups are Less Trusting Than Individuals But Just as Trustworthy. *Journal of Economic Psychology*, 28(6): 646–57.

March, J. (1994). *A Primer on Decision Making: How Decisions Happen.* New York: The Free Press.

Marengo, L., Alexia, D., & Ploner, M. (2016). I did it your way. An experimental investigation of peer effects in investment choices. *Journal of Economic Psychology*, 113-123.

McKinsey. (2009). *How companies make good decisions: McKinsey Global Survey Results.* McKinsey & Company.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 1281-1302.

Simon, H. A. (1955). A Behavioral model of rational choice. *The Quarterly Journal of Economics*, Vol. 69, No. 1 (Feb., 1955), pp. 99-118.

Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations.* United States: Doubleday; Anchor .

Treynor, J. (1987). Market Efficiency and the Bean Jar Experiment. *Financial Analysts Journal*, 50-53.

Van Huyck, J., Battalio, R., & Beil, R. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, 234-248.

Watson, J. (2013). *Strategy.* W. W. Norton & Company.

Watson, J. (2013). *Strategy, third edition.* W. W. Norton & Company, Inc.