# To Treat Others as They Treat You, or to Treat Them as You Treat Yourself?

## A Comparison between two Moral Principles

**Ramzi suleiman**
**Department of Psychology**
**University of Haifa**

Please address all correspondence to Dr. Ramzi Suleiman, University of Haifa, Haifa. 31509, Israel; Email: suleiman@psy.haifa.ac.il, Homepage: http://suleiman.haifa.ac.il
Mobiles: + 972-(0)50-474- 215 and +31-(0)6-816-4553

# To Treat Others as They Treat You, or to Treat Them as You Treat Yourself? - A Comparison between two Moral Principles


## Abstract

The present article compares two fundamental moral principles: The ***Reciprocity Principle***, known also as Confucius's Golden Rule, stating "***treat others as you want them to treat you***", and a core principle in some monotheistic religions, stating "***treat others as you treat yourself***", referred to here as the ***Self-Referencing principle***. Under the rationality assumption I derive predictions of the two principles for interactions involving the division of economic benefits. For asymmetric interactions which could be modeled as Ultimatum games I show that while the reciprocity principle prediction is inconclusive and depends on the individuals' idiosyncratic preferences and social values, the self-referencing principle yields a point-prediction prescribing a fair allocation at the ***Golden Ratio***, according to which the proposer should keep about 62% of the total amount and offer about 38% to the responder. Interestingly, the solution derived from the self-referencing rule is identical to the theoretical result derived from a recently proposed economic theory of cooperation between egoists, called "Economic Harmony Theory" (Suleiman, 2013a). Comparison between the two principles on various dimensions, including the levels of fairness, stability and efficiency in utilizing the common resource, reveals that on all the investigated dimensions the self-referencing principle is superior to the reciprocity principle. Nonetheless, social evolution has apparently selected the reciprocity principle. I argue that the reason for this anomaly stems from the fact that reciprocity includes an option for punishment, which the responder might resort to if treated unfairly. In contrast, the self-referencing principle lacks a similar option. I propose two solution for this anomaly: (1) the establishment of a *social contract*, according to which the majority of individuals agree on a "mutual coercion mutually agreed upon" norm (Hardin, 1968), and (2) the establishment of a central authority, or a ***Leviathan*** (Hobbes, 1651), which has the legitimacy and power to punish uncooperative individuals who defect and free-ride on others' efforts. In modern societies the second solution could be achieved through the establishment and consolidation of a ***Rule of Law***, under which ***all*** people are equal.



**Keywords**: Reciprocity, Golden rule, Kant's Categorical Imperative, Golden ratio, Religion, Ultimatum game, Fairness, Equality, Equity. Game Theory, Harmony, Leviathan, Tragedy of the Commons.

## 1. The Reciprocity Principle – "Treat others as they treat you"

The Reciprocity Principle has its origins in ancient philosophy and religious teachings. Its first formulation is attributed to the Chinese philosopher Confucius (551–479 BCE). When Zigong, one of his students, asked him: "Is there a word that could guide a man throughout his whole life?", Confucius answered "What about reciprocity, don't do to other what you wouldn't choose for yourself" (Hinton, 1998). This principle is known as the Golden Rule. Its positive version, stating "do to others as you would have them do to you," is also found in several places in the Christian Bible, including in Matthew 7-12 and Luke 6:27. It also constitutes an important principle in the Hindu religion, as well as in Judaism and Islam. In the Book of Mathew, this principle is placed in the center, where the Book says: "So in everything, do to others what you would have them do to you, for this sums up the Law and the Prophets" (Mathew 7:12).

In philosophy, Emanuel Kant revived the reciprocity principle and developed it further in his famous "Categorical Imperative", which in one of its formulations prescribes "Act only according to that maxim whereby you can, at the same time, will that it should become a universal law" (Kant, 1785). It is worth noting that both Confucius and the Kant's rules focus largely on morality, prescribing that individuals should behave towards others, as they want others, in similar situations, to behave towards them. In a broader perspective the reciprocity principle applies to a wider range of reciprocal behaviors, as it applies not only to others' *expected* behavior, but also to their *actual* behaviors. In real life the chain of reciprocal behaviors, of actions and reactions, is common in all aspects of human interactions.

In modern theory reciprocity had become one of the most investigated topics in the social and evolution sciences (e.g., Trivers, 1971; Gintis, 2000; Bolton & Ockenfels, 2000; Dufwenberg & Kirchsteiger, 2004; Falk & Fischbacher, 2006), and is considered by many as the major factor in the development and sustainability of human cooperation (Trivers, 1971; Gintis, 2000; Axelrod, 1984). Recent developments in the study of human social evolution and in experimental economics attest to the importance of reciprocity for understanding human behavior (Rapoport, & Chammah, 1965; Fahra & Irlenbusch, 2000; Fehr, Fischbacher, & Gächter, 2002; Seinen & Schram, 2006; Rockenbach & Milinski, 2006; Gächter & Herrmann, 2009).

Game Theory has a long tradition in the study of the role of reciprocity in repeated economic, social and political interactions. For the principle of reciprocity to apply, an interaction must be repeated for at least two rounds, one by each of the interacting parties. In repeated interactions in which each player, in turn, must divide a unit of divisible goods between herself and her counterpart the theory prescribes that each player should reciprocate by transferring to her counterpart an amount equal to the amount transferred to her by the counterpart in the previous round. Denote the amount kept by player's *i* in round *n* by $x_{i,\,n}$ and the amount transferred to player j by $1\text{-}x_{i,n}$. Similarly, denote the amount kept by player's *j* in round *n*-1 by $x_{j,n-1}$ and the amount transferred to player *i* by $1\text{-}x_{j,n-1,}$ the reciprocity principle prescribes that: $1\text{-}x_{i,n} = 1\text{-}x_{j,n-1}$ (or: $x_{i,n} = x_{j,n-1}$), implying that each player should transfer to her counterpart what the counterpart had transferred to her in the previous round. The reciprocity principle does not say how much a player, who is first to play, should transfer in the first round of the game, and other assumptions about the player's preferences and/or social orientation type (Van Lange, 1999) should be made. A player with other-favoring taste (e.g., a cooperative type) might transfer an equal share, while a completely self-favoring (individualistic) or a competitive type will transfer nothing or almost nothing. Since reciprocity implies mimicking the counterpart's previous move, the outcome of a mutually reciprocal behavior depends on the level of cooperativeness of the first mover. A famous and widely investigated strategy that implements the principle of reciprocity is the Tit for Tat strategy (TFT) (Milinski , 1987; Nowak & Sigmund, 1992, 1993; Goldstein, 1991; Kollock, 1993; Segal & Sobel, 2007; Brams & Kilgour, 2012). This strategy was first proposed by Anatol Rapoport in his extensive study of behavior in the Prisoner's Dilemma (PD) game (Rapoport & Chammah, 1965). It prescribes that a player should start the interaction by cooperation and reply to the counterpart's cooperation by cooperation, and to her defection by defection. The most well-known effort in this line of research is the famous tournaments conducted by Robert Axelrod's (Axelrod, 1981) and his computer simulation studies of the evolution of cooperation among egoists (Axelrod, 1984). Conversely, the extensive experimental work of Anatol Rapoport and many others yielded mixed results by demonstrating that in experimental settings participants are often caught in a trap of mutual defection, known as the D-D lock in effect, where the interacting parties are "locked" in a long series of mutual defection. The general picture that emerges from this research is that players usually start by a

reasonable level of cooperation (about 50% of the moves), followed by a phase in which cooperation drops sharply (the D-D lock-in-effect phase), and only after a long period of mutual defection that they inter a "recovery phase", in which their cooperation reaches a level which is slightly higher than the cooperation level of the initial phase (Rapoport & Chammah, 1965).

**2. The self-referencing principle – "Treat others as you treat yourself".**

In addition to the reciprocity or Golden Rule principle, Christianity, Islam and Judaism preach another principle, which is commonly confused with the Golden rule, although it differs from it and from Kant's Categorical Imperative in some fundamental aspects. The Jewish Bible, which might be considered the parent of Christianity and Islam, preaches to "Love others like you would yourself". Interpreters consider this commandment as the core principle of the Bible and of Judaism as a whole. Islam and Christianity also preach the same principle and prescribe to it the same degree of importance. Islam in teaches that "No one could be considered a believer until he desires for his brother what he desires for himself". The New Testament mentions the principle in various chapters, including Mark (2:28) and Mathew (22:35-40). In Matthew 22:35-40, we are told: "Then one of them, a lawyer, asked Him a question, testing Him, and saying, 'Teacher, which is the great commandment in the law?' Jesus said to him: 'You shall love the Lord your God with all your heart, with all your soul, and with all your mind. This is the first and great commandment. And the second is like it: You shall love your neighbor as yourself. On these two commandments hang all the Law and the Prophets."

In comparison to the Reciprocity (Golden rule) principle, which has attracted considerable theoretical and experimental research, the self-referencing rule had hardly been studied. In a recent paper, I proposed a new economic theory, called "Economic Harmony Theory" (EHT), which prescribes that egoist individuals who behave only according to self-favoring considerations and who has no taste for altruism or other favoring motives, can, and in fact must cooperate and behave fairly in order to maximize their personal benefit (Suleiman, 2103a, 2013b). In addition to the rationality principle, ETH rests on the following three assumptions:

1. When making their decisions, individuals consider their payoffs relative to *subjective reference points (SRPs),* rather than their absolute payoffs. A *SRP* can be social, when a player compares her payoff to the payoff(s) of another member, or members in her group (e.g., a co-

worker's salary), or non-social, when she compares her payoff with a neutral (non-social) reference point (e.g., her expected expenditure).

2. There exists a formal or informal sanctioning mechanism, by which effective sanctions are applied on deviants from the group's norms and rules.

3. Individuals are aware of the norms of equality and equity, as they are practiced in their social group.

The assumption of EHT that players are cognizant of the norms of equality and equity does not mean that they have a taste for fairness. Rather, it is assumed that *individuals use the information about existing social norms of equality and equity and the sanctions for deviations from these norms for their own self-interest*. The novelty of ETH lies in the plausible assumption that different players are motivated by different reference points, and that adherence to a certain reference point is influenced by the social structure of the interaction and by the position of the player in the social group. In real life, this assumption is not hard to defend, not only because different individuals have different motivations and intentions, but also because individual motivations and intentions are strongly affected by the determinants of their position in a specific social setting. An employee who earns a monthly salary of $x might compare her salary with the salary of another workmate, or with the average salary for workers with similar expertise (social *FP*s), but she might also compare her salary with the salary she could have received had she chosen another job offer, or with her monthly expenses (non-social *FPs*). Her employer might compare the net income generated by the employee with the salary she pays her employee, by the income generated by another equally paid employee (social *FP*s), or she might compare her net income with a projected profit which could guarantee a minimum growth rate of her workplace. The point to make here is that all intersections of employer-employee reference points are plausible.

According to EHT the existence of a sanctioning agency is a necessary condition for cooperation and fairness to emerge. In this respect, EHT is similar to the theory of Altruistic Punishment (Fehr & Gachter, 2002). In EHT the punishment agency can be a second party, as in dyadic interaction when one party can punish another for deviating from acceptable norms. It can also

be applied by a third party, as witnessed in experiments on Public Goods games with punishment (Fehr & Gachter, 2002; Fehr & Rockenbach, 2003; Egas & Riedl, 2008) or by a central, agreed upon authority (Yamagishi, 1986; Mulder, Van Dijk, Wilke, & De Cremer, 2005; Samid & Suleiman, 2008). I show elsewhere that ETH accounts well for a large set of experimental data from experiments on different games with two, three and five players and that it outperforms all other theories, including standard game theory (the SPE model) and major economic theories of cooperation like the theory of Equity Reciprocity and Competition (ERC) (Bolton & Ockenfels, 2000) and the theory of Inequality Aversion (Fehr & Schmidt, 1999).

In the present article I limit myself to the case of dyadic interactions. Specifically I shall present the theory's solution for the widely investigated Ultimatum Game (UG) (Güth, Schmittberger, & Schwartze, 1982; Camerer & Thaler, 1995; Suleiman, 1996) and compare it with the solution provided by the reciprocity principle. In the one-stage UG a proposer makes an offer (x, 1-x), for self a designated recipient, respectively. The recipient responds by either accepting the offer, in which case both players receive their offered shares, or by rejecting the offer, in which case the two players receive nothing. The UG structure meets EHT's conditions, since it embodies a sanctioning option, by which responders can punish unfair proposers (Xiao, & Houser, 2005). It is well documented that responders use their punishment option although it is costly for them. Experimental results show that offers of 20% or less are rejected with high probability and that the modal and mean offers in the game are about 50% and 40%, respectively, indicating that the proposers raise their offers in fear of being punished (Xiao & Houser, 2005; Camerer, 2003). Moreover, there seems to be a consensus among researchers that by proposing a reasonable offer, a rational player increases the probability that her offer will be accepted. The fact that proposers care more about appearing fair (out of self-interest) and less about being fair, has been nicely demonstrated in previous studies (e.g., Kagel, Kim, & Moser, 1996).

To derive ETH's solutions for the standard UG, we must consider all possible subjective reference points (*RPs*) of both proposer (P) and responder (R). This yields two symmetric [(P-social, R-social), (P-Non-social, R-Non-social)], and two asymmetric [(P-social, R-Non-social), (P-Non-social, R-social)] intersection points. Such consideration, detailed elsewhere (Suleiman, 2013a, 2013b) reveals that from the point of view of the proposer, the amount that she could

have gotten, had she retained the entire amount for herself, is most likely to be the preferred SRP to adhere to, from the standpoint of the recipient, comparison with the proposer's portion is more probable than other comparisons. In mathematical terms, P will evaluate her subjective benefit or utility in terms of the ratio $\frac{x}{1}$, whereas R will evaluate her own benefit in terms of the ratio $\frac{1-x}{x}$. Balance or harmony between the two utilities is achieved if $\frac{1-x}{x}$ equals $\frac{x}{1}$, or:

$$\frac{1-x}{x} = \frac{x}{1} \qquad \ldots\ldots (1)$$

Eq. 1 which I derived from completely rational considerations implies that harmony in the investigated situation will be achieved if, in relative terms, the proposer treats the responder as she treats herself. One conclusion that might be drawn at this point posits that the respective religious commandment discussed above could be sought not as one that preaches for benevolence or altruism, but as a rational choice by egoists who want to maximize their relative payoffs under the aforementioned assumptions. It is possible that the commonality between the monotheistic religions Judaism, Christianity and Islam, in their arriving at the same rational-economic rule, is attributed to commonalities in the cultural evolutionary processes of the peoples that inhabited the lands that were the cradle of the these religions.

Equation 1 could be rewritten as:

$$x^2 + x - 1 = 0. \qquad \ldots\ldots (2)$$

Solving for x, we obtain $x = \frac{-1 \pm \sqrt{1^2 - 4.1.(-1)}}{2} = (-\frac{\sqrt{5}+1}{2}, \frac{\sqrt{5}-1}{2})$, or approximately: - 1.618 and 0.618, respectively (see Endnote). For positive amounts, balance is achieved for a partition of $(\frac{\sqrt{5}-1}{2}, 1 - \frac{\sqrt{5}-1}{2})$, or about (0.62, 0.38) for P and R, respectively.

Another solution, not discussed here is the symmetric solution, according to which the entire amount is split equally between the two parties ($x = y = 1$). Worth noting that the asymmetric

solution seems fit as predictor of behavior in the UG, since the game structure, i.e., the game's set of rules is asymmetric.

Interestingly, the asymmetric solution $\frac{\sqrt{5}-1}{2}$ is the famous *Golden Ratio* (Livio, 2002; Olsen, 2006; Posamentier & Lehmann, 2007), which could be defined as $\lim_{n\to\infty}(\frac{f_n}{f_{n+1}}) = \frac{\sqrt{5}-1}{2}$, where $f_n$ is the $n^{th}$ term of the Fibonacci Series: 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144…, in which each term is equal to the sum of the two preceding terms $(f_n = f_{n-1} + f_{n-2})$ (41). While the equal split solution is in agreement with the prediction of the Equality Principle and with focal points considerations (Schelling, 1980; Messick & Sentis, 1983; Jost & Kay, 2010), the asymmetric solution at the Golden ratio is a novel one. It emerges from the model's assumption that different players may compare their payoffs to different reference points. Because the Golden Ratio is usually associated with harmony (Livio, 2002; Olsen, 2006), I call the points of "balance" "harmony points".

To summarize, the analysis of the UG game yields two points of balance or harmony: an asymmetric one at which the proposer receives a payoff of $\frac{\sqrt{5}-1}{2} \approx 0.62$ (62%) of the entire amount and the responder receives 1- $\frac{\sqrt{5}-1}{2} \approx 0.38$ (38%), and an symmetric solution, discussed elsewhere (Suleiman, 2013a, 2013b), which prescribes the equal division $(\frac{1}{2}, \frac{1}{2})$ between the two players.

Worth noting that the choice between the two solutions depends on how the situation is framed or perceived by the player in the role of proposer. Despite its structural asymmetry, if we can assume that players view the situation as symmetric, then the point-prediction is the 50/50 split. On the other hand if we can assume that the situation is perceived as what it is (i.e., asymmetric), then we should predict a 62/38 split. Because the theory is normative and does not relate to individual differences, the more plausible prediction is the asymmetric 62/38 split. In experimental settings, where the roles are assigned by a random device, participants might reason that there is no advantage for the proposer. Indeed, in almost all UG experiments the modal offer is the equal split (e.g., Camerer & Thaler, 1995; Camerer, 2003), indicating that the majority of participants in these experiments view the situation as symmetric. However, in the

same experiments the mean offer is about 60/40, which is in nice agreement with the theory's asymmetric prediction (for more details see, Camerer & Thaler, 1995; Suleiman, 1996; Suleiman, 2013a). In experiment in which the role of proposer is assigned by merit, either by winning an auction (Guth & Tietz, 1986) or by winning in a general knowledge test (Hoffman et al., 2004) the reported mean offers made by proposers become more egotistic.

It is important to note that the balance or harmony point at the Golden Ratio is not in equilibrium. For example the proposer will benefit from unilaterally deviating from a 62/38 split of 100 monetary units, say, to a 70/30 split. By doing so, she will increase her relative payoff from ≈62% to 70%, while reducing the responder's relative payoff from φ% ≈62% to 43%. Obviously, the best response for proposers is the SPE, of almost zero or zero offer. For the harmony point to be in equilibrium, it must be supported by an effective punishment, such that the proposer's negative payoff resulting from punishment, should, in relative terms, exceed her benefit from deviating from harmony. In mathematical notation, the punishment $P(x)$ inflicted on a proposer who increases her demand from the harmony point $x_h$ to $x$ ($x>x_h$) should satisfy $P(x)> x - x_h$.

Interestingly, the solution at the Golden Ratio and the Fibonacci series play an important role in many fields of science. Studies in biology have shown that the Fibonacci numbers plays an important role in determining the development of plants and animals, including the human body (Hammel, 1987; Klar, 2002) and human brain (Roopun, 2008; Conte, 2009; Weiss & Weiss, 2010) . It is also used in the visual arts, music and design due to the fact that humans find this ratio, harmonious and attractive both visually and auditory (Hammel & Vaughan, 1995; Pittard, Ewing, & Jevons, 2007). As examples, rectangles /height ratio of ≈ 0.618 (or ratio height/width of ≈ 1.618) are widely used in the design of TV and computer screens, as well as in credit cards due to feelings of pleasantness that they induce (Hammel & Vaughan, 1995). In physics it was recently shown that the golden ratio plays a key role both empirically and theoretically, in determining the critically quantum energy at which matter transforms to a wave (Coldea et al., 2010) or to dark matter (Suleiman, 2013c).

## 3. Experimental validation of the self-referencing rule

I tested the above derived prediction of the self-referencing rule using two large-scale datasets: (1) data from a Meta-analysis which integrated 37 studies conducted in 25 different countries, representing different cultures and social-political systems (Oosterbeek, Sloof & Van de Kuilen, 2004), and (2) data collected in a comprehensive study in 15 small communities, exhibiting a wide variety of economic and cultural conditions (Henrich et al. 2006). Figures 1(a & b) depict the distribution of proposers' offers, recipients' rejections, and final payoffs in the two investigated data-sets. As shown in the figure, the mean offers reported by the two studies are almost identical (40.5% and 39.5%, respectively), and quite close to the predicted Golden Ratio equilibrium (38.2%). Table 1 compares the predicted proposers' offers with the predictions of the Equality principle (EQ), predicting a 50/50 split, the sub-game perfect equilibrium (SPE), predicting an $\varepsilon \rightarrow 0$ offer, the Inequality Aversion (IA) theory (Fehr & Schmidt, 1999), predicting $\approx$ 50/50 split, and the Equity Reciprocity Conflict theory (ERC), which predicts that the proposer should offer any portion between an infinitesimally small positive portion and 50% (Bolton & Ockenfels, 2000).

The bottom row in the table depicts the various predictions' errors. It demonstrates that the prediction of the proposed EHT theory is superior to all the other theories, including the equality principle which predicts the modal offer in most ultimatum experiments (Güth, Schmittberger & Schwartze, 1982; Camerer & Thaler, 1995; Camerer, 2003; Kagel, Kim & Moser, 1996; Suleiman, 1996). To test the adequacy of the various theories in accounting for the experimental data, I used the *two one-sided test* (*TOST*). This equivalence analysis determines whether a variable's mean is sufficiently close to a hypothesized parameter. If the confidence interval $100(1-2\alpha)$ is within a defined interval, we conclude that equivalence exists. For a pre-specified equivalence interval of ±10% and a log-normal distribution of mean offers, the analysis yields a significant result ($p<0.01$) for the large industrial societies (Oosterbeek, Sloof & Van de Kuilen, 2004) and a marginally significant result for the small-scale communities (Henrich et al., 2006). Similar *TOST*s performed for the adequacy of all the other theories were statistically insignificant.

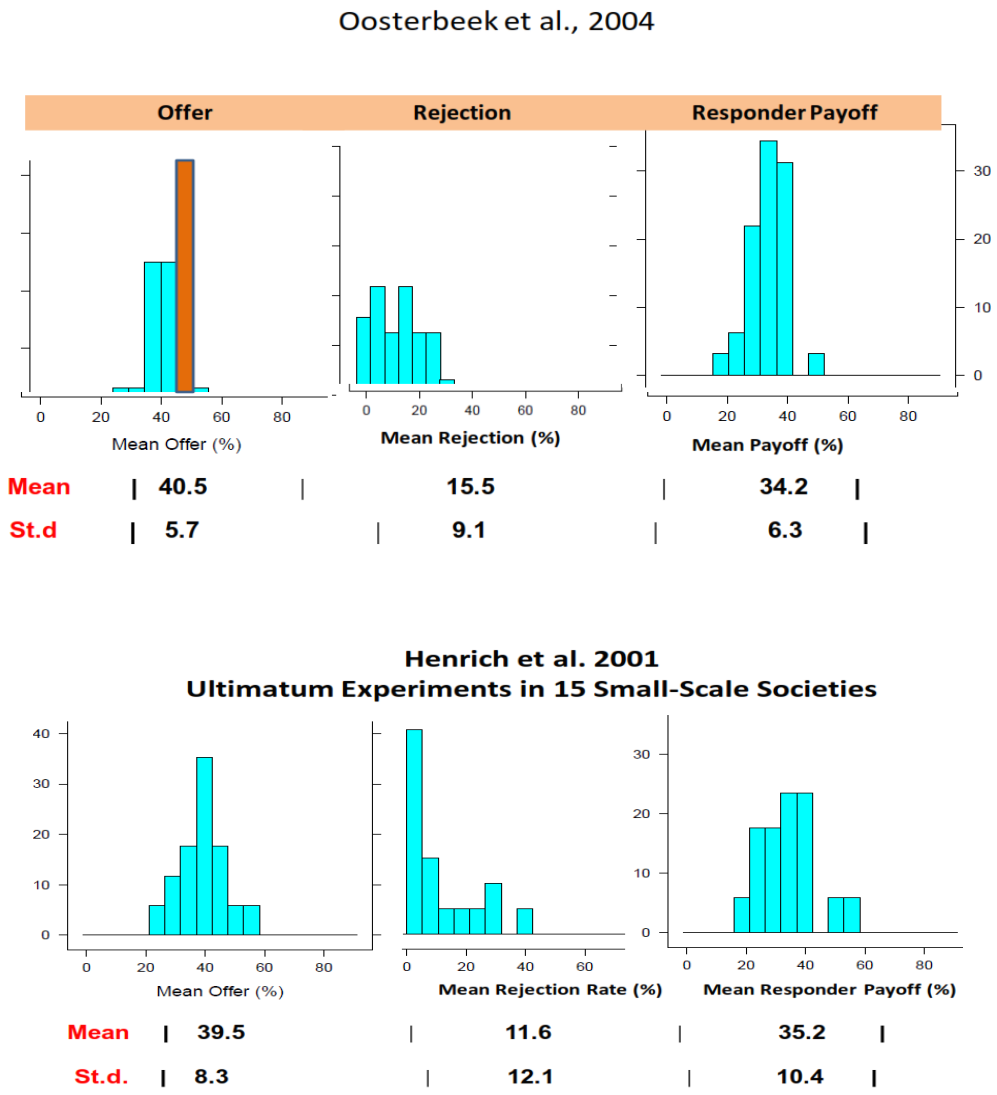


**Figure 1 (a &b):** Distributions of offers, rejection rates and final payoffs in two large-scale studies (a) Oosterbeek, Sloof & Van de Kuilen, (2004) and (b) Henrich et al. (2006)

| **Mean Offer (Experimental)** | **Subgame Perfect Equilibrium SPE** | **Inequality Aversion IA** | **Equity-Reciprocity Conflict ERC** | **Equality EQ** | **Economic Harmony Theory (ETH) (self-referencing rule)** |
|---|---|---|---|---|---|
| 1. Oosterbeek et al. (2004) **40.5 (5.7)** 2. Henrich et al.(2006) **39.5 (8.3)** | 0 | ≈ 50 | Any offer in [0, 50) | 50 | $(1-\frac{\sqrt{5}-1}{2}) \approx 38.2$ |
| **Overall Prediction Error (in %)** | ≈ 40 | ≈ 10 | ≈ 15 | ≈ 10 | 1.3 |

**Table 1.**Comparison between the predictability of ETH (self-referencing rule) with the predictions of four theories.

## 4. Reciprocity and Self-Referencing – A Comparison

In the present paper I discuss only economic interactions that could be modeled as an Ultimatum game. I compare between the two principled according to three criteria: (1) the degree of fairness in the distribution of the common wealth, (2) the degree of stability of the social contract resulting from the adherence to each principle, and (3) the degree of efficiency in managing the common resource.

**4.1. Fairness**: The Reciprocity Principle, or Golden Rule, prescribes that given the opportunity to allocate resources themselves and others, individuals should allocate to their counterparts as much as they allocate to themselves, had the situation been reversed. Because the Ultimatum game gives a special standing to the first mover, it is hard to make a point prediction on what an allocator, who follows the reciprocity principle, would do. If we assume that both players are rational and that rationality is common knowledge, then allocators will reason that responders would accept any offer, which they themselves would have accepted had they been in the role of responders. Consequently, a weak formulation of the reciprocity principle would prescribe that

the proposers will offer small portions of the entire cake to their responders, while a strong formulation would predict that the proposers will offer the smallest indivisible amount, if any.

In contrast the Self Referencing principle makes a precise point-prediction prescribing that proposers should offer to keep a Golden Ratio portion of the entire amount (≈ 62%) and offer the rest (≈ 38) to the responders. As mentioned above, experimental findings lend strong support to this prediction (for more details Suleiman 2013a and 2013b).

**4.2. Stability**: The reciprocity principle is vulnerable to noise and misunderstandings. For example, one side could mistakenly misinterpret a cooperative behavior by her counterpart as being non-cooperative or even aggressive, which might lead her to reply accordingly, causing a chain of retaliation which might be costly to both parties and to the collective goods as a whole. Situations of this type are well studied in a variety of disciplines in the social sciences. Many theoretical and empirical studies have demonstrated that the introduction of a low level of "noise" to the interaction was sufficient to severely impair the performance of reciprocal strategies like TFT (Wu & Axelrod, 1995; Bendor, Kramer, & Stout, 1991). In comparison, the self-referencing principle is considerably less susceptible to misunderstandings and noise. Except in extreme or abnormal cases, most people, most of the time, will have a reasonable picture of what or how much they want for themselves.

The aforementioned notwithstanding, individuals may lack an incentive to adhere to the self-referencing principle due to the temptation to maximize their own profit. Moreover, there is the question of who will pay the cost of a sanctioning agency to support the self-referencing principle; a problem known as a second-order dilemma (Yamagishi, 1988).

**4.3. Managing the Common Resource**: By this I mean providing, maintaining and increasing the common resource. Under the reciprocity principle misunderstandings like the one described in the previous section might be detrimental to the common wealth. Evidence of individuals and groups depleting their resources below critical levels due to an entrapment in cycles of revenge is unfortunately plentiful. Numerous experimental and field studies on the provision of public goods and on the management of common pool resources (CRPs) demonstrate that the lack of cooperation by a minority of free-riders eventually drives cooperators to significantly reduce

their levels of cooperation (e.g., Ostrom, 1990; Burger et al. 2001, Andreoni, 1988; Isaac, Walker, & Thomas, 1984; Kim & Walker, 1984). In contrast the self-referencing rule does not include a revenge option, since the option for punishment, which is essential to enforce and maintain cooperation could be placed in the hands of a third party or an institution.

The reciprocity and self-referencing principles could be compared on additional dimensions, some of which are mentioned hereby in brief. One dimension pertains to the duration of the interaction. While adhering to the reciprocity principle is possible only in repeated interactions, the self-referencing principle applies to one-period interactions as well. Another aspect concerns the group size. In groups with three players or more ($n \geq 3$), in which the group members have no information about the individual contributions of others, a retaliatory response to non-cooperators might hurt not only free-riders, but also others who did cooperate, leading them to reciprocate by lowering their contribution or even cutting it completely (Dawes, 1980). In contrast, applying the self-referential principle is not influenced by the group size.

## 5. Summary and concluding remarks

In the present paper I compared two moral principles: the reciprocity principle and a recently proposed self-referencing principle. The two principles date back to ancient times and are considered as core principles in most religious and moral teachings. While the reciprocity principle has attracted much research in a variety of disciplines, including economics, psychology and philosophy, the self-referencing principle, as a normative theory, was proposed only recently as an economic theory of cooperation (Suleiman, 2013a, 2013b). For the case of single interactions between two individuals, which could be modeled as an Ultimatum game, the predictions of the reciprocity principle are inconclusive and depend on our assumption about the first mover expectations and social type. Rational proposers might reason that had they been in the role of responders, they would have accepted any low offer, and this would inevitably lead them to behave in accordance with the standard game theoretic solution. This solution, known as the Sub-Game Perfect equilibrium prescribes that proposers should keep most of the amount to themselves. In contrast, the self-referencing principle yields a solution at the Golden Ratio, according to which proposers should keep about 62% of the entire amount and give offer about 38% to the recipients.

The comparison detailed in the previous section demonstrates that on several dimensions, the self-referencing principle is more advantageous than the reciprocity principle. Nonetheless, reciprocity is the principle that seems to dominate human interactions both on the interpersonal and intergroup levels.

This raises the two related questions: 1. why despite its inferiority in important aspects, including the degree of efficiency in utilizing the common wealth and its fair allocation, and despite the teachings of all big religions to "love for neighbors as you love for yourself", the reciprocity principle (Golden Rule) is by far the more prevalent principle among the two. 2. What remedial solutions one can suggest to promote the use of the self-referencing rule.

With regard to the first question I argue that from an evolutionary perspective, the superiority of the reciprocity principle could be attributed to the fact that in repeated interactions, reciprocity enables the interacting parties to positively reinforce their counterparts when they behave cooperatively. More importantly, it enables players to punish others who behave uncooperatively. In contrast the self-referencing principle lacks a similar punishment mechanism. Thus, social evolution has apparently selected the reciprocity principle despite its inferiority.

With regard to the second question, two possible solutions are proposed to rectify the anomaly. First, I propose that the self-referencing principle could be advanced and supported by the establishment of a *social contract* according to which the majority of individuals agree on a "mutual coercion mutually agreed upon" norm (Hardin, 1968), according to which group members establish a decentralized punishment mechanism, by which they can punish free-riders. An alternative or complimentary solution prescribes that groups establish a central authority, or in Hobbes's terms, a ***Leviathan*** (Hobbes, 1651; Levy, 1954), which has the legitimacy and power to punish uncooperative individuals who defect and free-ride on others. In modern societies such arrangement could be achieved by the establishment and consolidation of a ***Rule of Law***, under which ***all*** people become equal.

**References**


Andreoni, J.(1988). Why Free Ride? Strategies and learning in public goods experiments, Journal of Public Economics, 37, 3, 291-304.

Axelrod, R. (1981). The emergence of cooperation among egoists. *The American Political Science Review*, *75* (2), 306-318.

Axelrod, R. (1984). *The Evolution of Cooperation*, Basic Books.

Bendor, J., Kramer M, & Stout, S. (1991). When in doubt: Cooperation in a prisoner's dilemma game. Journal of Conflict Resolution, 35, 691-719.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A Theory of equity, reciprocity, and competition. *American Economic Review*, *90*(1), 166-193.

Brams, S.J. & Kilgour, D.M. (2012) Inducible Games: Using Tit-for-Tat to Stabilize Outcomes. MPRA Paper from University Library of Munich, Germany. http://econpapers.repec.org/scripts/search/search.asp?ft=tit+for+tat

Burger, J., Ostrom, E., Norgaard, R., Policansky, D., & Goldstein B. D. (2001). Protecting the Commons: A Framework for Resource Management in the Americas. Island Press.

Camerer, C., & Thaler, R. H. (1995). Anomalies: Ultimatums, Dictators and Manners. The Journal of Economic Perspectives, 9 (2), 209-219.

Camerer, C., & Thaler, R. (1995).Ultimatums, dictators and manners. *Journal of Economic Perspectives*, *9* (2), 209-219.

Camerer, C. (2003). Behavioral Game Theory: Experiments in Strategic Interaction. Princeton University Press, Princeton, NJ.

Coldea, R., Tennant, D. A., Wheeler, E. M., Wawrzynska, E., Prabhakaran, D., Telling, M., Habicht, K., Smeibid, P., & Kiefer, K. (2010). Quantum criticality in an Ising chain: Experimental evidence for emergent E8 symmetry. Science, 327(5962), 177-180.

Conte, E., Khrennikov, A., Federici, A., & Zbilut, J. P. (2009). Variability of brain waves: A new method based on a fractal variance function and Random Matrix Theory. Chaos, Solitons &Fractals, 41(5), 2790-2800.

Dawes, R. (1980). Social dilemmas, Annual Review of Psychology, 31, 169-193.

Dawes, R., & Messick, D. (2000). International Journal of Psychology, 35 (2), 111-116

Dufwenberg, M., & Kirchsteiger G. (2004). A theory of sequential reciprocity, Games and Economic Behavior, 47 (2), 268-298.

Egas, M., & Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. Proc. R. Soc. B, 275, 871-878.

Fahra, R., & Irlenbusch, B. (2000). Fairness as a constraint on trust in reciprocity: earned property rights in a reciprocal exchange experiment. *Economics Letters*, *66* (3), 275–282.

Falk, A., Fischbacher, U. (2006). *A theory of reciprocity*, *Games and Economic Behavior*, *54* (2), 293–315.

Fehr, E., & Schmidt, K. M. (1999).A theory of fairness, competition, and cooperation. Quarterly Journal of Economics, 114, 817-868.

Fehr, E. & Gachter, S. (2002). Altruistic Punishment in Humans. Nature, 415, 137–140.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, *13* (1), 1-25.

Fehr, E. & Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. Nature, 422, 137-140.

Gächter, S., & Herrmann, B. (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Phil. Trans. R. Soc. B*, *364*, 791-806.

Gintis, H. (2000). Strong Reciprocity and Human Sociality, *Journal of Theoretical Biology*, *206* (2),169–179.

Goldstein, J. (1991). Reciprocity in superpower relations: An empirical analysis. International Studies Quarterly, 35, 195-209.

Güth, W. Schmittberger, R. & Schwartze, B. (1982) An experimental analysis of Ultimatum games. Journal of Economic Behavior and Organization, 3, 367-388.

Guth, W., & Tietz, P. (1986). Auctioning ultimatum bargaining positions: How to act if rational decisions are unacceptable. In: R.W. Scholz (Ed.), Current Issues in West German Decision Research (pp. 173-185). Frankfurt: P. Lang.

Hammel, G. T. (1987). Fascinating Fibonaccis: Mystery and Magic in Numbers. Dale Seymour Publications. 49. Hammel, G. T., & Vaughan, K. C. (1995). Math and Music: Harmonious Connections. Dale Seymour Publications.

Hardin, G. (1968). The Tragedy of the Commons, Science, 162(3859), 1243-1248.

Henrich, J. et al. (2006). Costly punishment across human societies. *Science, 312*, 1767-1770.

Hinton, D. (1998). *The Analects of Confucius.* Counterpoint.

Hobbes, T. (1651). Leviathan, or The Matter, Forme, & Power of a Common-Wealth Ecclesiasticall and Civill. London: Green Dragon in St. Pauls Church-yard.

Hoffman, E., McCabe, K. Shachat, K., & Smith, V. (1994). Preferences, property rights and anonymity in bargaining games. Games and Economic Behavior, 7, 346-380.

Isaac, R. M., Walker, M, & Thomas, S. H. (1984). "Divergent evidence on free riding: An experimental examination of possible explanations," Public Choice, 43:2 113-149.

Jost, J. T., & Kay, A. C. (2010). Social Justice: History, Theory, and Research. Handbook of Social Psychology.

Kagel, J. H., Kim, C., & Moser, D. (1996). Fairness in ultimatum games with asymmetric information and asymmetric payoffs. Games and Economic Behavior, 13, 100-110.

Kant, E. (1785). *Grounding for the Metaphysics of Morals 3rd ed.* (translated to English by James W. Ellington, 1993.

Kim, O., & Walker, M. (1984). The Free Rider Problem: Experimental Evidence, Public Choice, 43:1 3-24. Keywords: experiments, public, voluntary contributions, free riding.

Klar, A. J. S. (2002). Fibonacci's flowers. Nature, 417(595), 177-180.

Kollock, P. (1993). An eye for an eye leaves everyone blind: Cooperation and accounting systems. American Sociological Review, 58, 768-786.

Levy, A. (1954). Economic Views of Thomas Hobbes. Journal of the History of Ideas, 15(4), 589–595.

Livio, M. (2002). The Golden Ratio: The Story of Phi, the World's Most Astonishing Number. New-York: Broadway Books.

Messick, D. M. & Sentis, K. P. (1983). Fairness, preference, and fairness biases. In D. M. Messick & K. S. Cook (Eds.), Equity theory: Psychological and sociological perspectives, (pp. 61-64). New-York: Praeger.

Milinski , M.(1987). Tit for Tat in sticklebacks and the evolution of cooperation. *Nature, 325*, 433 - 435.

Mulder,L., Van Dijk, E.,Wilke, H. A. M., & De Cremer, D. (2005). The effect of feedback on support for a sanctioning system in a social dilemma: The difference between installing and maintaining the sanction. Journal of Economic Psychology, 26, 443–458.

Nowak, M., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature, 355*:250-253.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature 364*:56-58.

Olsen, S. (2006). The Golden Section. New-York: Walker & Co.

Ostrom, E. (1990). Governing the Commons: The Evolution of Institutions for Collective Action. New York: Cambridge University Press.

Oosterbeek, H., Sloof, R., & Van de Kuilen, G. (2004). Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Journal of Experimental Economics*, *7*(2), 171-88.

Pittard, N., Ewing, M., & Jevons, C. (2007). Aesthetic theory and logo design: examining consumer response to proportion across cultures. International Marketing Review, 24(4), 457-473.

Posamentier, A. S., & Lehmann, I. (2007). The fabulous Fibonacci numbers. Prometheus Books.

Rapoport, A. & Chammah, A.M. (1965). *Prisoner's Dilemma*, The University of Michigan, Ann Arbor, MI.

Rockenbach, B., & Milinski, M. (2006).The efficient interaction of indirect reciprocity and costly punishment. *Nature*, *444*, 718-723.

Roopun, A. K., Kramer, M. A., Carracedo, L. M., Kaiser, M., Davies, C. H., Traub, R. 25. D., Kopell , N. J., & Whittington, M. A. (2008). Temporal interactions between cortical rhythms. Frontiers in Neuroscience, 2(2), 145-154 .

Samid, Y., & Suleiman, R. (2008). Effectiveness of coercive and voluntary institutional solutions to social dilemmas. In A. Biel, D. Eek, T. Garling, & M. Gustafsson (Eds.). New Issues and Paradigms in Research on Social Dilemmas (pp. 124-141). Springer.

Schelling, T. C. (1980). The Strategy of Conflict. Harvard University Press.

Segal, U. & Sobel, J. (2007). Tit for Tat: Foundations of Preferences for Reciprocity in Strategic Settings. No 9905, UWO Department of Economics Working Papers from University of

Western Ontario, Department of Economics. http://econpapers.repec.org/scripts/search/search.asp?ft=tit+for+tat.

Seinen, I., & Schram, a. (2006). Social status and group norms: Indirect reciprocity in a repeated helping experiment. *European Economic Review*, *50* (3), 581–602.

Suleiman, R. (1996). Expectations and fairness in a modified ultimatum game. Journal of Economic Psychology, 17, 531-554.

Suleiman, R. (2103a). Economic Harmony: A theory of cooperation between egoists. Manuscript under review in *Psychological Review*.

Suleiman, R. (2103b). Why Proposers offer 60/40? - The Golden Ration Solves the Ultimatum Puzzle. Manuscript under review in the Journal of Behavioral Decision Making.

Suleiman, R. (2013c).The Dark Side Revealed: A Complete Relativity Theory Predicts the Content of the Universe. To appear in Progress in Physics, 3, 2013 (forthcoming).

Trivers, R. (1971). The Evolution of Reciprocal altruism, *The Quarterly Review of Biology*, *46* (1), 35–57.

Van Lange, P. A. M. (1999). The Pursuit of Joint Outcomes and Equality in Outcomes: An Integrative Model of Social Value Orientation, *Journal of Personality and Social Psychology*, *77* (2), 337-349.

Weiss, H., & Weiss, V. (2010). When frequencies never synchronize: the golden mean and the resting EEG. Behavioral Brain Research, 1335(4), 91-102.

Wu, J., & Axelrod, R. (1995) How to cope with noise in the iterated Prisoner's Dilemma. Journal of Conflict Resolution, 39 (1), 183-189.

Xiao, E., & Houser, D. (2005). Emotion expression in human punishment behavior. Proceedings of the National Academy of Sciences of the United States of America, 102, 7398-7401.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. Journal of Personality and Social Psychology, 51, 110–116.

Yamagishi, T. (1988). Seriousness of Social Dilemmas and the Provision of a Sanctioning System, Social Psychology Quarterly, 51 (1), 32-42.

**Endnote**

The quadratic equation in its general form could be written as:

$ax^2 + bx + c = 0$

And its solutions are given by: x = $\frac{-b \pm \sqrt[2]{b^2-4\ a\ c}}{2a}$

Substituting a=b=1 and c= -1 yields:

$x^2 + x - 1 = 0$،

and x=$\frac{-1 \pm \sqrt[2]{1^2-4\ .1.9(-1)}}{2}$ = $\frac{-1 \pm \sqrt[2]{5}}{2}$ , or:

$x_1$= φ = $\frac{\sqrt[2]{5}-1}{2}$ ≈ 0.618 and $x_2$= φ + 1= Φ = $\frac{\sqrt[2]{5}+1}{2}$ ≈ - 1.618

Where φ and Φ are the golden ratio and its transverse Φ = $\frac{1}{\varphi}$ ≈ 1.618