INTERNATIONAL JOURNAL OF PSYCHOLOGY AND PSYCHOLOGICAL THERAPY 2011, 11, 2, pp. 183-199

Coping with Bargains in the Ultimatum Game: The Triple Circuit Hypothesis

Álvaro Machado*, Luiz Roberto Britto, and Eduardo Oda

[1]*Universidade de São Paulo, Brazil*

Abstract

This paper aims to present a new hypothesis on cognitive and neurobiological processes involved in the evaluation of offers, based on Ultimatum Game (UG). Recently, different studies have linked serotonin and serotonin-related compounds to rejection rates in this game, through the mediation of intolerance to unfairness, thus leading to the *serotonin hypothesis* of UG. Despite the great interest of these findings, the current paper shows that the behavior of a responder in the game is much more complex than originally thought, and that are needed at least three cognitive schemas and neurobiological processes to properly cope with that behavior. This paper is designed as a classic scientific hypothesis. First, it defines the epistemological basis of the hypothesis, which is introduced in relation to limitations of the field that are expected to be overcome by this endeavor. Next, it presents evidence for the hypothesis, and finally it makes predictions that can be used to test it. The new hypothesis is named triple-circuit hypothesis; it states that at the cognitive level, the minimum schemas to represent the responder's behavior are: pre-consciousness discrepancy; attributional schema based upon valence activation; and the counterfactual tendency to repress impulsive behaviors. At the neurobiological level, it proposes that the essential circuit relies on: transient decreases in phasic activity of neurons located in the dorsolateral portion of midbrain (error processing); MPFC excitatory firings toward the limbic system (especially the amygdala), mainly through glutamatergic pathways; dopaminergic activity toward the MPFC, generating inhibitory activity, which disinhibits limbic activity. Several evidences in support our hypothesis are presented.
Key words: ultimatum game, neurobiology of bargains, fairness, decision-making.

Resumen

Se presenta una nueva hipótesis sobre los procesos cognitivos y neurobiológicos implicados en la evaluación de ofertas en el Juego del Ultimatum (UG). Recientemente, diversos estudios han relacionado la serotonina y los compuestos relacionados con ella a las tasas de rechazo en el UG, a través de la mediación de la intolerancia a la injusticia, lo que ha conducido a formular la "hipótesis de la serotonina" del UG. A pesar del gran interés de estos hallazgos, este artículo pretende mostrar que la conducta de un jugador en este juego es mucho más compleja y son necesarios, al menos tres esquemas cognitivos y procesos neurobiológicos para explicarla. Este artículo ha sido diseñado como la presentación de hipótesis científica clásica, en la que primero se definen las bases epistemológicas presentadas teniendo en cuenta las limitaciones del campo a la espera de que esfuerzo ayude a clarificarlas. En segundo lugar, se presentan las evidencias de la hipótesis para, finalmente, elaborar las predicciones que pueden usarse para verificarla. La nueva hipótesis presentada

* Correspondence concerning this article should be addressed to: Dr. Álvaro Machado Dias, Institute of Psychyatry, Dept. of Neuroimage, University of São Paulo, Rua Dr. Ovidio Pires Campos s/n, 05508-030 Cerqueira Cesar, São Paulo, Brazil. Email: alvaromd@usp.br

se denomina hipótesis del triple circuito, y afirma que, a nivel cognitivo, los esquemas mínimos para representar la conducta del jugador son: la discrepancia preconsciente, el esquema atribucional basado en la activación, y la tendencia contrafactual para controlar las conductas impulsivas. A nivel neurobiológico, la hipótesis propone que el circuito esencial descansa sobre un descenso transitorio en la actividad fásica de las neuronas localizadas en la porción dorsolateral del cerebro medio (procesamiento de errores), disparos excitatorios de la MPFC hacia el sistema límbico (especialmente la amígdala), principalmente a través de las rutas glutamatérgicas, y actividad dopaminérgica hacia la MPFC que genera actividad inhibitoria que desinhibe la actividad límbica. Se presentan diversas evidencias en apoyo de nuestra hipótesis.
Palabras clave: juego del ultimatum, neuroeconomía, imparcialidad, toma de decisiones.

In the ongoing discussion on the neurobiological basis of decision-making, several attempts have been made to define how different patterns of thought (more specifically named 'cognitive schemas'), brain circuits, and neurochemical compounds affect the magnitude of offers and the choice to accept/reject one. One of the most remarkable findings to date is that intrinsically cultural cognitive schemas (e.g. moral feelings and associated representations) bias the tendency to maximize immediate gains, at the same time that specific brain lesions reverse this tendency (Greene, 2007; Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006).

Many behavioral tasks are usually applied to the understanding of the nature and structure of cognitive schemas that may affect the tendency to maximize gains for both proposers and responders, but none has received the attention that has been directed to the Ultimatum Game (UG), created by Rubinstein (1982). The UG is a zero-sum, two-player game in which participants alternate as proposer and respondent of monetary offers, and together meet the challenge of either reaching an agreement or facing a mutual loss. The game is at the core of the emerging field of neuroeconomics (for a presentation of the field: Sanfey, Loewenstein, McClure, & Cohen, 2006; for a discussion on the role of the UG: Sanfey, 2007), where it is often designated as a Theory of Mind game (Behrens, Hunt, & Rushworth, 2009), in the sense that it is based upon the prospection of the intentions of the opponent.

Several studies have linked the cognitive schemas activated during rounds of UG to the activation of specific brain areas (Knoch, *et al*., 2006; Michael Koenigs & Daniel Tranel, 2007; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). More recently, neurophysiological studies based on the 'acute tryptophan depletion' paradigm (ATD) showed a direct relation between tryptophan depletion and rejection rates (Crockett, Clark, Tabibnia, Lieberman, & Robbins, 2008; Emanuele, Brondino, Bertona, Re, & Geroldi, 2008), hence suggesting a pivotal role for serotonin in the evaluation of bargains. These findings complement conclusions regarding the positive correlation between testosterone levels and rejection rates (Burnham, 2007; for a replication: Zak, *et al*., 2009; for a study challenging these findings: Eisenegger, Snozzi, Heinrichs, & Fehr, 2010), as well as inverse correlations to the presence of omega-3 fatty acids (Emanuele, Brondino, Re, Bertona, & Geroldi, 2009).

Considering that both omega-3 (Yao, *et al*., 2004) and testosterone (Fink, Sumner, Rosie, Wilson, & McQueen, 1999) directly affect serotonin levels, it is reasonable to assume that both could be part of a neurobiological hypothesis, as addressed by Enzo Emmanuele and collaborators with their 'serotonergic hypothesis' (Emanuele, Bertona, Re, & Brondino, 2009; Emanuele, *et al*., 2008), or even further, by a hypothesis capable of integrating both the neuroanatomical and neurophysiological findings.

From that perspective, it is interesting to note that in order to achieve its maximum soundness, a neurobiological hypothesis of the kind must first model the most significant stages of the decision-making dynamics in the UG and then conceive the neurobiological basis of each of these stages. Many important features of this integrated process can be lost, as one disregards the stages of decomposing and understanding each step, in order to present it as a direct function of neurobiological activity.

This paper takes for granted that to achieve such goal, we need a new and broader cognitive and neurobiological model, which must be unambiguous, falsifiable, and at the same time includes much more than the direct relation between behavioral tendencies and specific neurochemical compounds. Our starting point is that the UG can represent an appropriate venue to study the cognitive and neurobiological basis of fairness. This premise follows from the fact that the rules of the game are quite simple, leading nearly all researchers to believe that rejection of unfair offers cannot be thought as the result of a diminished capacity to scrutinize costs/benefits, but rather as the result of the activation of cognitive schemas related to morality, in relation to which offers are thought of as fair/unfair (Sanfey, *et al*., 2003).

Considering that unfairness and moral violations in general have different meanings for proposers and responders, and that only the latter can be characterized as one who suffers from a moral violation and therefore should react to it, it follows that while playing the responder will turn the participant especially sensitive to cognitive schemas that may compel him to neglect the maximization of expected utility. In that sense, our model mirrors the assumption that the rationale behind the game needs to be split in two. Support for this premise comes both from findings that the process of decomposing the cognitive schemas that are activated during the game provide (as we will see) and from previous findings from related fields, akin to the conclusion that oxytocin (known to increase human interaction and empathy) augments the magnitude of mean offers while it does not affect the player's lower limit for acceptance (Zak, Stanton, & Ahmadi, 2007), therefore suggesting that the mechanisms of the brain whose activations are associated with the act of offering an amount that stems from values that maximize the expected utility are different from mechanisms whose activations lead to rejections that departure from the maximization thereof.

This paper is mainly focused on the cognitive and neurobiological basis of the responder's behavior, which we conceive as a three-stage process: information processing at a non-representational and unconscious level; activation of a moral/attributional schema (triggered as a certain cut-off is reached); counterfactual tendency to consciously repress the association between the moral/attributional schema and the behavior of rejecting an offer. The latter is a remarkable and under-explored trend that stands beyond the focus of near all medical, psychological, and anthropological studies on the relation between impulsivity and morality.

We also assume that the establishment of any conclusion regarding these stages depends on an appropriate understanding of the strategies that can be used in the UG, with a particular approach on the type of strategy that leads to the rejection of offers alleged to be unfair, as it tends to run against the maximization of the Expected Utility (EU).

FORMAL CONCEPTION OF MORAL STRATEGIES IN THE UG

A game is a situation wherein two or more individuals (players) have predefined objectives and face the need to follow formal rules, which limit the scope of their actions. In the UG game, the rules are fairly simple and relate to the division of a sum or 'pie': the first player (proposer) has to make an offer, the second player (responder) may either accept or reject; for that reason, it is said that the UG has only one parameter, the amount of goods presented in the pie.

Games are traditionally discussed in terms of strategies, which are expected to be rated in accordance with their potential success. Without loss of generality, we can presume that the amounts of goods involved in any round of the UG equals 1 and that a strategy consists of a pair (p,d), where "p" belongs to [0,1] and d: [0,1]--> {"Accept","Reject"}. If d(p)= "Accept", it follows that the first player receives "p"and the second receives "1-p", otherwise both receive "0" (see Figure 1).

An important feature of game analysis is the definition of the Nash Equilibrium and/or its properties for the game (general equilibrium analysis). By Nash Equilibrium of a game we mean the strategy that prevents a player from increasing his gains simply by changing his strategy in disregard to the opponent's strategy. In the very case of UG, it is quite easy to see that a strategy (p,D) such that D(p)= "Accept" and D(x)= "Reject" for all x>p is a Nash Equilibrium, and that the strategy (1,D), satisfying D(x)= "Reject" for all x>0, is another Nash Equilibrium.

We can also search for strategies that maximize payoffs. To reach this objective, the responder should accept any amount, since rejection implies receiving nothing. Conversely, the optimal strategy for the proposer is to offer the smallest amount ofgoods, limited by the smallest fraction of goods that is expected to be accepted by the other;

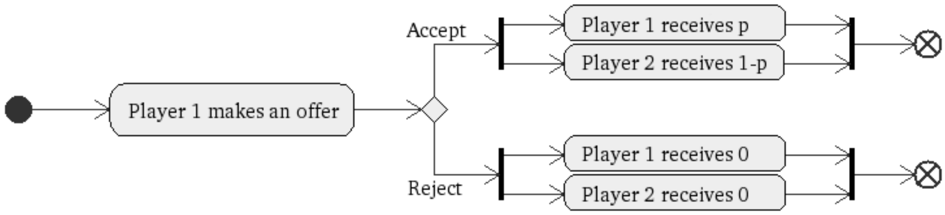

Figure 1. UML Activity Diagram representing UG.

when these strategies become dominant we say that a perfect Nash equilibrium subgame has been established.

Based upon this principle, experimental researches on the Ultimatum Game reveal that it deviates from perfect Nash equilibrium subgames due to the fact that proposers tend to offer and responders tend to solely accept pieces of the pie that depart from the minimum. This is the starting point to the inclusion of moral feelings in the conceptual framework that must drive the conception of the players' behavior, especially that of the responder provided that he is the one who 'reacts', and thus is expected to show the most intense feelings, as argued before.

Simply stated, the existence of a trend toward the rejection of unfair offers means that moral feelings take part in the establishment of utility, which on the other hand should approximate the function that maximizes monetary gains in accordance to the extent to which the player can be considered rational and less sensitive to these feelings. However, this is a purely tautological statement, which may lead to uncharacteristic conclusions (e.g., that chimpanzees are more rational than humans: Jensen, Call, & Tomasello, 2007), and to go beyond it we need to investigate the kind of strategy that the sensitivity to moral feelings represents.

Considered in broad terms, the responder's strategy can be represented by the following schema: if the value is 1, then the responder accepts the proposal; if the value is 0, he rejects it; at the intermediate values (between 0 and 1) the criterion to cope with the decision may be conceived as a cutting line (Alpha) that divides the curve in accordance with contingent demands. Based on that, we can mathematically summarize the discussion on the relation among fairness, rejection rates, and neurobiological activity, as the discussion about the role of internal and external determinants for three axes of the decision-making dynamics: general guidelines of acceptances (where '1' is fixed), general guidelines of rejection (where '0' is fixed) and determinants of the Alpha cutting line (which should be the most influenced by psychological demands), as represented in Figure 2.

Moving forward with this idea, we consider that the best way to analyze the interactive decision making behavior of near all players is to consider a specific version of the game, made of several rounds and the strategies showed in Table 1.

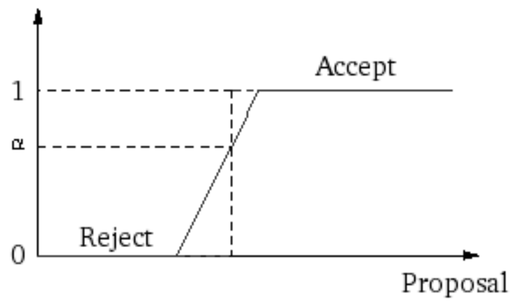

Figure 2. A model of a decision schema that accepts contingent inputs.

Table 1. Strategies used to simulate series of games.

Strategy	Offer	Response
Rational	The minimum.	Accepts any value greater than or equal the minimum.
Non-rational	Offers 0.5 minus the minimum.	Accepts above 0.5 minus the minimum.
Mimic	The same value that was previously offered by the other player. If he is proposer in the first round, then he offers 0.5.	If the other player accepted his previous proposal, he accepts any value above that proposal. Otherwise, he accepts above that proposal plus the minimum. In the first round, if he is the responder, he accepts above 0.5.
Just	Offers the same value the other player offered in the previous round. In the first round, if he is the proposer, he offers 0.5.	Accepts any value greater than or equal the minimum.
Altruist	Offers 1.	Accepts any value.

Table 2. Results of the simulations

	Rational	Non-rational	Mimic	Just	Altruist
Rational	0.5 x 0.5	0.245 x 0.255	0.49901 x 0.49999	0.5 x 0.5	0.995 x 0.005
Non-rational	0.255 x 0.245	0.5 x 0.5	0.49949 x 0.49951	0.5 x 0.5	0.755 x 0.245
Mimic	0.4995 x 0.4995	0.4995 x 0.4995	0.5 x 0.5	0.5 x 0.5	0.5005 x 0.4995
Just	0.49951 x 0.50049	0.49999 x 0.50001	0.5 x 0.5	0.5 x 0.5	0.5005 x 0.4995
Altruist	0.005 x 0.995	0.245 x 0.755	0.5 x 0.5	0.5 x 0.5	0.5 x 0.5

In order to analyse their roles in this game, we simulated 1000 rounds of UG confronting all these strategies, and obtaining the following average payoffs showed in Table 2.

According to our analysis, whereas rigid strategies -rational, non-rational and altruist- lead to poor performances in several cases, the strategies which adapt offers and responses to the opponent's actions (mimic) led to payoffs that most approximate to an optimum strategy. This is not to say that it was possible to prove that mimic strategies are optimal (in fact, they prove to be disadvantageous in some contexts), but that if instead of trying to focus on the maximization of the payoffs, we try to guide the payoffs to a Nash Equilibrium's payoff, we can expect to have advantageous results in the long run.

Bearing this analysis in mind, we add that the inclusion of moral feelings in the utility of the game is represented by the inclusion of psychological demands toward the abstract idea of 'fairness' and the objective perception of reciprocity (whose absence take part in the chimps' rationality), leading players to a provisional departure from immediate maximization of payoffs (e.g.: Knoch, *et al*., 2006; Koenigs & Tranel, 2007; Krueger, Grafman, & McCabe, 2008) in order to follow a decision-making pattern (D(p)) that equalizes gains. Without loss of generality, it is possible to add that the mimic strategy is the most suited to achieve such a pattern and that this is not generally irrational, but rather represents a specific type of adaptive best-reply seeking approach strategy (for a discussion, Roth & Erev, 1995) that can be quite efficient in the long run, as suggested in our simulations.

As one may note, this strategy is based on gradual learning (operant conditioning), supported by the computation of different degrees of reciprocity that players face along the rounds. Hence, it can be defined as a strategy that is best fitted for games with several rounds -as suggested by the evolutionary psychologists, who have raised the idea that moral behaviors are advantageous to the individual in the long run (e.g.: de Quervain, *et al*., 2004; Garcia & Ostrosky-Solis, 2006; Haidt, 2001; Kohlberg, 1981); for a simulation (genetic algorithm) based on learning in the UG, see Calderon and Zarama (2006). In line with this idea, a study by Hoffman, McCabe, and Smith (1996) revealed that "when a double-blind procedure intended to guarantee the complete social isolation of the individual's decision was used, 64% of the offers were $0, with only 8% offering $4 or more" (p. 654); e.g., when perspectives of reciprocity are unclear, the strategy tends to be less prominent.

These perspectives lead to the conclusion that the maximization of payoffs is not defined solely by the strategy that is being used, but by the number of trials or some information from the past behavior of the other player (Nowak, Page, & Sigmund, 2000), in relation to which it is suggestive that the process of coping with bargains in games consisted of few rounds (e.g., one-shot UGs), can be seen as an exception to the much more common many-rounds games (in real life), which may influence the behavior in these exceptional conditions. A history of positive outcomes applying the 'mimic strategy' should affect the way naïve players behave in games consisted of few rounds and even one-shot games, therefore suggesting that we should reconsider what these moral rejections actually represent: instead of being irrational (Jensen, *et al*., 2007), they are functional cognitive schemas that become rigidly manifested. Thus, our question turns out to be more quantitative (what leads to such intense activation) than a qualitative (why is the schema activated); and that is precisely our starting-point to model the cognitive and the neurobiological basis of the process of coping with bargains in the UG.

Triple-Circuit Hypothesis Cognitive Basis: Focus On Post-Moral Schemas

When modeling the cognitive basis of the process of coping with bargains in the UG, the most fundamental perspective is that acceptance/rejection relates to the evaluation of offers, while fairness relates to the evaluation of attitudes, which result from the perception that the offer was placed beyond the boundaries of expectancy. Unfair offers stand beyond these boundaries, while very advantageous offers stand above. This is a valid perspective for any value above the individual threshold, in face of which a player becomes motivated by the payoff of the game, but not necessarily valid in the context where the payoff does not generate an attempt to maximize the expected utility (e.g., in the context where a player that is insensitive to fake money is invited to play using such).

The rejection of an offer assumed to be unfair can only emerge after the biological computation of a discrepancy between the offer and certain (implicit) boundaries; consistent with this idea, a study showed that different offers generate evoked-related potentials (ERPs) of different magnitudes during the first 250-350 ms after the event (Polezzi, *et al*., 2008), which is a time frame known to be insufficient for conscious process; while a recent study revealed that expectation violations in any direction predicted the extent to which healthy subjects were able to recall the faces of players with who they have interacted in recent games (Chang & Sanfey, 2009); for a discussion Dias (2009). The meaning of these different ERPs can be linked to the fact that moral feelings also generate different ERPs in specific brain sites (Brazdil, *et al*., 2002; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Holroyd & Coles, 2002), and broadly to the premise that biological error-processing supports the phenomenological structure of all these processes (for classical approaches see Goldstein, 2004; Johnson-Laird, 1988; Newell, 1990).

These biological errors are mostly negative (below expectations) and promote the emergence of attributional schemas that provide meaning to them. Hence, moral feelings are not part of the intrinsic set of mental processes that are implicated in this game, but rather stand as a secondary cognitive schema. This is in agreement with the idea that punishment of unfair offers is observed even when the responder plays on the behalf of a third-party individual to whom he is unrelated (Civai, Corradi-Dell'Acqua, Gamer, & Rumiati, 2010) and regardless of the possibility of been noted as the agent of such behavior (Yamagishi, *et al*., 2009). That is, it is not correct to say that rejections simply follow from first person emotions (e.g., anger, vengeance), but that they tend to be guided by these attributional schema, which can be conceived as forms of interpreting and producing meaning to a phenomenon that is tied to error processing at the unconscious level.

At the point where an attributional schema is activated, the information process gains the level of experience and becomes available to some degree of conscious evaluation. This stage is chiefly associated with executive behavior (Baars & Franklin, 2003) and thus with a cognitive schema represented by the intention to maximize payoffs through the relative inhibition of the association between moral feeling and the rejection of the offer. Accordantly, a recent study revealed that rejections of offers in the UG are associated to emotional arousal, measured by the salivary alpha-amylase levels (Takagishi, Takayuki, Kameshima, Koizumi, & Takahashi, 2009), while another study revealed that this arousal is correlated to the perspective that the offer was placed by another human being, in opposition to a computer or robot (van't Wout, Kahn, Sanfey, & Aleman, 2006).

Conversely, the neglect of this dimension represents the assumption that players are intrinsically impulsive, which is simply not necessarily true (e.g.: Reynolds, Ortengren, Richards, & de Wit, 2006). Moral feelings are dimensional traits, as suggested by the fact that the BAS scores ('Behavioral Activation System', composed of 'Reward Responsiveness', among others) correlates with rejection rates (Scheres & Sanfey, 2006). In spite of countless studies showing that people depart from the immediate maximization of the expected utility, it is important to bear in mind that it would not be reasonable to assume that playing the game simply abolishes the capacity to inhibit impulsive behavior. Such a capacity accompanies the development of cognitive functions

(Casey, *et al*., 1997; Meyer-Lindenberg, Buckholtz, & Kolachana, 2006; Nigg, 2000) and consciousness (Zelazo, 2004), whose abnormalities are associated to disorders that are precisely related to disadaptive/impulsive decision-making (Siever & Davis, 1991), leading players to achieve certain cognitive control over the association between moral feelings and the behavior of rejecting offers.

Assuming that the strategy to cope with a game made of few rounds derives from the strategy applied to games made of several rounds, it is fair to consider that as the number of trial decreases, the role of a cognitive schema that may control/repress the association between moral feelings and behavior increases, regardless of the fact that the overall behavior stems from amoral patterns. That is precisely what characterizes the stage of relative inhibition of the association between the attributional schema that emerges from the neurocomputation of errors and the impulsive behavior; a struggle between what Haselhuhn and Mellers (2005) named 'pleasure from fairness' and 'pleasure from greediness' in the UG (Haselhuhn & Mellers, 2005).

We propose to call this schema as 'strategic repression'. It is based upon the idea that moral strategies tend to be characterized as impulsive tendencies in accordance to the (diminished) number of interactive rounds of a game, and that they are somehow mediated by counter-impulsive tendencies to increase immediate gains; unlike chimpanzees do, but by the association of cognitive schemas that diminish the intensity of the tendency towards rejection of offers conceived as unfair. Just like there are tendencies toward the rejection of unfair offers, there is a tendency to conceive these rejections as potentially unfavorable (pleasure from greediness). This tendency drives the human organism a bit closer to the maximization of expected utility, through its consciousness and will; and in terms of strategic modeling represents a psychologically-driven approximation of the mimic strategy to an ideal strategy. Note that we are not advocating to be the first to conceive the UG and related games as a function of successive and conflictive cognitive schemas (e.g., Ainslie & Monterosso, 2002); our contribution relies on the articulation of these stages that guide the final decision.

Altogether, we advocate that the acceptance of any offer is a 'simple choice'; its association with attributional schemas transforms that choice into a 'simple strategy' (not solely focused on the game, as it is involved with the necessity to respond to moral feelings). Finally, the association with a counterfactual tendency, turns that simple strategy into a 'complex decision-making process' and a real challenge for both mathematicians and cognitive scientists.

Triple-Circuit Hypothesis Neurological Basis

As suggested above, the cognitive process of coping with bargains is not as simple as it seems, although it starts as a simple error message that is yielded prior to conscious perception, which is driven in the negative direction (under expectancy) and then mapped into attributional schemas, which generate 'sense' (moral feelings) and a counterfactual tendency to inhibit its association with the behavior of rejecting the offer.

This type of simple-error message is represented in the central nervous system by decreased activity of phasic neurons in many brain sites (Niv & Schoenbaum, 2008). There is solid evidence suggesting that the starting point to this process is the midbrain (Matsumoto & Hikosaka, 2007), and that negative and positive directions are biologically dissociated: negative errors are preferentially encoded by dopaminergic neurons located in the dorsolateral portion of midbrain (in the substantia nigra pars compacta), while positive errors are strongly related to ventromedial and ventral tegmental dopaminergic activity (Matsumoto & Hikosaka, 2009).

This activity spreads mainly to the ventral striatum (Delgado, Nystrom, Fissell, Noll, & Fiez, 2000) (for an example specifically related to financial error prediction, see: Delgado, Li, Schiller, & Phelps, 2008) and the medial prefrontal cortex (MPFC) (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Modirrousta & Fellows, 2008). While the former seems to be involved also with the generation of a cognitive response to such an output, the latter seems to be involved mostly with the first steps of the evaluation of the discrepancy (Tobler, O'Doherty, Dolan, & Schultz, 2006). Moreover, the dopaminergic neurons along the midbrain-striatum-cortex axis exert a neuromodulatory effect on synapses encoding the mismatch (Joshua, Adler, Mitelman, Vaadia, & Bergman, 2008), up to the point when a response begins to be prepared.

This response does not constitute the behavior of rejecting or accepting the offer, but rather the activation of the attributional schema, which is precisely the focus of the 'serotonergic hypothesis' and related findings. In accordance with this hypothesis, the anterior cingulate cortex (ACC) shows a very high concentration of terminals expressing the serotonin transporter, while the rostral part of the limbic system delineates the most significant circuit in order to "assess the motivational content of internal and external stimuli" (Mantere, *et al*., 2002, p. 604).

From that point on, the integration of this partes extra partes schema with the tendency to inhibit impulsive behavior is biologically represented by the inhibitory firing of prefrontal (PFC) neurons. "Similar to the role that the prefrontal cortex plays in regulating the interaction with the external world, this region is crucial for attention to and inhibitory control of internal mental representations engaged during working memory, employment of strategies, planning and decision-making" (Knight, Richard Staines, Swick, & Chao, 1999, p. 162). And this closes the circuit.

Wherefore, it is important to bear in mind that despite the fact that cognitive processes maintain a direct relation with specific neurobiological activity, sometimes this is not necessarily a one-to-one relation. Such activity cannot be reduced to a struggle between neuronal activity related to the reactivity of attributional schema and the inhibitory activity of specific PFC neurons; instead, the literature suggests that it must include both the selective firing of PFC neurons to excite synapses of the inhibitory neurons of the limbic system (Quirk, Likhtik, Pelletier, & Pare, 2003; Rosenkranz & Grace, 2002) and the selective firing of other PFC neurons toward the inhibitory synapses of the former prefrontal neurons, thus leading to a counterfactual tendency of disinhibiting limbic activity (Rosenkranz & Grace, 2002).

The main network within which these integrated effects take place is represented by the activation of the basolateral synapses on the amygdala (BLA) by the glutamatergic neurons of the VMPFC (Rosenkranz & Grace, 2002), which increases limbic activity (and thus feelings), in association with the dopaminergic inhibition of VMPFC neurons, which act to inhibit the glutamatergic activity and, thus, produce inverse effects. "DA receptor stimulation can remove MPFC inhibition of the BLA and potentially allow a sensory cortical-driven affective response to be produced" (Rosenkranz & Grace, 2001, p. 4090).

The VMPFC neurons exert a complex role in regard to limbic drives and impulsive decision-making; and the literature on the matter needs to be analyzed with care. For example, a recent article emphasizes that "ventromedial prefrontal cortex (VMPFC) damage is reliably associated with defective emotion modulation; specifically, poorly controlled emotional responses that are incommensurate with the precipitating event" (Koenigs & Tranel, 2007, p. 951) and that "the hyper-irrational rejection of unfair Ultimatum offers observed in the VMPC patients, therefore supports the view that emotion regulation is an underlying component of normal economic decision-making in situations where financial considerations conflict with emotional responses" (p. 955). However, in another paper, the same main author along with other collaborators state that: "patients with VMPC lesions exhibit generally diminished emotional responsivity and markedly reduced social emotions", and from that standpoint conclude that they were more utilitarian than the average (Koenigs, *et al*., 2007, p. 908).

The apparent contradiction among these positions did not pass unnoticed by the authors, who stated: "in this study, the VMPC patients' abnormally high rate of utilitarian judgments is attributed to diminished social emotion, whereas in a recent study of the Ultimatum Game, the VMPC patients' abnormally high rate of rejection of unfair monetary offers was attributed to poorly controlled frustration, manifested as exaggerated anger" (p. 908).

It is not correct to say that ventromedial activity in most cases increases or decreases rationality, but rather that the proper functioning of the structure makes people more sensitive to social demands (e.g. to morality in general), thus making them less rational in decision-making in such contexts (due, e.g., to raised empathy). Notwithstanding, VMPFC lesions make people less emphatic and more irritable, increasing feelings of unfairness whenever they receive an unfair proposition.

Based on those assumptions and in our cognitive model, we present a representation of our 'triple-circuit model' (see Figure 3), which includes the general lines of the neurobiological basis of decision-making for responders. In its most general aspects, the triple-circuit hypothesis proposes that increased levels of glutamate and serotonin will decrease rejection rates, while increased levels of dopamine shall have the opposite effect. Note that the hypothesis does not aim to deny previous accounts on the role of serotonin or any other compound in the neurobiology of the process of copying with bargains in the UG, but rather tries to incorporate them into a broader, more complete framework.

EVIDENCE AND PREDICTIONS OF THE TRIPLE STRATEGY HYPOTHESIS

The fact that the BIS/BAS neuropsychological test is correlated to decision patterns in the UG (Scheres & Sanfey, 2006) is highly suggestive of the existence of a stage

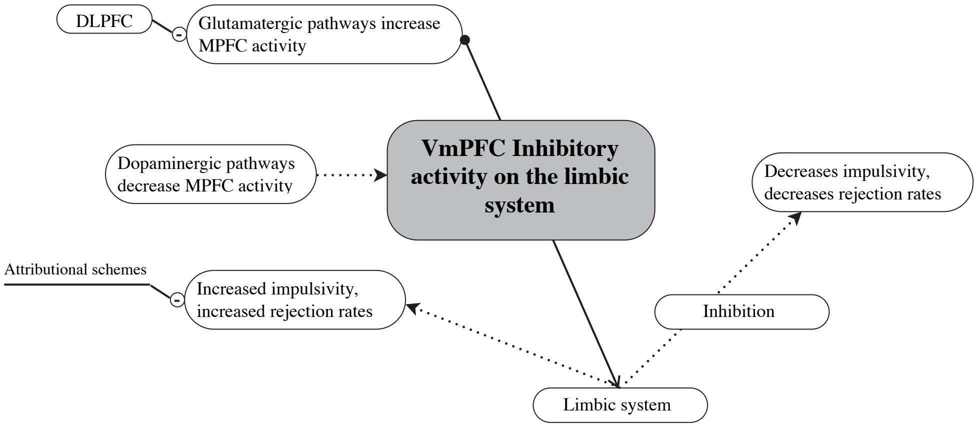

Figure 3. simplified algorithm representing the some of the most important dopaminergic and glutamatergic pathways that affect rejection rates and the generation of attributional schemas.

where some degree of inhibition, and thus of a struggle between a 'socially-desirable decision pattern' and an 'individually-desirable decision pattern'.

Also important is that VMPFC lesions increase rejection rates (Koenigs & Tranel, 2007) and this has to do basically with glutamatergic and dopaminergic pathways. In support of this argument, a recent article revealed that decreased right VMPFC size is correlated with diminished impulsivity control in young men (Boes, *et al*., 2009).

Dopamine agonists increase impulsivity (Cools, 2008; Potenza, 2007), while antipsychotic medication reverses this tendency (Potenza, 2007), which is a paradigm that suggests a parallel with the ATD paradigm (5-HT↑/DA↑).

Genetic studies present evidence showing that genes regulate the expression of both dopaminergic and serotonergic receptors, regulating decision-making processes that involve reward and social interactions. "The genetic mechanisms regulating dopaminergic and serotonergic synaptic transmission might underlie individual differences in behaviors and neural circuits implicated in reinforcement learning and therefore contribute to individual variability in social decision-making" (Lee, 2008, p. 407).

Strategies presented by those who adopt a more simplistic approach (defined as strategies based solely on their own will or impression, disregarding the role of the other player) tend to be more associated with ACC activity (a massively serotonin-innervated cortical area), while more sophisticated strategies tend to be associated with higher levels of mentalization and widespread prefrontal activity, as the medial prefrontal cortex seems to play a pivotal role (Coricelli & Nagel, 2009).

PREDICTIONS OF THE MODEL

If this model proves to be correct, we expect that:

- Neuropsychological tests that evaluate executive control in decision-making (e.g., Preference for Intuition or Deliberation -PID: Plessner, Betsch, & Betsch, 2008); adaptive decision-making (e.g., Iowa Gambling Task: Bechara & Damasio, 2005; Bechara, Damasio, Tranel, & Damasio, 2005), and response inhibition (e.g., Stroop Test: Stroop, 1935) should reveal positive correlations to adaptive decision-making in the UG, as one plays the responder, independently of any moral evaluation.

- Glutamatergic supplementation shall mimic the effects of serotonergic supplementation, while dopaminergic supplementation should have the opposite effect. Justification: the first two shall increase limbic inhibition, while the last one should have the opposite effect (this prediction should be considered on a broad basis, since there are many biding sites for each of these compounds).
- Brain stimulation (e.g. excitatory transcranial magnetic stimulation -TMS) in the medial prefrontal and orbitofrontal sites should decrease rejection rates, while inhibitory transcranial magnetic stimulation should produce the opposite effect. Justification: excitatory TMS increases neuronal activity in the sites that are stimulated, which in this case means control over limbic impulses. On the other hand, inhibitory TMS should block executive control.
- Excitatory TMS over the left DLPFC should decrease rejection rates, much in the same way that its effects on mood resembles the action of antidepressants (Avery, *et al*., 2008); beta wave neurofeedback should produce the same effect. Justification: the same as aforementioned, along with the perspective that the VMPFC and dorsolateral prefrontal cortex (DLPFC) work in coordination, with the former more closely associated with the computation of inputs associated with feelings and emotions, and the latter with abstract thinking (Meeks & Jeste, 2009).
- In multi-round games (Güth & Tietz, 1990), the effect of ATD should be proportionally greater in relation to the first unfair offer than to the game as whole. Justification: rejections that contradict expected utility are mainly associated with attributional schemas and thus to error prediction. As offers assumed to be unfair are presented, less correction is needed. This perspective may also provide a partial explanation for the existing differences in rejection rates in single-round vs. multiple-round UG (Meyer, 1992).

FINAL REMARKS

In this paper we have introduced a new hypothesis to explain the behavior of players rejecting offers in the UG. This hypothesis relies on the perspective that a player's decision in this case is associated with the use of three different types of mental schemas, each of which is associated with different patterns of brain activity produced both in coordinated and competitive fashions. In a broader sense, our paper posits a note of caution to those who assume that an observed association between neurochemical activity and behavior means the feasibility of establishing a one-to-one relation between the two -a tendency that may be termed simplistic reductionism.

Moreover, it challenges the idea that the definitive cognitive schema in the process of coping with bargains in the UG is represented by moral feelings and associated representations. In this sense, regardless of our adherence and empirical evidence of the principle (raised by evolutionary psychologists) that morality is an efficient strategy in zero-sum games, we are far less optimistic in relation to the forces that drive humans, at least as we found in westerns of the current century. As suggested by Mazar and collaborators (2008): "people typically solve this motivational dilemma adaptively by finding a balance or equilibrium between the two motivating forces, such that they derive some financial benefit from behaving dishonestly but still maintain their positive self-concept in terms of being honest" (Mazar, Amir, & Ariely, 2008, p. 634).

Based on the same general idea, we suggested that it is paramount to differentiate proposers and responders in terms of their relation to moral feelings; this is a necessary perspective not only in our hypothesis, but also in the understanding of findings reported in the literature, which extend to the role of different brain structures and specific neurotransmitters. Not only we expect that the specific roles of glutamatergic and dopaminergic cascades start to be explored in the same way as the ATD, as we expect that this paper stimulates new developments regarding the neurocognitive basis of non-cooperative decision-making.

REFERENCES

Ainslie G & Monterosso J (2002). Will as intertemporal bargaining: Implications for rationality. *University of Pennsylvania Law Review, 48*, 825-862.

Avery DH, Isenberg KE, Sampson SM, Janicak PG, Lisanby SH, Maixner DF, *et al*. (2008). Transcranial magnetic stimulation in the acute treatment of major depressive disorder: clinical response in an open-label extension trial. *Journal of Clinical Psychiatry, 69*, 441-451.

Baars B, & Franklin S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences, 7*, 166-172.

Bechara A & Damasio AR (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior, 52*, 336-372.

Bechara A, Damasio H, Tranel D, & Damasio AR (2005). The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *Trends in Cognitive Sciences, 9*, 159-162.

Behrens TEJ, Hunt LT, & Rushworth, M. F. S. (2009). The Computation of Social Behavior. *Science, 324*, 1160-1164.

Boes AD, Bechara A, Tranel D, Anderson SW, Richman L, & Nopoulos P (2009). Right ventromedial prefrontal cortex: a neuroanatomical correlate of impulse control in boys. *Social Cognitive & Affective Neuroscience, 4*, 1-9.

Brazdil M, Roman R, Falkenstein M, Daniel P, Jurak P, & Rektor I (2002). Error processing evidence from intracerebral ERP recordings. *Experimental Brain Research*, *146*, 460.

Breiter HC, Aharon I, Kahneman D, Dale A, & Shizgal P (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron, 30*, 619-639.

Burnham TC (2007). High-testosterone men reject low ultimatum game offers. *Proceedings of the Royal Society B: Biological Sciences, 274* (1623), 2327-2330.

Calderon JP, & Zarama R (2006). How Learning Affects the Evolution of Strong Reciprocity. *Adaptive Behavior, 14*, 211-221.

Casey B, Castellanos F, Giedd J, Marsh W, Hamburger S, Schubert A, *et al*. (1997). Implication of right frontostriatal circuitry in response inhibition and attention-deficit/hyperactivity disorder. *Journal of American Academy of Child & Adolescent Psychiatry, 36*, 374.

Chang LJ, & Sanfey AG (2009). Unforgettable ultimatums? Expectation violations promote enhanced social memory following economic bargaining. *Frontiers in Neuroscience, 3*, 36.

Civai C, Corradi-Dell'Acqua C, Gamer M, & Rumiati RI (2010). Are irrational reactions to unfairness truly emotionally-driven? Dissociated behavioural and emotional responses in the Ultimatum Game task. *Cognition, 114*, 89-95.
Cools, R (2008). Role of Dopamine in the Motivational and Cognitive Control of Behavior. *Neuroscientist, 14*, 381-395.
Coricelli G, & Nagel, R (2009). Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proceedings of the National Academy of Sciences, 106*, 9163-9168.
Crockett MJ, Clark L, Tabibnia G, Lieberman MD, & Robbins TW (2008). Serotonin Modulates Behavioral Reactions to Unfairness. *Science, 320*, 1739.
de Quervain DJ-F, Fischbacher U, Treyer V, Schellhammer M, Schnyder U, Buck A, *et al*. (2004). The Neural Basis of Altruistic Punishment. *Science, 305*, 1254-1258.
Delgado MR, Li J, Schiller D, & Phelps EA (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*, 3787-3800.
Delgado MR, Nystrom LE, Fissell C, Noll DC, & Fiez JA (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology, 84*, 3072-3077.
Dias AM (2009). Reconsidering the modularity of social cognition in the human brain. *Frontiers in Behavioral Neuroscience, 3*, 1-2.
Eisenegger C, Snozzi R, Heinrichs M, & Fehr E (2010). Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature, in press*.
Emanuele E, Bertona M, Re S, & Brondino N (2009). Human economic and financial behavior: The serotonergic hypothesis. *Bioscience Hypotheses, 2*, 109-110.
Emanuele E, Brondino N, Bertona M, Re S, & Geroldi D (2008). Relationship between platelet serotonin content and rejections of unfair offers in the ultimatum game. *Neuroscience Letters, 437*, 158-161.
Emanuele E, Brondino N, Re S, Bertona M, & Geroldi D (2009). Serum omega-3 fatty acids are associated with ultimatum bargaining behavior. *Physiology & Behavior, 96*, 180-183.
Falkenstein M, Hohnsbein J, Hoormann J, & Blanke L (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography & Clinical Neurophysiology, 78*, 447-455.
Fink G, Sumner B, Rosie R, Wilson H, & McQueen J (1999). Androgen actions on central serotonin neurotransmission: relevance for mood, mental state and memory. *Behavioural & Brain Research, 105*, 53-68.
Garcia AEV & Ostrosky-Solis F (2006). From morality to moral emotions. *International Journal of Psychology, 41*, 348-354.
Goldstein L (2004). Philosophical integrations. *Language Sciences, 26*, 545-563.
Greene JD (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences, 11*, 322-323; author reply 323-324.
Güth W & Tietz R (1990). Ultimatum bargaining behavior : A survey and comparison of experimental results. *Journal of Economic Psychology, 11*, 417-449.
Haidt J (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834.
Haselhuhn MP & Mellers BA (2005). Emotions and cooperation in economic games. *Cognitive Brain Research, 23*, 24-33.
Hoffman E, McCabe K, & Smith V (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review, 86*, 653-660.
Holroyd C & Coles M (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*, 679-708.
Jensen K, Call J, & Tomasello M (2007). Chimpanzees Are Rational Maximizers in an Ultimatum Game.

Science, 318, 107-109.
Johnson-Laird PN (1988). *The Computer and The Mind*. Cambridge: Harvard University Press.
Joshua M, Adler A, Mitelman R, Vaadia E, & Bergman H (2008). Midbrain Dopaminergic Neurons and Striatal Cholinergic Interneurons Encode the Difference between Reward and Aversive Events at Different Epochs of Probabilistic Classical Conditioning Trials. *Journal of Neuroscience, 28*, 11673-11684.
Knight RT, Staines RW, Swick D, & Chao LL (1999). Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta Psychologica, 101*, 159-178.
Knoch D, Pascual-Leone A, Meyer K, Treyer V, & Fehr E (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science, 314*, 829-832.
Koenigs M & Tranel D (2007). Irrational Economic Decision-Making after Ventromedial Prefrontal Damage: Evidence from the Ultimatum Game. *Journal of Neuroscience, 27*, 951-956.
Koenigs M, Young L, Adolphs R, Tranel D, Cushman F, Hauser M, *et al*. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature, 446*, 908-911.
Kohlberg L (1981). *Essays on Moral Development. Vol. 1 & 2*. San Francisco: Harper & Row.
Krueger F, Grafman J, & McCabe K (2008). Neural correlates of economic game playing. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*, 3859-3874.
Lee D (2008). Game theory and neural basis of social decision making. *Nature Neuroscience, 11*, 404-409.
Mantere T, Tupala E, Hall H, Sarkioja T, Rasanen P, Bergstrom K, *et al*. (2002). Serotonin Transporter Distribution and Density in the Cerebral Cortex of Alcoholic and Nonalcoholic Comparison Subjects: A Whole-Hemisphere Autoradiography Study. *American Journal of Psychiatry, 159*, 599-606.
Matsumoto M, & Hikosaka O (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature, 447*, 1111-1115.
Matsumoto M & Hikosaka O (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature, 459*, 837-841.
Mazar N, Amir O, & Ariely D (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research, 45*, 633-644.
Meeks TW & Jeste DV (2009). Neurobiology of Wisdom: A Literature Overview. *Arch Gen Psychiatry, 66*, 355-365.
Meyer-Lindenberg A, Buckholtz J, & Kolachana B (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences, 103*, 6269.
Meyer HD (1992). Norms and self-interest in ultimatum bargaining: The prince's prudence. *Journal of Economic Psychology, 13*, 215-232.
Modirrousta M & Fellows LK (2008). Dorsal Medial Prefrontal Cortex Plays a Necessary Role in Rapid Error Prediction in Humans. *Journal of Neuroscience, 28*, 14000-14005.
Newell A (1990). *Unified theories of cognition*. Cambridge: Harvard University Press.
Nigg J (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*, 220-246.
Niv Y & Schoenbaum G (2008). Dialogues on prediction errors. *Trends in Cognitive Sciences, 12*, 265-272.
Nowak MA, Page KM, & Sigmund K (2000). Fairness Versus Reason in the Ultimatum Game. *Science, 289*, 1773-1775.
Plessner H, Betsch C, & Betsch T. (2008). *Intuition in Judgment and Decision Making*. New York: Lawrence Erlbaum Associates.
Polezzi D, Daum I, Rubaltelli E, Lotto L, Civai C, Sartori G, *et al*. (2008). Mentalizing in economic decision-making. *Behavioural Brain Research, 190*, 218-223.
Potenza MN (2007). To Do or Not To Do? The Complexities of Addiction, Motivation, Self-Control, and Impulsivity. *American Journal of Psychiatry, 164*, 4-6.

Quirk GJ, Likhtik E, Pelletier JG, & Pare D (2003). Stimulation of Medial Prefrontal Cortex Decreases the Responsiveness of Central Amygdala Output Neurons. *Journal of Neuroscience, 23*, 8800-8807.

Reynolds B, Ortengren A, Richards JB, & de Wit H (2006). Dimensions of impulsive behavior: Personality and behavioral measures. *Personality and Individual Differences, 40*, 305-315.

Rosenkranz JA & Grace AA (2001). Dopamine Attenuates Prefrontal Cortical Suppression of Sensory Inputs to the Basolateral Amygdala of Rats. *Journal of Neuroscience, 21*, 4090-4103.

Rosenkranz JA & Grace AA (2002). Cellular Mechanisms of Infralimbic and Prelimbic Prefrontal Cortical Inhibition and Dopaminergic Modulation of Basolateral Amygdala Neurons In Vivo. *Journal of Neuroscience, 22*, 324-337.

Roth AE & Erev I (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior, 8*, 164-212.

Rubinstein A (1982). Perfect Equilibrium in a Bargaining Model. *Econometrica, 50*, 97-109.

Sanfey AG (2007). Social Decision-Making: Insights from Game Theory and Neuroscience. *Science, 318*, 598-602.

Sanfey AG, Loewenstein G, McClure SM, & Cohen JD (2006). Neuroeconomics: cross-currents in research on decision-making. *Trends in Cognitive Sciences, 10*, 108-116.

Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, & Cohen JD (2003). The Neural Basis of Economic Decision-Making in the Ultimatum Game. *Science, 300*, 1755-1758.

Scheres A & Sanfey AG (2006). Individual differences in decision making: Drive and Reward Responsiveness affect strategic bargaining in economic games. *Behavioral and Brain Functions, 2*, 35.

Siever L & Davis K (1991). A psychobiological perspective on the personality disorders. *American Journal of Psychiatry*, 148, 1647-1658.

Stroop J (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General, 18*, 643-662.

Takagishi H, Takayuki F, Kameshima S, Koizumi M, & Takahashi T (2009). Salivary alpha-amylase levels and rejection of unfair offers in the ultimatum game. *Neuro Endocrinological Letters, 30*, 643-646.

Tobler PN, O'Doherty JP, Dolan RJ, & Schultz W (2006). Human Neural Learning Depends on Reward Prediction Errors in the Blocking Paradigm. *Journal of Neurophysiology, 95*, 301-310.

Van't Wout M, Kahn RS, Sanfey AG, & Aleman A (2006). Affective state and decision-making in the Ultimatum Game. *Experimental Brain Research, 169*, 564-568.

Yamagishi T, Horita Y, Takagishi H, Shinada M, Tanida S, & Cook KS (2009). The private rejection of unfair offers and emotional commitment. *Proceedings of the National Academy of Sciences, 106*, 11520-11523.

Yao JK, Magan S, Sonel AF, Gurklis JA, Sanders R, & Reddy RD (2004). Effects of omega-3 fatty acid on platelet serotonin responsivity in patients with schizophrenia. *Prostaglandins, Leukotrienes and Essential Fatty Acids, 71*, 171-176.

Zak PJ, Kurzban R, Ahmadi S, Swerdloff RS, Park J, Efremidze L, *et al*. (2009). Testosterone administration decreases generosity in the ultimatum game. *PLoS One, 4*, e8330.

Zak PJ, Stanton AA, & Ahmadi S (2007). Oxytocin Increases Generosity in Humans. *Plos ONE, 2*, e1128.

Zelazo P (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences, 8*, 12-17.

Received, 16 July, 2010
Final Acceptance, 8 March, 2011