Impaired decision-making is a core problem in several psychiatric disorders including attention-deficit/hyperactivity disorder, schizophrenia, obsessive–compulsive disorder, mania, drug addiction, eating disorders, and substance abuse as well as in chronic pain. To ensure progress in the understanding of the neuropathophysiology of these disorders, animal models with good construct and predictive validity are indispensable. Many human studies aimed at measuring decision-making capacities use the Iowa gambling task (IGT), a task designed to model everyday life choices through a conflict between immediate gratification and long-term outcomes. Recently, new rodent models based on the same principle have been developed to investigate the neurobiological mechanisms underlying IGT-like decision-making on behavioral, neural, and pharmacological levels. The comparative strengths, as well as the similarities and differences between these paradigms are discussed. The contribution of these models to elucidate the neurobehavioral factors that lead to poor decision-making and to the development of better treatments for psychiatric illness is considered, along with important future directions and potential limitations.

Keywords: Iowa gambling task, animal model, validity, dopamine, serotonin, neurobiology

INTRODUCTION

Decision-making plays a pivotal role in daily life. Impairment in this process, as observed in several psychiatric disorders, results in an inability to make profitable long-term decisions that incorporate expectations of future outcomes (for review, see Dunn et al., 2006). Impaired decision-making is recognized as a core problem in disorders like substance abuse (Grant et al., 2000; Bechara, 2001; Bechara and Damasio, 2002; Bechara et al., 2002; Ernst et al., 2003; Bechara and Martin, 2004; Whitlow et al., 2004; Dom et al., 2006; Hanson et al., 2008), pathological gambling (Cavedini et al., 2002b; Brand et al., 2005; Goudriaan et al., 2005, 2006), schizophrenia (Bark et al., 2005; Shurman et al., 2005; Kester et al., 2006; Lee et al., 2007b; Sevy et al., 2007), obsessive–compulsive disorder (OCD; Lawrence et al., 2006; da Rocha et al., 2008; Cavedini et al., 2010; Starcke et al., 2010), eating disorders (Cavedini et al., 2004; Tchanturia et al., 2007; Liao et al., 2009; Brogan et al., 2010), attention-deficit hyperactivity disorder (ADHD; Toplak et al., 2005; Garon et al., 2006; Malloy-Diniz et al., 2007; Luman et al., 2008) chronic pain (Apkarian et al., 2004; Verdejo-García et al., 2009), and Parkinson’s disease (Brand et al., 2004). Therefore, understanding the mechanisms underlying poor decision-making is key to successful treatment of neurological and psychiatric disorders. This is why decision-making is studied intensively by disciplines ranging from neuroscience, cognitive and social psychology, to experimental and behavioral economics (Kahneman and Tversky, 1979) leading to the recent interdisciplinary field of neuroeconomics that combines all these disciplines (Glimcher et al., 2009).

The Iowa gambling task (IGT) is the most commonly used task to assess decision-making performance in a clinical setting (Bechara et al., 1994, 1999). The IGT is particularly interesting because it mimics the complexity of the choices that we are confronted with in everyday life. Its design incorporates the unpredictability of the consequences of a choice, the need to weigh short- and long-term gains and losses, and the necessity to exert behavioral control to maximize gains in the long-term. Successful performance requires the integration of several executive functions: individuals must demonstrate flexibility in planning to account for various outcomes, constantly monitor incoming information, evaluate the risk–reward ratio for various decision-making options, and refrain from choosing the options that are immediately more rewarding. For these reasons, successful performance requires the integration of several executive functions. The IGT was originally developed to assess the specific cognitive impairments of prefrontal cortex-damaged individuals (Bechara et al., 1994), but impaired performance has subsequently been observed following damage to other brain regions such as the amygdala, insula, and more controversially, the dorsolateral prefrontal cortex (dlPFC) (Bechara et al., 1998, 1999; Mannes et al., 2002; Clark et al., 2003; Fellows and Farah, 2005), as well as in a range of psychiatric disorders.
populations (see above). Several (psychiatric) conditions induce various kinds of deficits in the IGT, like disadvantageous deck preference (schizophrenia, OCD, pathological gambling, substance abusing individuals, psychopathic individuals, ADHD, chronic pain) with preference for infrequent punishments (ADHD, schizophrenia); no preference (anxiety, (Miu et al., 2008; de Visser et al., 2010)); slower learning (mania, substance-dependent individuals) or deficits in reversal learning (schizophrenia; for review, see Dunn et al., 2006). The effects of genetic polymorphisms, pharmacological treatments, as well as functional neuroimaging data, have significantly expanded our knowledge of the neural substrates underlying decision-making, and how their functions are compromised in individuals with decision-making deficits.

Interest is growing in the development of rodent models of decision-making for several practical reasons (see also Potenza, 2009). First, such models are indispensable for the dissection of precise mechanisms involved in decision-making, such as the role of specific brain regions and circuits, modulation by the monoaminergic systems, and neurodevelopmental events. Second, rodents are particularly suitable for screening and identifying risk or protective factors for poor decision-making. Rodent studies are not subject to the same time constraints associated with longitudinal studies in human populations and easily allow the study of inter-individual differences in behavioral and cognitive capacities (Rivalan et al., 2009b). Third, animal models are particularly valuable since environmental conditions as well as genetic variation can be carefully controlled. As such, animal models of the IGT have recently been developed (van den Bos et al., 2006; Pais-Vieira et al., 2007; Rivalan et al., 2009a; Zeeb et al., 2009). These models are largely complementary, yet have distinct strengths. Here we aim at describing the current state of these novel models with respect to face, predictive, and construct validities. We will present recent findings that demonstrate their potential to investigate the neuropsychobiological mechanisms of decision-making as well as directions for future research.

THE IOWA GAMBLING TASK

TASK CHARACTERISTICS

The IGT requires individuals to choose cards, one by one, between four different decks to earn money. Two decks are equally advantageous in the long run because cards chosen from these decks provide immediate moderate monetary gains but also moderate or low losses according to two different probability schedules. Two decks are equally disadvantageous in the long run because even though immediate gains are higher, the unpredictable losses are also higher, according to two different probability schedules (Figure 1). Thus, a conflict is induced between immediate high rewards and long-term gains. Participants are not provided with any information as to which choice is optimal, but they are instructed to try to maximize their gains as much as possible by freely choosing cards from each deck, and have the ability to switch between decks at any time (Bechara et al., 1999). Subjects therefore need to discover the task contingencies by trial and error. This sets the IGT apart from tasks that overtly signal the odds of winning such as the Cambridge Gambling Task (Clark et al., 2003).

When performing the IGT, healthy human subjects usually display a shift from primarily explorative behavior at the beginning of the task, during which they sample from all decks, toward a more exploitative strategy involving substantially more choices of the advantageous options associated with the best long-term outcome (Bechara et al., 1994, 1999; Brand et al., 2007b). Thus, decision-making is first made under ambiguous conditions, in that the subjects do not know what the reinforcement contingencies are. Following repeated sampling from the decks, it can be assumed that subjects are more aware of the chances of winning or losing associated with each deck, and therefore risky decision-making can take place (Stoltenberg and Vandever, 2010, but see Fellows and Farah, 2005).

Patients suffering from psychiatric disorders in which decision-making is compromised typically persevere in their choice of the disadvantageous options that yield immediate large rewards, despite larger losses in the long-term. Interestingly, a subset of healthy individuals also makes poor decisions in the IGT, suggesting a continuum between normal and pathological conditions (Brown and Barlow, 2005). Therefore, it can be hypothesized that poor decision-making in clinical and non-clinical populations shares common neuropsychological characteristics. As such, identification of these markers could improve our understanding of the transition from a healthy vulnerable state to psychiatric conditions.

NEURAL SUBSTRATES

Studies using brain-lesioned patients and imaging techniques have provided consistent evidence that decision-making depends on...
the integrity of, and functional connectivity between, many brain areas. The main structures are the amygdala, the insula, the striatum (STR), and several frontal cortical regions, including ventromedial prefrontal cortex (vmPFC), orbitofrontal cortex (OFC), the anterior cingulate cortex (ACC), and dlPFC (Bechara et al., 1999; Menas et al., 2002; Bolla et al., 2004; Tucker et al., 2004; Fellows and Farah, 2005; Hsu et al., 2005; Brand et al., 2007a).

The somatic marker hypothesis proposes that emotion-based biasing signals arising from the body are integrated in higher brain regions, notably the vmPFC, the amygdala, the insula, and the somatosensory cortex to regulate complex decision-making (Bechara et al., 1997; Dunn et al., 2006). This hypothesis is based on the fact that successful IGT performance is related to the development of somatic marker signals, as indexed by the magnitude of anticipatory skin conductance responses, before any conscious knowledge of the adapted choices (Bechara et al., 1994, 1997). These signals serve as an indicator of the value presented. If they are ineffective, like in vmPFC lesioned people, solving the task is no more possible. Within the vmPFC, the OFC is involved in the treatment, evaluation and filtering of perceptual, social, and emotional information (Elliott et al., 2000). This region is strongly interconnected with areas within the limbic system, particularly the basolateral amygdala (BLA), and receives prominent inputs from sensory association cortices. This pattern of connectivity suggests that the OFC plays a role in integrating potentially salient information about environmental contingencies (Ongur and Price, 2000), and uses this information to assign a value to a reward and signal outcome expectancies which can thus influence action selection (Schoenbaum et al., 2003; Rolls and Grabenhorst, 2008; Mainen and Kepcs, 2009; Takahashi et al., 2009). Thus, the OFC allows the representation of the reinforcing consequences of a choice to adapt goal-directed behaviors (Mainen and Kepcs, 2009) and modulates this value according to the contingency changes (Schoenbaum et al., 1998; Murray et al., 2007; Hayton and Olmstead, 2009).

The ACC is a converging area for cognitive and motor commands (Paus et al., 1993) that monitors and detects the presence of conflicts related to actions (Magno et al., 2006; Oliveira et al., 2007), whereas the dlPFC, tightly connected with ACC and OFC (Barbas, 2000), engages a top-down process required to monitor and implement change (Hyafil et al., 2009). The ACC signals error-likelihood (Sallet et al., 2007) and has a key role in choosing and updating appropriate actions when the environment is uncertain or dynamic (Kennerley et al., 2006; Quilodran et al., 2008), and in combining information about the costs and benefits associated with alternative actions (Rudebeck et al., 2006). This update is made in combination with the dlPFC, which is critically involved in the temporary maintenance of recently acquired information (Lee et al., 2007a), and in the detection of action-outcome contingencies (Balleine and O’Doherty, 2010), a major aspect of associative learning that allows the elaboration of goal-directed responding.

It has been suggested that the exploration and exploitation phases of the IGT involve different brain areas (van den Bos et al., 2007). The exploration phase may be mediated by the amygdala and ventral STR and their projections to the OFC (Bechara et al., 1999; Knutson et al., 2001). When the task progresses and a preference for the advantageous decks is emerging, the ACC, dlPFC, and dorsal STR may be recruited to engage in cognitive control of the once established choice in order to maintain and exploit this strategy to secure long-term payoff (Bush et al., 2000; Ernst et al., 2002; McClure et al., 2004; Ridderinkhof et al., 2004; Pezawas et al., 2005). However, large inter-individual differences in brain areas recruitment according to performances in the IGT, certainly occurs. Animal models of the IGT are uniquely placed to assess the validity of these theories, as brain imaging studies or lesions to particular areas can be selectively implemented at different stages of training, thereby preferentially targeting either the exploitation or exploration phases (de Visser et al., 2011b; Rivalan et al., 2011; Zeel and Winstanley, 2011).
have yielded conflicting results (He et al., 2010; Stoltenberg and Vandever, 2010; Lage et al., 2011). Because the SERT is responsible for serotonin reuptake into the presynaptic nerve terminal, the s-allele is hypothesized to be associated with increased extracellular serotonin levels (Lesch et al., 1996; Kalueff et al., 2009). Of interest, the s-allele, compared to the l-allele, is associated with amygdala hyperactivity in response to fearful stimuli and at rest (Hariri et al., 2002; Canli et al., 2006). This hyperactivation correlates with reduced volume of the ACC, as well as a functional and anatomical uncoupling between the ACC and amygdala (Pezawas et al., 2005; Pacheco et al., 2009) and hyperactivity of the vmPFC (Heinz et al., 2005). One speculation is that people carrying the 5-HTTLPR s-allele are hypervigilant, which is advantageous when environmental stimuli are controllable and manageable. Indeed, in a risky decision-making task with choice probability outcomes made explicitly s/s subjects chose more advantageous due to increased risk adversity (Kuhnen and Chiao, 2009). Conversely, under conditions in which stimuli are uncertain, such as during the IGT, s-allele carriers engage in maladaptive behavior (Homberg and Lesch, 2010). Overall, these data indicate that serotonin signaling can negatively affect decision-making on the IGT, with possible sex-dependent effects during the explorative phase of the task.

**Dopaminergic system and decision-making in the IGT**

Changes in the dopaminergic system have likewise been shown to modulate IGT performance. As previously noted, poor IGT performance has been observed in several pathologies related to dopamine dysfunction, such as schizophrenia, Parkinson’s disease, drug addiction, and pathological gambling (Comings et al., 1996; Rogers et al., 1999; Mimura et al., 2006). Dopamine is critically involved in associative learning (Schultz, 2002) time perception (Meck, 2006) and signaling within the reward system (Di Chiara and Bassareo, 2007), all of which are fundamental processes required for decision-making. Consequently, it is not surprising that reduction of dopaminergic levels impairs decision-making in healthy individuals. Acute administration of a branched-chain amino acids (BCAA) mixture lowers the plasma ratio of dopamine’s precursor amino acids and decreases dopaminergic activity (Harmer et al., 2001; Gijsman et al., 2002). This treatment also orientates healthy male participants toward disadvantageous decks in later trials as their dopamine levels reduce, thus indicating a fundamental function of dopaminergic signaling in advantageous guiding decision-making during the IGT (Sevy et al., 2006). This impairment could be related to reduced perception of probability and time, given that attention was shifted toward more recent events compared to more distant events (van den Bos et al., 2007). PET imaging data also indicate a positive correlation between dopamine release in the ventral striatum and IGT performance in healthy male participants (Linnet et al., 2010), yet a negative correlation in problem gamblers (Linnet et al., 2011). However, the D2 antagonist haloperidol increased the drive to play slot machines in pathological gamblers, but not in healthy controls (Zack and Poulos, 2007).

Two genetic factors mediating dopamine signaling, namely the catechol-O-methyltransferase (COMT) enzyme and the D4 dopamine receptor (DRD4), are also related to IGT performance. The rs4818 C/G polymorphism of the COMT gene, which results in an 18-fold divergence of enzymatic activity between the high activity variant (G allele) and low activity variant (C allele; Nackley et al., 2006), has been shown to significantly affect performance in the IGT. In particular, male subjects homozygous for the high activity variant performed better than those carrying the low activity variant (Roussos et al., 2008). In a financial risk decision-making task in which subjects were informed on the payoff of choice options carriers of the 7-repeat allele of DRD4 were significantly more risk seeking relative to those individuals without the 7-repeat allele (Kuhnen and Chiao, 2009), indicating that the 7R allele is associated with novelty-seeking regardless of choice conditions (uncertain or certain). Another COMT gene polymorphism, Val158Met, is associated with IGT performance in healthy females. The Met/Met variant, related to lower COMT activity and higher constitutive dopamine prefrontal cortical levels, lead to poorer performances compared to Val/Val (Lotta et al., 1995; Mannisto and Kaakkola, 1999; Chen et al., 2004; van den Bos et al., 2009b); but see (Kang et al., 2010). Interestingly, carriers of both the Met/Met and 5-HTTLPR-s/s genotypes displayed the worst IGT performance among all possible COMT and 5-HTTLPR genotype combinations (van den Bos et al., 2009b), indicating that the dopaminergic and serotoninergic systems may interact in the modulation of IGT choice behavior.

Regarding the variable number of tandem repeat (VNTR) polymorphism in the DRD4 gene, healthy male carriers of the seven repeats (7R) allele of this gene choose significantly more cards from disadvantageous decks in the IGT compared to participants exhibiting the 4R allele (Roussos et al., 2009). This 7R allele is associated with lower transcriptional/translational levels and diminished in vivo receptor responsivity compared to the four repeats (4R) allele (Hutchison et al., 2003; Hamarman et al., 2004; McGough, 2005; Brody et al., 2006; Ebstein, 2006).

Collectively, these data suggest that dopamine, possibly in interaction with serotonin, can modulate decision-making as measured in the IGT. Serotonin may be inversely associated with IGT performance, whereas directional consensus for dopamine is not yet fully clear. Pathological gamblers and healthy controls appear to react differently to dopaminergic manipulations, and may also differ in their baseline and gambling-induced changes in DA release. However, our understanding of the monoaminergic modulation of IGT performance is still limited, and consequently, pharmacogenetic treatment of decision-making deficits in patients is currently a distant goal.

**MODELING THE IGT IN RODENTS**

It is comparatively easy to perform pharmacological, genetic, and environmental manipulations in mice and rats. Hence, rodent IGT models (RGTs) can make a crucial contribution to scientific and medical advances in the field of decision-making. For this purpose, paradigms which capture the essence of the IGT have been developed for use in rodents in order to establish animal models of human decision-making with high face and construct validities. The IGT (Bechara et al., 1994) has several key features that had to be reproduced. As reviewed below, four different types of rodent IGT models have been designed that incorporate several of these key features (Table 1). These models differ both in the equipment used (mazes vs. operant chambers), the task duration (from a single session to several daily sessions), the learning of task contingencies and the ratio between advantageous vs. disadvantageous...
options. The way in which loss is signaled also differs: making sugar pellets aversive through the addition of the bitter-tasting substance quinine (van den Bos et al., 2006; rats and mice), simple non-reward (Pais-Vieira et al., 2007; rats), or the delivery of time-out periods on loss trials which minimize the number of pellets the animals can earn (Rivalan et al., 2009a; rats, Zeeb et al., 2009; rats, Young et al., 2011; mice).

**CONFLICT BETWEEN PROBABILITIES OF HIGH FOOD REWARD VS. QUININE: A FOUR-ARM BOX MAZE MODEL (RGTREWARD OR QUININE)**

The first experimental protocol aiming to reproduce the characteristics of the IGT in rodents was developed by van den Bos et al. (2006). This task measures the choice between four goal arms, two of them containing either food rewards or punishments per choice. High amount of rewards is combined with the incurrence of a high amount of punishment in the disadvantageous arm, as opposed to the advantageous one. Uncertainty is given by varying the sequence of sugar or quinine pellet presentation. The apparatus (Figure 2) consists of a box divided in three different areas: a starting zone, a choice zone, and an arena divided in four parallel arms. The goal arms, labeled A, B, C, and D, are provided with internal visual cues (symbols of different shapes and colors) to help animals in differentiating them during the task apart from the spatial location. They contain “monetary rewards” in the form of sugar pellets, or “monetary punishments” that are quinine-treated (bitter) sugar pellets. Before testing, animals are habituated to the sugar pellet taste and briefly allowed to explore the maze (10 min). Any animals which continue to eat the quinine-treated sugar pellets are removed from the experiment. Test trials are initiated by removing the slide door from the start box. The animal is allowed to freely explore the choice area. A choice for one of the arms is made when the animal has walked into the arm for at least one-third of the length of the arm. Once a choice is made, the arm is closed to prevent the animal from walking back without investigating the reward cup. During test, every other time. In the latter case, the possibility of receiving low immediate rewards but having a net gain over multiple choices is achieved by administering one sugar pellet eight times in 10 selections and one quinine-treated pellet every other time. The

<table>
<thead>
<tr>
<th>Task features</th>
<th>IGT</th>
<th>RGT-reward or quinine</th>
<th>RGT-reward probabilities</th>
<th>RGT-one session reward and time-out</th>
<th>RGT-reward or time-out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apparatus</td>
<td>Computerized</td>
<td>Manually operated maze</td>
<td>Manually/automated arena</td>
<td>Automated operant chamber</td>
<td>Automated operant chamber</td>
</tr>
<tr>
<td>No. of choice options</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Reward</td>
<td>Monetary gain</td>
<td>Sucrose pellets</td>
<td>Sucrose pellets</td>
<td>Palatable food pellets</td>
<td>Sucrose pellets</td>
</tr>
<tr>
<td>Reward occurence</td>
<td>Each trial</td>
<td>Alternate with punishment</td>
<td>Alternate with punishment</td>
<td>Each trial</td>
<td>Alternate with punishment</td>
</tr>
<tr>
<td>Punishment</td>
<td>Monetary loss</td>
<td>Quinine pellets</td>
<td>No reward</td>
<td>Time-outs</td>
<td>Time-outs</td>
</tr>
<tr>
<td>Conflict immediate rewards</td>
<td>100 (A and B) vs. 50 (C and D)</td>
<td>3 (A) vs. 1 (B)</td>
<td>3 (A) vs. 1 (B)</td>
<td>2 (A and B) vs. 1 (C and D)</td>
<td>3:4 (A and B) vs. 1:2 (C and D)</td>
</tr>
<tr>
<td>Conflict long-term payoff</td>
<td>Per 10 cards −250 vs. +250</td>
<td>Per 10 trials 3 vs. 8 pellets ratio = 2.7</td>
<td>Per 10 trials 9 vs. 8 pellets ratio = 0.9</td>
<td>Total test: 60 vs. 300 pellets ratio = 5</td>
<td>Total test: 353 vs. 117 (average) ratio = 3</td>
</tr>
<tr>
<td>Task duration</td>
<td>Single session (100 trials)</td>
<td>10 daily sessions (10–20 trials)</td>
<td>Single session</td>
<td>Single session (1 h)</td>
<td>25 daily sessions for stability (approximately 100 trials per 30 min session)</td>
</tr>
<tr>
<td>Pre-training procedure</td>
<td>None</td>
<td>10-min habituation</td>
<td>20–25 days</td>
<td>5–7 days</td>
<td>10–15 days incl. 7 forced-choice sessions</td>
</tr>
<tr>
<td>Prior knowledge of contingencies</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Motivational aspects</td>
<td>n/a</td>
<td>90–95% of FFW</td>
<td>80% of FFW</td>
<td>95% of FFW</td>
<td>85% of FFW</td>
</tr>
</tbody>
</table>

Values for the conflict long-term payoff are hypothetical calculations for RGT-reward or quinine and RGT-reward or time-out based on a fixed trial duration of 9 or 5 s, respectively. FFW, free-feeding weight.
uncertainty of rewards and punishments per choice provided by the human task is maintained by varying the sequence of sugar or quinine pellet presentation between blocks of 10 trials in each arm. Consequently, the net gain ratio between advantageous and disadvantageous options is 2.67. The distribution of punishments is randomized per 10 trials and disadvantageous options are not “stacked” toward the end of the task (see Fellows and Farah, 2005). Control for non-specific exploration is assessed by entries into the two empty arms. A mouse version of the IGT has also been designed by van den Bos et al. (2006), using an eight-arm radial maze. This task presents main characteristics that similar to those of the rat gambling task as described above.

Recently, an operant version of RGT\textsuperscript{reward or quinine} has been developed using specially designed operant panels that are placed in the home cage of rats (Koot et al., 2009, 2010). Next to measuring decision-making, this task allows assessment of impulsivity (Koot et al., 2010) and is currently being further validated.

PREFERENCE FOR HIGH VS. LOW PROBABILITIES TO GET SIMILAR AMOUNT OF REWARD: A TWO-LEVER OPERANT CHAMBER MODEL (RGT\textsuperscript{reward probabilities})

In 2007, a second rodent task that modeled some features of the IGT was established by Pais-Vieira et al. (2007). This task measures the preference for an infrequent high amount of food compared to a more frequent, lower amount of food, both options leading to an almost similar amount of total reward. Another task was recently established, based on the same principle (Roitman and Roitman, 2010), but with a lower level of unpredictability.

The apparatus (Figure 3) consists of an octagonal arena connected to a starting corridor through a guillotine door. The arena is divided in a right and a left side by a central separator, and each half of the arena is provided with one food cup and one lever connected to an automated pellet dispenser. Training consists of two phases: first, animals undergo a series of sessions to get familiarized with the testing apparatus and to learn the association between lever presses and food delivery. Subsequently, rodents are subjected to five sessions of 30–90 trials in which the outcome values are identical for both sides of the arena (one food pellet in 8 out of every 10 consecutive presses). These sessions serve as a control for individual spatial bias: animals preferring one side of the arena in more than 70% of the trials in two of the five sessions are removed from the study. Each trial begins with the animal in the starting corridor; upon lifting of the guillotine door the animal chooses between going to the left or the right side of the arena and presses the corresponding lever. Lever pressing results in either the delivery of chocolate flavored sucrose pellets or no food delivery. Subsequent presses of the levers have no effect (retractable levers have also been used without any change in task performance). Each trial lasts 20 s after which the animal is hand-removed from the arena and placed back in the closed starting corridor. Then, a new trial will start after a variable inter-trial interval (ITI) of 5–10 s to prevent behavioral adjustment to the session limit. Trials in which the animal fails to press a lever within the 20-s are counted as “incomplete.”

Final RGT\textsuperscript{reward probabilities} testing consists of a single probe session of 90 trials. Reward contingencies associated with the levers are set so that one leads to frequent small rewards and the other to infrequent large rewards. Both levers have non-rewarded trials and both lead to similar long-term gains in a pseudo-random order, levers being counterbalanced between animals. Options are not “stacked” to favor reward delivery at the beginning of the session (see Fellows and Farah, 2005). During this test session, one lever remains at the reward settings of the training (one food pellet in 8 out of every 10 consecutive presses) while the other lever is set for rewarding three food pellets in 3 out of every 10 consecutive presses. The maximum gain of pellets per 10 trials is comparable.
for the low and high risk levers [8 vs. 9; but an equal maximum gain of nine pellets for both levers does not change task performance (Ji et al., 2010)]. As in the human IGT, control animals display a preference for the lever with infrequent large rewards in the beginning of the session, but shift their preference to the lever with frequent small rewards in the second half of the test.

**ONE SESSION CHOICE FOR HIGH REWARD/IMPROBABLE LONG PENALTIES VS. LOW REWARD/IMPROBABLE SHORT PENALTIES: THE FOUR-HOLE OPERANT CHAMBER MODEL (RGTONE SESSION REWARD AND TIME-OUT)**

Rivalan et al. (2009a) recently proposed an alternative automated rodent IGT in an operant chamber, with the goal to test complex decision-making processes within a single session, as in the original human IGT. This model offers the advantage to allow a rapid assessment of the decision-making process, measuring the time-course of decisions, from random choice to a majority of choices for the preferred options, within 1 h. Because decision-making can be measured in only one session, this task is particularly suitable for quickly identifying individual differences, or for the search of the neural bases of the time-course process of decision-making, using cellular brain imaging or PET scanning. Performances in this task are stable and reproducible (Rivalan et al., 2009a, 2011). This task can be solved by the majority of rats, whereas some poor decision-makers that prefer larger immediate rewards despite suffering large losses can be identified, a result also observed in humans (Bechara et al., 1999, 2001; Bechara and Damasio, 2002; Davis et al., 2007; Glicksohn et al., 2007).

The test requires the rat to deduce, by trial and error, among four options, the two that are the more rewarding on the long-term and tracks the continuous and dynamic process of deduction and readjustment of choice. The principle of the task tightly mimics that of the human IGT: the contingencies are arranged so that the two options (holes chosen by nose-poking) that steadily offer bigger immediate food reward, are disadvantageous in the long run due to higher unpredictable penalties (frustrating time-outs during which no reward can be obtained). Inversely, the two advantageous options steadily offer smaller reward, but unpredictable penalties that can follow are short.

The testing apparatus (Figure 4) consists of an operant chamber lightly adapted from a standard five-hole operant chamber usually used for the five-choice serial reaction time task (Imetronic, Pessac, France). The adaptation consists of blocking the access to the central hole by a panel, and adding a transparent vertical partition containing a central opening that divides the chamber in half to allow an equal distance to each hole. The four holes, that can be dimly illuminated, are available on a curved wall, with a food dispenser on the opposite wall. The holes and food magazine are equidistant from the central opening.

Rewards are represented by the delivery of palatable food pellets (TestDiet, formula P), while punishments are associated with time-out periods, during which nose-pokes are inactive. As in the original IGT, each selection is rewarded, and some are also unpredictably punished. Similarly to RGT (reward probabilities) training for the acquisition of the basic operant responses is performed before testing. During these daily sessions (30 min each, repeated until a learning criterion is obtained within a session) animals are trained to associate two consecutive nose-pokes (ensuring voluntary choice performance) in any of the four illuminated holes with the delivery of one food pellet. Once the learning criterion is reached, animals are habituated to variable reward amounts (one/two pellets) per selection (two 15-min block per reward amount). Although a preference for a side could be observed individually, it has repeatedly been shown that this preference had no consequences on performances during test, nor differences in the level of exploration. Moreover, because palatable food pellets are used, food restriction facilitate training but is not necessary for testing (Rivalan et al., 2009a). During the subsequent test session (1 h of duration), animals are again requested to freely choose among the available options, but advantageous and disadvantageous selection schedules are now introduced. On one side of the curved wall, advantageous options are associated with the immediate delivery of only one pellet, that can eventually be followed by short time-outs (“deck C”: 12-s time-out 25% of the trials; “deck D”: 6-s time-out 50% of the trials). These two choices result in the same theoretical maximum gain of about 300 pellets in 1 h.

On the other side of the curved wall, disadvantageous options correspond to two response holes that always provide two pellets as a reward upon selection, but that can eventually be followed by
long time-outs (“deck A”: 222-s time-out 50% of the trials; “deck B”: 444-s time-out 25% of the trials). These two options have the same theoretical maximum gain of about 60 pellets in 1 h. Punishments are assigned to each selection in a pseudo-random manner, so to maintain immediate outcome unpredictability, thus, options are not “stacked” to favor reward delivery at the beginning of the session (see Fellows and Farah, 2005). A choice also results in the deactivation of all stimulus-lights except for the chosen hole, until the reward is collected. This is particularly important during the time-outs to facilitate the association between hole response and its consequences. Based on the reinforcement schedule employed during testing, choosing either advantageous or disadvantageous options during this session would result in a theoretical overall payoff ratio of 5.00 (300 or 60 pellets per response hole, considering a standard trial duration of 9 s). The task difficulty can be easily modified, i.e., decreasing the ratio to 3 by increasing time-out duration associated with favorable options, that lead to a slower decision-making process (see Rivalan et al., 2009a).

**MULTIPLE SESSION CHOICE FOR CONFLICT BETWEEN OPTIONS VARYING IN REWARD AND TIME-OUT FREQUENCY AND DURATION:**

**THE FIVE-HOLE OPERANT CHAMBER MODEL (RGTREWARD OR TIME-OUT)**

This rodent gambling task also signals loss on non-rewarded trials through the presentation of frustrating time-out periods, and uses sucrose pellets as a reward. This model presents the animals with a choice between four distinct options on each trial which are loosely analogous to the four deck of cards in the IGT, i.e., the options differ in terms of the probability and magnitude of expected gains and losses, such that the two options which deliver the smaller units of reward are ultimately advantageous. However, this model differs both in the duration of training, the reinforcement contingencies associated with the different options and the way in which the task is learned. Although both the RGTone session reward and time-out and RGTreward or time-out are complementary, they also have different strengths and can be best used to answer different questions. Notably, the RGTreward or time-out model was designed in a manner that could be optimized for behavioral pharmacology experiments and other manipulations, which benefit from repeated daily testing.

Testing takes place in standard five-hole operant chambers available from a number of vendors (e.g., Med Associates, Coulbourn; the same chambers used for the five-choice serial reaction time task; Figure 5). The defining feature of such boxes is that an array of five stimulus—response holes is located on one wall of the chamber, although only four holes are used during the task. Each response hole can be illuminated by a stimulus light located therein, and nose-poke responses into a hole are detected by an infrared sensor. A food tray, also equipped with an infrared sensor and a tray light, is located in the middle of the opposite wall, into which sucrose pellets can be delivered via an external pellet dispenser. The entire chamber can also illuminated using a house-light.

As is often standard for operant-based tasks, animals are first trained to make the basic operant response, in this case to make a nose-poke response into a single illuminated response hole within 10 s. Animals are then trained on a forced-choice version of the task for five to seven sessions, during which only one of the four possible options is presented on each trial (Zeeb et al., 2009; Zeeb and Winstanley, 2011). This training stage ensures that all animals equally experience each of the four reinforcement contingencies, in order to prevent simple biases toward a particular hole from developing. Although this forced-choice contingency does not occur in the IGT, subjects are verbally instructed that there are decks which result in greater losses than others and that winning may be accomplished if the worst decks are avoided (Bechara et al., 1999).

Following forced-choice training, all animals are then tested on the free-choice task, where all four options are presented (i.e., all four response holes are illuminated). Training occurs once daily, and each session lasts for 30 min. Animals initiate each trial by making a nose-poke response into the illuminated food tray. This response initiates a 5-s ITI, where all lights are off and the chamber is in darkness. A nose-poke response in any hole of the array during this time is termed a premature response – a measure of motor impulsivity – and is signaled by illumination of the house-light for 5 s. The animal must then re-start another trial (by making a nose-poke response in the illuminated food tray). This measurement is analogous to that of the premature responses measured in the five-choice serial reaction time task (Robbins, 2002). Therefore, for the first time, both motor impulsivity and decision-making can be concurrently assessed – and dissociated – within the same task (Zeeb et al., 2009).

At the end of the ITI, four holes are concurrently illuminated (left to right: holes 1, 2, 4, and 5) for a maximum of 10 s. The animal signals its preference by nose-poking in one of the illuminated holes, at which point all the stimulus-lights are then extinguished.

---

[Figure 5](#): The standard five-hole operant chamber, as used for the RGT reward or time-out model. (A) Side view of a Med Associates five-hole box. These chambers are modular and can be configured for multiple behavioral paradigms and manipulations, hence the arm assembly for drug self-administration/microdialysis and additional levers visible. However, these components are not necessary for the RGT. (B) A close-up of the standard five-hole array showing stimulus light location and infrared beam.
If the trial is rewarded, the tray light illuminates and the corresponding number of sucrose pellets are delivered immediately. As in the five-choice serial reaction time task, responding at the food tray to collect reward also initiates the next trial. However, if the trial is a loss, no reward is delivered and a time-out period begins during which the light in the hole chosen flashes at a frequency of 0.5 Hz. The animal is unable to initiate any more trials or earn reward until the time-out is over, at which the animal can initiate the next trial by responding in the now-illuminated food tray. The reinforcement contingency linked to each option (response hole) varies in terms of both the number of sugar pellets available, and the duration and probability of a punishing time-out period. The probability of receiving reward or punishment remains constant throughout each session, i.e., the options are not “stacked” to favor reward delivery at the beginning of the session (see Fellows and Farah, 2005). The spatial location of these options are balanced, such that one advantageous and one disadvantageous option is located on each side of the chamber, therefore the correct strategy cannot be reduced to a side bias. The optimal strategy is to select the two-pellet option (P2), associated with a 10-s time-out period that occurs 20% of the time (80% chance of winning). The next best option is P1 (5 s time-out, 90% chance of winning). The two highly disadvantageous options are both associated with larger immediate gain – three or four sucrose pellets – but also longer time-out periods (P3: 30 s time-out, 50% chance of winning; P4: 40 s time-out; 40% chance of winning). Occurrence of gains and losses on each trial is determined pseudo-randomly to ensure a constant distribution of gains and losses throughout each session.

It is possible to calculate the hypothetical amount of reward that could be obtained if an option is chosen exclusively in a 30-min training session, and if each trial is initiated and performed as quickly as possible (i.e., within 5 s; Zeeb et al., 2009). Based on these calculations, persistent selection of the optimal choice (P2) yields a hypothetical maximum of 411 sugar pellets, whereas P1 yields 295 pellets. The two disadvantageous options are associated with significantly less possible reward (P3: 135 pellets; P4: 99 pellets). Therefore, the optimal strategy is to prefer the advantageous options – P2 and P1 – which are associated with smaller, immediate gain, but also less punishment resulting in more reward in the long-term, while avoiding the tempting, yet disadvantageous, large reward options associated with greater loss – P3 and P4.

**RGT Validity**

**Criteria for the validity of an animal model**

The goal of an animal model of human decision-making is to develop, in laboratory animals, behavior that closely resembles that observed in human subjects, thereby allowing researchers to translate findings across species. The validity of animal models depends on the strict definition of its potential applications, taking into account the models’ biases and limits. This validity is commonly assessed using the concept of face, construct and predictive validities (McKinney and Bunney, 1969; Willner, 1995). Face validity refers to similarities between animals and humans in symptomatology, construct validity concerns similarities in underlying psychobiological and physiological processes, and predictive validity concerns the model’s potential to predict these processes in human, most often with regards to identifying efficient pharmacological compounds in humans. With respect to face validity, numerous studies using the IGT in both healthy and patient populations provide knowledge regarding task performance and distinct behavioral patterns of impairment that allows direct comparison to response patterns in the rodent tasks. However, when it comes to construct and predictive validities, current knowledge from pharmacological or imaging human studies is incomplete, making these attributes more difficult to assess. The gap in our current understanding of the neurobiological underpinnings of decision-making in fact underlines the need for animal models that allow more thorough and controlled investigation of these mechanisms.

**Face Validity of the RGTs**

**Response patterns in RGTs**

In all four RGTs, despite the considerable differences in task characteristics, animals were able to evaluate which options are advantageous in the long-term, and to adapt their choice behavior accordingly. Although the end point of behavioral testing differed across tasks, from decision-making measure during a single session to a multiple session learning process, most rats finally developed a significant preference for the more advantageous options (van den Bos et al., 2006; Pais-Vieira et al., 2007, 2009; Homberg et al., 2008; Zeeb et al., 2009; de Visser et al., 2011b; Zeeb and Winstanley, 2011), as also observed in C57BL/6 mice employing the eight-arm radial maze model (van den Bos et al., 2006) and mice performing a version very similar to RGT reward or time-out (Young et al., 2011).

In RGT reward or quinine rats start off with an explorative search profile, displaying equal preference for either the advantageous or disadvantageous arm. After 60–80 trials, a stable preference for the advantageous arm emerges. Rats’ performance of RGT one session reward and time-out indicates that, although most animals learn the contingencies and gradually develop a preference for the advantageous options (Rivalan et al., 2009a), about 40% of the animals failed to solve the task, either because they did not develop any option preference, or because they developed a significant preference for disadvantageous options. These inter-individual differences are stable across time, and reproducible across groups. Interestingly, failure to solve the IGT by a portion (20–30%) of healthy humans has also been observed (Bechara et al., 1998, 2001, 2002; Crane and van der Molen, 2004; Dunn et al., 2006). These individuals have been characterized as impulsive and sensation-seekers (Davis et al., 2007; Franken et al., 2008), some behavioral traits that are related to those of poor decision-makers in rats, i.e., risk-taking and sensitivity to reward (Dellu et al., 1993, 1996; Rivalan et al., 2009a).

In RGT reward or time-out, although an initial preference for the advantageous options is already established in the initial training sessions, this preference continues to develop as testing proceeds. In this experimental design, learning of options’ contingencies is made prior the decision-making test, by systematically exposing animals to the task contingencies during forced-choice training sessions (Zeeb et al., 2009). This ensures that all animals have equal exposure to the different reinforcement contingencies associated with all the options, and theoretically prevents the development of any biases due to inadequate sampling. As a result, it can be argued that performance in this task predominantly captures the second phase of the IGT when contingencies are known (exploitation), particularly later in training.
when performance has stabilized. However, the fact that choice between the options still varies between sessions earlier in training implies that learning (exploration) still occurs. Furthermore, recent evidence suggests that acquisition and performance of this task is controlled by somewhat dissociable neural circuitry (Zéeb and Winstanley, 2011) which supports the distinction between brain areas involved in exploration vs. exploitation outlined in Section “Neural Substrates.”

In the RGT,

- the outcome difference between advantageous and disadvantageous options is minimal (8 vs. 9 pellets), an inconsistency with respect to the human IGT where choice options have marked long-term outcome differences. As such, this task may be more similar to models of probability discounting (e.g., St Onge and Floresco, 2009b) and other aspects of decision-making under uncertainty, rather than the IGT. Nevertheless, control rats preferentially choose smaller but more reliable rewards over larger unreliable rewards, similar to the choice pattern observed in humans. In summary, in all four RGTs animals are able to learn which options are advantageous in the long-term, but the nature of the learning patterns differs between the tasks. This source of variation could be used to address more specifically one aspect of a deficit in animal models of psychiatric conditions, as observed in mental disorders. For example, preference for disadvantageous options, as observed in schizophrenia, OCD, pathological gambling, substance abusing individuals, or ADHD, can be tested in RGTs with a high ratio between advantageous and humans, sex differences emerge in the second part of the task (2004; van den Bos et al., 2007; de Visser et al., 2010). Both in rats and decision-making performance has been investigated in both healthy and clinical populations exhibiting poor decision-making, reward hypersensitivity appears to underlie deficits in IGT performance (Must et al., 2006; Davis et al., 2007; Suhr and Tsanidis, 2007; Kobayakawa et al., 2010). Furthermore, abnormal levels of risk-taking and hypersensitivity to reward are found in psychiatric disorders associated with poor decision-making and impulsivity, such as ADHD, substance abuse, pathological gambling, or mania (Mazas et al., 2000; Bechara et al., 2001; Drechsler et al., 2008; Kathleen Holmes et al., 2009). Hence, it could be argued that poor decision-making in both humans and rats may stem from common risk factors to develop these disorders.

In a recent study, high levels of anxiety as measured with standard parameters in the elevated plus-maze (% time and visits in open arms) has been associated with poor decision-making using the RGT, reward or quinine model (de Visser et al., 2011b). Based on a detailed analysis of rat choice strategies, it was suggested that highly anxious subjects may have a shifted bias toward responding to the negatively valued stimuli, i.e., the quinine pellets, and under-appraised the positively valued stimuli, i.e., the sucrose rewards, leading to suboptimal decision-making. These findings are in line with human data, in that highly anxious healthy subjects performed worse on the IGT compared to their less anxious counterparts (de Visser et al., 2010). These reports also highlight the potential confound in the human data, in that both high anxiety (preferentially measured in the experimental paradigm of de Visser et al., 2011b), and high risk seeking (preferentially measured by Rivalan et al., 2009a) are associated with poor IGT performance, an issue that animal models may be able to resolve, and future work will no doubt address the biological basis of these findings.

Taken together, RGT performance, similar to the IGT, is associated with inter-individual differences related to sex differences and behavioral traits. The study of these inter-individual differences provide a reliable and valuable tool to investigate, for instance, the neurobiological features of subjects at risk to develop mental disorders related to poor decision-making.

**Construct Validity of the RGTs**

Construct validity of the RGTs can be evaluated based on the partial knowledge of the neurobiological mechanisms underlying IGT performance in humans. Key brain areas have been identified from lesion and imaging studies that include parts of the prefrontal cortex, the ventral striatum, and limbic structures such as the amygdala (see Neural Substrates). Furthermore, gene polymorphisms related to serotonergic and dopaminergic systems have been identified that modulate IGT performance and affect the function of corticolimbic neural circuits. Thus, construct validity of the RGTs will be discussed based on these findings (see Table 2 for overview).
Table 2 | Effect of manipulations on rodent IGT performance.

<table>
<thead>
<tr>
<th>Manipulation</th>
<th>RGT reward or quinine</th>
<th>RGT reward probabilities</th>
<th>RGT one session reward and time-out</th>
<th>RGT reward or time-out</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C-FOS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orbitofrontal cortex</td>
<td>No dissociation between good and bad performers (de Visser et al., 2011b)</td>
<td>Dissociation good and bad performers Fitoussi et al. (in preparation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial prefrontal cortex</td>
<td>Dissociation good and bad performers (de Visser et al., 2011b)</td>
<td>Dissociation good and bad performers Fitoussi et al. (in preparation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ventral striatum</td>
<td>Dissociation good and bad performers (de Visser et al., 2011b)</td>
<td>Dissociation good and bad performers Fitoussi et al. (in preparation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LESIONS/INACTIVATIONS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orbitofrontal cortex</td>
<td>Increased choice of higher risk option during exploitation phase</td>
<td>Perseverative responding (Rivalan et al., 2011)</td>
<td>Lesion before acquisition: delayed development of preference for the correct option.</td>
<td>Lesion after acquisition: no effect (Zeeb and Winstanley, 2011)</td>
</tr>
<tr>
<td>Prelimbic cortex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior cingulate cortex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial prefrontal cortex</td>
<td>Impaired task-progression during exploitation phase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amygdala</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GENETIC ALTERATIONS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SERT knockout rat</td>
<td>Increased choice of advantageous option during exploitation phase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAT knockout mouse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PHARMACOLOGY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-OH-DPAT (5-HT1A agonist)</td>
<td></td>
<td>Less advantageous choices (Zeeb et al., 2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SKF 81297 and quinpirole or bromocriptine (D1, D2/3 agonists)</td>
<td></td>
<td>No effect (Zeeb et al., 2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eticlopride (D2 antagonist)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amphetamine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBI27914 (CRF1 antagonist)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Intra amygdala administration (BLA) reversed poor performance (Ji et al., 2010)
Serotonergic system and RGT
It is possible to draw a parallel between performances of rats lacking the SERT in the RGT\textsubscript{reward} or quinine (Homberg et al., 2008) and the 5-HTTLPR (serotonin-transporter-linked polymorphic region) IGT studies in humans. Homozygous (SERT\textsuperscript{−/−}) and heterozygous (SERT\textsuperscript{+/−}) knockout rats are suggested to model stressed and unstressed 5-HTTLPR s-allele carriers, while wild-type control rats may correspond to 5-HTTLPR l-allele carriers (Kaluff et al., 2009; Homberg and Lesch, 2010). SERT\textsuperscript{−/−} and SERT\textsuperscript{+/−} rodents show gene-dose-dependent increases in extracellular serotonin levels due to reduced serotonin reuptake (Homberg et al., 2007). It was found that SERT\textsuperscript{+/−} and SERT\textsuperscript{−/−} rats demonstrated better decision-making compared to SERT\textsuperscript{+/+} animals, particularly in the second half of the trials. This seemingly contrasts findings that human 5-HTTLPR s-allele carriers performed worse during the task compared to l/l subjects (Must et al., 2007; da Rocha et al., 2008; Homberg et al., 2008; van den Bos et al., 2009b; He et al., 2010). If serotonin modulates vigilance and is responsible for the integration of relevant environmental stimuli (Branchi, 2010; Homberg and Lesch, 2010), this discrepancy may be reconciled by the fact that RGT\textsubscript{reward} or quinine employs two baited choice options, as opposed to four in the human IGT. In the RGT\textsubscript{reward} or quinine, the SERT\textsuperscript{−/−} and SERT\textsuperscript{+/−} rats may not have been distracted by environmental stimuli, such that they could focus on their long-term goal: obtaining maximum gain. It would be interesting to test this hypothesis in one of the RGT models which require animals to discriminate between four options concurrently, such as the RGT\textsubscript{one session reward and time-out} or RGT\textsubscript{reward or time-out}.

Dopaminergic system and RGT
As discussed earlier, increased choice of the disadvantageous options on the IGT has been reported in bipolar patients. Reduced dopamine transporter (DAT) function has been hypothesized as a contributing factor to bipolar disorder on the basis of both genetic linkage studies (Kelsoe et al., 1996; Greenwood et al., 2001, 2006) and analysis of DAT expression in patients (Horszchitz, 2005). Using a mouse version of the RGT\textsubscript{reward} or time-out, increased choice of the disadvantageous, or risky, options has been observed in mice lacking the DAT (Young et al., 2011). Furthermore, this risky decision-making correlated with specific exploratory activity in a mouse behavioral pattern monitor (BPM), a pattern of motor behavior that is also observed in acutely manic patients (Perry et al., 2009). Hence, this RGT model has proven useful in demonstrating similar behavioral phenotypes in a putative mouse model of bipolar disorder as compared to the clinical condition.

Prefrontal cortex and RGT
Evidence of involvement of the prefrontal cortex in rodent IGT-like decision-making was reported in three RGT studies. OFC lesioned animals preferred the higher risk lever during the second phase of the task in the RGT\textsubscript{reward probabilities} (Pais-Vieira et al., 2007). This is in accordance with impaired decision-making and high risk-taking observed in human patients with damage to the vmPFC (Bechara et al., 2000; Clark et al., 2008), and the relationships between OFC activity and IGT performance in healthy subjects in fMRI studies (Bolla et al., 2004; Lawrence et al., 2009). Theoretically, OFC lesion effect on risk-taking could be expressed in the RGT\textsubscript{reward probabilities} since this task is based on a two-option simple choice as opposed to the RGT\textsubscript{one session reward and time-out} (Rivalan et al., 2011). In this latter task, measure of risk-taking is combined with the capacity to perceive the changes in contingencies between training (all options have the same consequences) and test (four different consequences), a function also associated with OFC (see below). Consequently, OFC lesioned rats exhibited perseverative responding, as if they failed to encode the change in contingency. This was demonstrated by an absence of sampling of the four options at the beginning of the test, and a marked preference for the holes on the side that they preferred during training. Such inflexible responding following OFC lesions has been reported multiple times, particularly in the context of impaired reversal learning (Dias et al., 1996; Schoenbaum et al., 2002; McAlonan and Brown, 2003; Ragozzino, 2007; Rudebeck and Murray, 2008; Kazama and Bachevalier, 2009; Robbins and Arnsten, 2009).

Indeed, the original demonstration that vmPFC-damaged patients were impaired on the IGT used stacked decks, such that no losses were experienced in the first block of trials, and all subjects initially preferred the large reward decks (Bechara et al., 1999). Patients with comparable vmPFC damage were not impaired on a "shuffled" version of the IGT, in which gains and losses were distributed randomly through the decks, suggesting that the original impairment arose from patient’s inability to switch preferences away from the disadvantageous decks once punishments were introduced (Fellows and Farah, 2005). Considerable evidence suggests that this region is critical in generating outcome expectations, and updating them as the reward value of available options changes (for review see Schoenbaum et al., 2009). Notably, the degree of OFC involvement may depend on the level of ambiguity experienced by the subject, i.e., the OFC may be important when individuals are learning the reinforcement contingencies and ambiguity is high. Once uncertainty becomes expected, such that the individual figures out the odds of risk and reward associated with each option, the OFC may play less of a role in maintaining the optimal choice strategy. This hypothesis is supported by recent findings using RGT\textsubscript{reward or time-out} in which OFC lesions performed prior to acquisition of the task slowed learning such that rats took longer to develop a strong preference for the correct option (Zeeb and Winstanley, 2011). However, if lesions were performed once the task had already been learned, then no choice impairment was observed. Furthermore, in RGT\textsubscript{reward or quinine} final task performance was not related to OFC activation, as measured by the levels of expression of the immediate early gene c-fos (de Visser et al., 2011b), again indicating that activity within the OFC is not a critical determinant of decision-making under risk. Conversely, under risk and ambiguity in the RGT\textsubscript{one session reward and time-out}, c-fos expression differentiated between good and poor decision-makers (Fitoussi et al., in preparation).

In addition to the OFC, damage to two areas of the medial PFC (mPFC) were found to affect performance of the RGT\textsubscript{one session reward and time-out}: the prelimbic cortex [a primitive version of the dIPFC of the primate (Vertes, 2006)] and the ACC (Rivalan et al., 2011). Lesions of the ACC mainly delayed good decision-making in this task whereas lesions of the prelimbic cortex either led to an inability to chose between good and bad....
options (undecided behavior) or induced inflexibility in behavior, similarly to OFC lesions. Moreover, c-fos activity in the mPFC was found to differentiate between good and poor performers in the RGT\textsubscript{reward} or quinine and RGT\textsubscript{one session reward and time-out} (de Visser et al., 2011b; Fitoussi et al., in preparation). Interestingly, the rat mPFC shares anatomical and functional homology with the ACC and dlPFC in humans (Uylings and van Eden, 1990; Brown and Bowman, 2002; Uylings et al., 2003). These areas are involved in IGT performance in humans (Ernst et al., 2002; Bolla et al., 2004; Fukui et al., 2005; Lin et al., 2008; Lawrence et al., 2009). The ACC and dlPFC are specifically involved in a negative feedback circuit of cortical control over limbic areas (Ridderinkhof et al., 2004; Bechara et al., 2005). This top-down cognitive control circuit controls decision-making on the basis of reward and punishment (Quirk et al., 2006; Miller and Cohen, 2001; Rogers et al., 2004; St Onge and Floresco, 2009a; Davis et al., 2010) and is suggested to mediate predominantly the second part of the IGT, when a preference for the advantageous decks has developed and performance is mainly characterized by maintenance and exploitation of the advantageous choice strategy (van den Bos et al., 2007). Thus, activation differences in the mPFC in rats in the RGT may be related to a weaker cognitive control system in poor vs. good performers. In line with the aforegoing, inactivation of the mPFC using a mixture of GABA-agonists muscimol and baclofen, hampered task-progression in those animals which already showed task-learning but not in those animals which were still in the exploratory phase of the task (de Visser et al., 2011a). Further experiments are needed to substantiate the role of the mPFC in decision-making using the different RGT models.

**Sub-cortical areas and RGT**

Apart from cortical areas, the ventral striatum was found to be differentially recruited in good vs. poor performers in RGT\textsubscript{reward} or quinine (de Visser et al., 2011b). More specifically, c-fos induced activation was higher in the nucleus accumbens core (NaC) but lower in the nucleus accumbens shell (NaS) in good performers compared to poor performers, suggesting distinct roles for the nucleus accumbens subareas in rat decision-making. As Yin et al. (2008) argued, the NaC may be involved in more advanced decision-making processes than the shell, thus reflecting the more advanced performance of the good decision-makers. Moreover, the NaC has been implicated in impulse control and behavioral flexibility (Cardinal et al., 2001; Christakou et al., 2004; Potthuizen et al., 2005; Floresco et al., 2006), but see (Murphy et al., 2008), which may suggest that good decision-makers are better at developing a behavioral strategy that is directed to the long-term gain of the advantageous option as opposed to the immediate gain of the disadvantageous option.

The amygdala has also been associated with IGT performance, in that patients with bilateral lesions to this brain region showed a very similar pattern of choice on the IGT as vmPFC patients, choosing more often from the disadvantageous decks (Bechara et al., 1999). A similar pattern of choice has recently been observed using the RGT\textsubscript{reward} or time-out following bilateral lesions to the BLA, in that lesions made after animals had acquired the task lead to an increase in preference for the disadvantageous options. Interestingly, lesions made prior to acquisition of the task strongly resembled the effects of OFC lesions made at the same time point, in that both groups of lesioned animals were slower to adopt the correct strategy as compared to sham controls (Zeeb and Win-stanley, 2011). Such data suggest that the OFC and BLA may be working together to optimize choice behavior when the odds of reinforcement are still unclear. These results are in accordance with the finding that the level of ambiguity in choices was positively correlated with the level of activity in the OFC and amygdala (Hsu et al., 2005; Lawrence et al., 2009). The fact that BLA lesions still affected decision-making once the task had been learned suggests that decreased activity in this region may precipitate an increase in risky choice, and that this area is involved in maintaining an optimal decision-making strategy under risk even after the cost–benefit contingencies have been acquired.

In conclusion, neural circuitry comprising the PFC, striatum, and amygdala appears to modulate decision-making in RGT models, and the effects observed are largely consistent with findings in humans. Altered connectivity between cortico-striatal-limbic circuitry has been implicated in various disorders, such as schizophrenia, drug addiction, OCD, and anxiety disorders. The RGTs may therefore contribute to an increased understanding of the pathophysiology of these disorders. Comparing the findings of lesion and c-fos activation studies across the RGT models will help to further validate their use, and may provide new insight into the neural circuitry involved in this form of decision-making. Moreover, the somatic marker hypothesis could be addressed in all the RGTs during the exploration phase of the tasks, i.e., when choices evolve with time before reaching a stable level (exploitation phase). Measures of blood pressure, heart rate, associated with behavioral activity could serve as somatic markers in freely moving rats, by radiotelemetry.

**PREDICTIVE VALIDITY OF THE RGTs**

Given that low serotonin (5-HT) function has been observed in problem gamblers, and that selective 5-HT reuptake inhibitors are currently used as a treatment for this disorder, Zeeb et al. (2009) investigated the effects of acutely decreasing 5-HT release using the 5-HT\textsubscript{1A} receptor agonist 8-OH-DPAT on performance of the RGT\textsubscript{reward or time-out}. This compound impaired performance, significantly increasing choice of less advantageous options (P1, P3) and decreasing choice of the best option (P2). These effects were effectively blocked by co-treatment with the highly selective 5-HT\textsubscript{1A} receptor antagonist WAY100655. However, it remains to be established whether these effects are due to activation of inhibitory 5-HT\textsubscript{1A} receptors located pre-synaptically in the raphe nuclei, or post-synaptically in limbic and cortical brain regions.

In addition to the effects of acute 5-HT manipulations, the effects of receptor-specific dopamine agonists and antagonists have also been explored using RGT\textsubscript{reward or time-out} (Zeeb et al., 2009). Interestingly, D\textsubscript{1} and D\textsubscript{2} receptor agonists (SKF 81297 and quinpirole or bromocriptine) did not affect animals’ choice preferences on the task. However, acute administration of the D\textsubscript{2} receptor antagonist, eticlopride, significantly improved task performance by increasing the animals preference for the optimal option (P2) while decreasing choice of both high reward–high punishment options (P3, P4). This effect appears to be selective to the blockade of D\textsubscript{2} receptors, as administration of a D\textsubscript{1} receptor
antagonist, SCH 23390, did not alter decision-making (Zeeb et al., 2009).

The psychostimulant amphetamine has been shown to prime the motivation to play slot machines in pathological gamblers (Zack and Poulos, 2009). Amphetamine treatment significantly impaired the animals' ability to perform the RGTreward or time-out optimally. However, in contrast to the effects of the 5-HT1A receptor agonist 8-OD-DPAT, amphetamine caused animals to become more punishment/loss-sensitive, illustrated by an increased preference for P1, the option with the smallest amount of immediate reward, but also the least amount (frequency and duration) of punishment. Therefore, amphetamine may amplify animals' sensitivity to punishment (Zeeb et al., 2009). Although this behavior cannot necessarily be classified as "risky," placing too much emphasis on a potential loss may be unfavorable in the long-term (Kuhnen and Knutson, 2005) and can contribute to loss-chasing behavior in real-life gambling situations (see discussion in Campbell-Meiklejohn et al., 2008).

Overall, there are both similarities and discrepancies when comparing these results to human findings. As previously discussed, high levels of prefrontal dopamine and corticolimbic 5-HT levels may correlate with worse IGT performance. Therefore, the fact that amphetamine impaired decision-making on RGTreward or time-out supports the hypothesis that increased levels of dopamine or 5-HT impairs decision-making. However, acutely decreasing dopamine levels in healthy volunteers by BCAA administration causes subjects to choose the disadvantageous options on the IGT, especially during the exploitation phase of testing (Sevy et al., 2006). One explanation provided by Zeeb et al. (2009), is that the effects of various drug manipulations may rely on both the basal levels of dopamine as well as task-related changes in dopamine release. Therefore, the effects of dopamine, and perhaps 5-HT, may follow an inverted U-shaped curve (see Neuro-modulators). Interestingly, it is unclear whether an acute dose of eticlopride is modulating activity of inhibitory D2 autoreceptors and/or post-synaptic receptors. A blockade of D2 autoreceptors may stimulate the firing of dopaminergic neurons, while suppressing dopamine transmission post-synaptically (Seamans and Yang, 2004). The net effect may be an enhancement of prediction error signals, which would thus improve decision-making. It should be noted that high PFC dopamine levels are proposed to enhance exploration in the direction of alternative options that might yield higher gains (Frank et al., 2009); however pharmacological manipulations were performed once animals had been trained on the RGTreward or time-out, and therefore did not assess the ability of these drugs to alter task acquisition.

As amphetamine increased the saliency of the punishment signals in RGTreward or time-out, it may be suggested that this effect was caused by an abnormal increase in dopamine. However, this explanation contrasts with the finding that COMT Met/Met individuals have an attentional imbalance in favor of rewards (van den Bos et al., 2009b). One possible explanation is that amphetamine may be causing rats to become more risk-averse through increases in other neurotransmitters (such as 5-HT). Another possibility is that animals received an acute treatment in the study by Zeeb et al. (2009), whereas the human subjects tested in the study by van den Bos et al. (2009a) was observing the effects of long-term genetic abnormalities. Furthermore, the COMT polymorphism is known to target the prefrontal cortex, whereas an acute dose of amphetamine would have more widespread effects.

5-HT mediates a variety of central processes, such as emotion-regulation, learning and memory, motivation, and behavioral inhibition; these traits may be unified by sensitivity to external and internal environmental stimuli, and integrated in order to facilitate associative learning processes (Branchi, 2010; Homberg and Lesch, 2010). While serotonin may modulate behavioral flexibility (e.g., Borg et al., 2009; Jedema et al., 2009) by such a mechanism, it may have less of a role when a subject has to deal with a myriad of stimuli. However, the 5-HT system is known to be involved in the emotional response to aversive events (Cools et al., 2008). As both the IGT and RGTs require subjects to integrate multiple stimuli (e.g., reward magnitude, probability of reward, punishment duration), 5-HT may play a large role in incorporation the concept of loss, and further research should be conducted to determine the role of acute and chronic manipulations of 5-HT.

LIMITATIONS AND PITFALLS OF THE RGTs

All of the RGT models capture one or more features of the human IGT, and have been validated to varying degrees. However, each of the RGTs also has limitations, and no single task fully captures all factors present in the IGT. If these limitations are taken into account when interpreting the data, the rodent models show great promise in being able to address research questions that cannot be easily studied in humans.

One general limitation of all RGTs is that rewards are represented by food pellets – a primary reinforcer – rather than a secondary reinforcer akin to money in the IGT. Given that rodents foraging for food in uncertain environments and human behavior in decision-making tasks share several features, it could be argued that the use of food as a reward may add to the ethological validity of the tasks. Moreover, utilizing food pellets presents some practical advantages in comparison to other types of rewards, such as the possibility of precise magnitude quantification, easiness in administration, and low impact on general psychological/physical functions (in contrast with psychoactive drugs, such as amphetamine, which mimic the effects of primary reinforcers at central level but also produce major side effects, e.g., the emergence of hyperactivity). However, the incentive value of food reward normally depends on an animal's motivational state (hunger/satiety; Cardinal et al., 2002). Therefore, interpretation of choice behavior in RGT models may be confounded by this highly uncontrolled factor. Although manipulation of the animals' drive for food through the employment of different food deprivation levels has been found to have no impact on decision-making in the RGTone session reward and time-out model (Rivalan et al., 2009a), this has yet to be determined for the other RGT versions. On the other hand, the incentive value of money, as used in human IGTs (see below) is also subjective to individual differences in money-triggered incentive salience, therefore this may be less of a concern.

By far the biggest concern with using food pellets as rewards relates to the accuracy of any of the RGTs to model the concept of loss. In the IGT, subjects materially experience financial "wins" and "losses" every time a selection is made. The probability associated
with each trial of incurring financial penalties appears to be central for IGT performance (Fernie and Tunney, 2006): for instance, a high frequency of losses can lead human subjects to discard decks that are advantageous in the long-term (Chiu et al., 2008; Lin et al., 2009). Given that sugar pellets are instantly consumed, rather than accumulated over time and then eaten, it is impossible to take sugar pellets away from animals once they have been won, i.e., truly reproducing the sensation of loss. As iterated above, in order to model the concept of loss in the RGTs, punishment is accomplished by delivery of quinine-treated instead of normal food items (RGTreward or quinine), the absence of reward (RGTreward probabilities), or delays (RGTone session reward and time-out and RGTreward or time-out). In all cases except for RGTreward probabilities, an actual decrease of a positive reinforcer is achieved, since choosing disadvantageous options has a negative impact on the total amount of food pellets consumed during the test. Nevertheless, employing unpalatable food or delays cannot reproduce an absolute resource deficit as a final outcome.

Gain/loss frequencies associated with each response option can be an important determinant of choice behavior during the IGT (Chiu et al., 2008; Lin et al., 2009). Regarding this task variable in the RGT models here examined, a number of discrepancies with the original human task can be identified. For both the RGTreward or quinine and RGTreward or time-out models, the disadvantageous options are associated with large but less frequent immediate rewards but larger more frequent punishments, while the advantageous options are associated with more frequent, although smaller, gains and infrequent but smaller losses. In contrast, within the IGT, the frequency of punishment delivery is varied within advantageous and disadvantageous options, but does not vary between them, such that there is both an advantageous and disadvantageous options, but does not vary between them, such that there is both an advantageous and disadvantageous options associated with a high and low frequency of punishment; the only variable that differs is the size of the respective rewards and punishments. From this perspective, the RGT model which best captures the reinforcement contingencies present in the IGT is the RGTone session reward and time-out which uses two reward sizes (1 vs. 2 pellets) and two probabilities of a penalty (0.5 and 0.25). This task would therefore be most appropriate for the investigation of the “prominent deck B” phenomenon, in which individuals find it difficult to avoid responding at the high reward deck associated with the lowest probability of punishment. However, it could be argued the probability of receiving a penalty in this RGT version is still higher than that in the IGT (0.4 and 0.1 for decks A and B respectively).

The fact that the reinforcement contingencies in the RGT models are not exactly the same as the IGT may offer some unexpected benefits. For example, in the RGTreward or time-out model, one of the more advantageous options (P1) involves a higher frequency of wins and shorter punishing time-out periods than the most optimal strategy (P2). Choice of P1 yields more reward than the “high risk–high reward” options (P3/P4) and certainly does not represent a “risky” choice, but is nonetheless suboptimal and may reflect risk-averse/overly loss-sensitive decision-making. There is no such component in the IGT, but it may be of interest nonetheless. Likewise, the fact that rats in the RGTreward probabilities model prefer the safe option even though there is only a marginal difference in net gain between the risky and advantageous options, may likewise be informative when considering decision-making biases under uncertainty.

It can be argued for both human and rodent IGT that subjects may perform worse as a result of working memory deficits and therefore have problems incorporating previous outcomes in their subsequent choices. Indeed, humans with impaired working memory were found to perform worse on the IGT (Bechara et al., 1998; Suhr and Hammers, 2010). Furthermore, discrimination learning, reversal learning and attentional capacities, punishment and reward sensitivity may all affect performance but are difficult to dissect within the IGT. In rodents, specific tasks have been widely employed to address these processes and may be combined with RGT to elucidate how different learning and decision-processes are interwoven in a complex decision-making task like the IGT.

**CONCLUSION**

Poor decision-making is a core deficit of major psychiatric disorders, and the identification of the underlying neural mechanisms will importantly advance both the diagnosis and treatment of these disorders. Animal models of affective decision-making provide an important tool in achieving this goal. In this review we discussed four RGTs that model specific aspects of the human IGT. In all these tests animals appear to use strategies that resemble those used by humans. That is, initially the animals explore the different choice options, but thereafter show a consistent behavioral pattern in which they attain to a given strategy. Lesion and immunohistochemistry studies have thus far shown that RGT performance is modulated by a similar neural circuitry as in humans, involving parts of the PFC, the nucleus accumbens, and BLA. Factors like the perceived value of the wins and losses, probability and time, and the integration of this information by neuromodulators dopamine and serotonin play an important role in guiding choice.

Because carefully controlled longitudinal studies in humans are hampered by practical issues, the RGTs provide new opportunities to investigate to what extent pre-existing changes in decision-making predict the development and treatment outcome of psychopathologies under influence of genes, stress, or the availability of drugs of abuse (Potenza, 2009). Although prevention is hard to achieve, even when risk factors have been identified, improvements in diagnosis may aid in the design of individualized therapies. Obviously, much remains to be done before we are able to use RGTs for this purpose, but the growing interest, and concomitantly, the development and validation of RGTs, provide heuristically useful data. While the available RGTs need further validation, one option that should be considered is whether the different RGTs should be used in parallel, or integrated into a more uniform model to investigate the factors and mechanisms associated with impaired decision-making. One potential advantage to maintaining the different models is that they all have distinct strengths. Together, comparisons between results obtain in the different RGTs are expected to provide significant contributions to our understanding of a broad range of neuropsychiatric illnesses, which will be of great benefit to society.


cognition of chronic amphetamine abusers, opiates abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms.

Neuropsychopharmacology 20, 322–339.


knockdown mice: further support for a mouse model of mania. J. Psychopharmacol. 25, 934–943.


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 June 2011; paper pending published: 28 June 2011; accepted: 29 August 2011; published online: 10 October 2011.


This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.

Copyright © 2011 de Visser, Homberg, Mitsogiannis, Zeeb, Rivalan, Fitoussi, Galhardo, van den Bos, Winstanley and Dellu-Hagedorn. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.