Decision-making during gambling: an integration of cognitive and psychobiological approaches

Luke Clark*

Department of Experimental Psychology, Behavioural and Clinical Neuroscience Institute, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

Gambling is a widespread form of entertainment that may afford unique insights into the interaction between cognition and emotion in human decision-making. It is also a behaviour that can become harmful, and potentially addictive, in a minority of individuals. This article considers the status of two dominant approaches to gambling behaviour. The cognitive approach has identified a number of erroneous beliefs held by gamblers, which cause them to over-estimate their chances of winning. The psychobiological approach has examined case-control differences between groups of pathological gamblers and healthy controls, and has identified dysregulation of brain areas linked to reward and emotion, including the ventromedial prefrontal cortex (vmPFC) and striatum, as well as alterations in dopamine neurotransmission. In integrating these two approaches, recent data are discussed that reveal anomalous recruitment of the brain reward system (including the vmPFC and ventral striatum) during two common cognitive distortions in gambling games: the near-miss effect and the effect of personal control. In games of chance, near-misses and the presence of control have no objective influence on the likelihood of winning. These manipulations appear to harness a reward system that evolved to learn skill-oriented behaviours, and by modulating activity in this system, these cognitive distortions may promote continued, and potentially excessive, gambling.

Keywords: gambling; addiction; risk; reward; cognition; emotion

1. INTRODUCTION

The term gambling refers to a form of entertainment where a wager, typically a sum of money, is placed on the uncertain prospect of a larger monetary outcome. As a form of recreation, gambling has been widespread for several centuries, and across many cultures (Raylu & Oei 2004). The 2007 British Gambling Prevalence Survey found that 68 per cent of respondents reported gambling at least once in the past year, and 48 per cent reported gambling on games other than the state lottery (Wardle et al. 2007). To economists and psychologists, the popularity of gambling represents an enduring paradox, as the vast majority of gamblers are well aware of the popular saying 'the house always wins'. This refers to the fact that gambling odds are carefully arranged to ensure a steady profit for the bookmaker, casino or slot machine; something that can only be achieved at the expense of the gambler. In economic terms, the expected value of gambling is negative, such that an accumulating debt is inevitable over a large number of trials. Thus, the widespread tendency to accept such gambles may provide some useful insights into the mechanisms of human irrationality. However, in addition to the financial considerations, it is probable that gambling is also motivated by cognitive and emotional factors. Unpredictable monetary wins are a potent form of positive reinforcement that strengthen the instrumental response. Gambling is associated with physiological arousal that is manifested in heart rate increases and elevated cortisol levels (Anderson & Brown 1984; Meyer et al. 2004). Environmental cues (e.g. flashing lights, the chime of coins) that are associated with this arousal become conditioned stimuli via Pavlovian processes. Gambling may also serve to alleviate unpleasant states of boredom, anxiety or low mood (i.e. negative reinforcement). These emotional learning mechanisms will play a key role in shaping gambling behaviour (Blaszczynski & Nower 2002).

Gambling is also a behaviour that can spiral out of control in some individuals. As gambling becomes excessive, there are observable harms including debt, illegal activity and interpersonal conflict. In its most extreme form, pathological gambling is a recognized psychiatric diagnosis in the Diagnostic and statistical manual, version 4 (text revision) (DSM-IV-TR; American Psychiatric Association 2000), with a prevalence of around 1 per cent (Petry et al. 2005). The severity of gambling involvement is thought to lie on a continuum, and the label of 'problem gambling' is used to denote the less severe form. The US prevalence of problem gambling is estimated between 1 and 4 per cent (Shaffer et al. 1999; Welte et al. 2002).
The current psychiatric system places pathological gambling within the impulse control disorders, a heterogeneous ‘rag-bag’ of conditions that also includes kleptomania (compulsive stealing) and trichotillomania (compulsive hair-pulling). Accumulating data point to a re-alignment of pathological gambling within the addictions (Potenza 2006). The diagnostic criteria themselves were closely modelled on the features of substance dependence, and there is evidence of cravings (Tavares et al. 2005), withdrawal symptoms (Wray & Dickerson 1981) and tolerance (Griffiths 1993b) in severe gamblers. In addition to clinical phenomenology, several other lines of evidence indicate aetiological overlap between problem gambling and drug addiction: there is substantial comorbidity between the conditions (Petry et al. 2005), shared genetic liability (Slutske et al. 2000), and prospective data identify personality traits that predict the development of both problem gambling and substance use disorders (Vitaro et al. 1999; Slutske et al. 2005). The critical difference is that problem gambling does not involve the ingestion of a psychoactive substance. Long-term drug administration causes an array of changes in the brain, so that in current users, it is difficult to disentangle the mechanisms by which the addiction developed. As a putative ‘behavioural addiction’, problem gambling may represent a model for studying addiction vulnerability, in brains that are not confounded by the damaging effects of drugs (Bechara 2003).

Research into gambling behaviour can therefore address two broad issues. First, given the general prevalence of this behaviour, what does gambling tell us about the fallibility of decision-making mechanisms in the healthy human brain? Second, from a clinical perspective, how does this common recreational behaviour become dysfunctional? An overarching theory of gambling should be able to explain both its general popularity, and its potential to become pathological. The aim of the present article is to integrate two approaches to gambling behaviour that have gained considerable popularity in recent years, but which are rarely linked and command quite separate research literatures. The cognitive approach emphasizes thought content and a distorted appraisal of control during gambling. The psychobiological approach assumes a disease model of problem gambling, and has sought to identify group differences between pathological gamblers and healthy controls on measures of brain chemistry and brain function. I will provide an overview of the current status of each approach, before reviewing recent findings that suggest a synthesis of the two approaches may be warranted.

2. THE COGNITIVE APPROACH

The cognitive formulation of gambling argues that the problem gambler continues to play because they possess distorted beliefs about gambling that cause them to overestimate their chances of winning (Ladouceur & Walker 1996). Several kinds of erroneous beliefs have been identified (Toneatto et al. 1997; Raylu & Oei 2004a), which ultimately give rise to an ‘illusion of control’ where the gambler confuses a game of chance with a game of skill (Langer 1975; Thompson et al. 1998). In games where there is some genuine skill involvement, such as blackjack, the gambler comes to believe that skill is excessively influential (Ladouceur & Walker 1996). In believing that they are acquiring the necessary skills to win (or even that such skills exist in principle), the gambler is able to justify continued play.

Much of the evidence for the cognitive approach has used the ‘think aloud’ procedure developed by Gaboury & Ladouceur (1989). In this paradigm, the gambler is asked to verbalize all thoughts during a brief period of gambling in a naturalistic setting, such as a casino. They are encouraged to speak continuously and to avoid censoring their speech. Their speech output is recorded by the experimenter, and statements are categorized subsequently as accurate (e.g. ‘It’s a machine, we have no control over it, it’s all luck’) or erroneous (‘I’m getting good at this game. I think I’ve mastered it’; Ladouceur & Walker 1996). In regular gamblers, around 70–80% of strategic statements about the game were erroneous, with similar figures obtained in slot-machine players and roulette players (Gaboury & Ladouceur 1989; Walker 1992). High rates of erroneous thoughts were even present in players who were clearly aware that the outcomes were determined by chance, given their responses on a questionnaire administered before and after the gambling session.

While these erroneous thoughts are evident in infrequent and controlled gamblers, one tenet of the cognitive approach is that cognitive distortions are exacerbated in problem gamblers, and are used to justify ongoing excessive play (Ladouceur & Walker 1996). A number of studies support this (Walker 1992; Griffiths 1994; Baboushkin et al. 2001; Joukhador et al. 2003). Using the think-aloud procedure, Griffiths (1994) found that regular (at least once per week) fruit machine players reported more erroneous thoughts than non-regular players (less than once per month). Baboushkin et al. (2001) found that university students classified as probable pathological gamblers on the widely used South Oaks gambling screen (SOGS; Lesieur & Blume 1987) reported more erroneous thoughts during computerized games of roulette, blackjack and a slot machine. In addition, a programme of research by Ladouceur et al. (2002) has shown efficacy of a form of cognitive therapy for pathological gambling that aims to correct these erroneous beliefs.

At a psychological level, it is important to understand how these faulty beliefs develop, in both occasional and problem gamblers. There appear to be at least two mechanisms at work. On the one hand, humans are generally poor at processing probability and judging randomness. On the other hand, various features of gambling games directly foster these distorted beliefs. It is widely accepted that humans are highly error-prone at judging probabilities (Gigerenzer 2002). Classic studies from experimental psychology show that people are poor at generating, and recognising, random sequences, such as the outcomes of a series of coin tosses (Tversky & Kahneman 1971; Wagenaar 1972). Subjects prefer sequences without long runs of the same outcome, and with balanced overall frequencies of heads and tails. This may arise because subjects fail to appreciate the independence of turns, and expect small samples
to be representative of the populations from which they are drawn (Wagenaar 1988). Impaired processing of randomness may give rise to the ‘Gambler’s Fallacy’, where the gambler believes that a win is ‘due’ after a series of losses. Cohen (1972; cited in Ladouceur & Walker 1996) looked at betting strategies in roulette players as a function of the previous outcome. Players were more likely to bet on black if the previous outcome was red (75%) than if the previous outcome was black (50%). In a study of university students choosing lottery tickets, it was shown that players preferred tickets of apparently random numbers over tickets containing consecutive numbers (14–19), clusters of numbers (e.g. six numbers between 20 and 30), and numbers involving patterns (16–21–26–31–36–41; Hardoon et al. 2001). Recent work in sport fans has also looked at winning and losing ‘streaks’, arguing that most people perceive a ‘streak’ on the third consecutive win or loss event (Carlson & Shu 2007).

In addition to these generic difficulties in processing chance, various features of gambling games (referred to as ‘structural characteristics’) promote gambling (Griffiths 1993a), potentially via the promotion of cognitive distortions. As a simple example, slot-machine wins are routinely accompanied by bright flashing lights and loud noises. Wagenaar (1988) suggested that this sensory stimulation fuels an ‘availability heuristic’, where the gambler can more easily recall past wins than past losses. By distorting their memory of past outcomes, this may bias the decision to continue play. In the next sections, we focus on two further structural characteristics that appear to manipulate the player’s perceptions of winning in a particularly profound manner.

(a) Personal control
Personal control refers to the gambler’s level of involvement in arranging their gamble. On a game of chance, the gambler is equally likely to win if they arrange their gamble, or if another agent places the gamble for them. For example, in a lottery, one’s favourite numbers are as equally likely to win as a ‘lucky dip’ ticket. However, it has been reliably observed across many forms of gambling that players have inflated confidence when they are given the opportunity to arrange the gamble themselves. In a seminal study by Langer (1975), subjects were invited to buy a lottery ticket, and the experimenter later asked to buy back their ticket. Subjects who were initially able to choose their ticket from a bag demanded more money ($9) to exchange compared with a group who were allocated a ticket at random ($2). In a follow-up experiment, subjects who had chosen their ticket were more likely to refuse a swap for a ticket in a second lottery with a higher chance of winning. This illustrates how perceived control can actually cause subjects to reject a genuine opportunity to increase their chances of winning.

Similar findings have been reported in craps and roulette. In craps, gamblers play in a team where they take turns to throw the dice (‘shooting the dice’) onto the craps table. They can place bets on certain numbers being rolled, on any player’s throw including their own. Regular craps players display a range of superstitious behaviours when throwing the dice, such as blowing on the dice, and using more force in their hand movements when trying to throw a high number (Henslin 1967). Consistent with an effect of personal control, when it is a player’s turn to shoot the dice, they are more likely to place a bet, place higher bets, and place more risky bets compared with when other players are shooting (Davis et al. 2000). Similarly, a study of roulette players found that higher bets were placed when the player was given the opportunity to throw the roulette ball, compared with trials where the experimenter acted as a croupier and threw the ball (Ladouceur & Mayrand 1987). In each of these examples, the presence of personal control has no effect whatsoever on the likelihood of winning.

(b) The near-miss effect
Near-misses occur when an unsuccessful outcome is proximal to a win. They occur across all forms of gambling, such as when a slot-machine payline displays two cherries with the third cherry just coming into view. Near-misses are salient events to the gambler. Reid (1986) found that in student volunteers watching a computerized horse-race, races with a close neck-to-neck finish were rated as ‘better’ than races with a clear winner from early on. Gamblers often interpret near-misses as evidence that they are mastering the game, and in this sense, near-misses appear to foster an illusion of control. As a consequence of the near-misses, the gambler feels that he is ‘not constantly losing but constantly nearly winning’ (Griffiths 1991).

A number of research studies have investigated the behavioural effects of near-miss outcomes on gambling play. In the first study of its kind, Strickland & Grote (1967) used a slot-machine simulation where the three reels stopped sequentially. The reels contained red and green stimuli, and wins were awarded for three reds. One group of subjects played a game where the chances of a red icon appearing on reels 1–3 was 70, 50 and 30 per cent, and hence there was a high likelihood of a near-miss. A second group played the same game but with reels 1 and 3 reversed, so that it was evident early on that the trial was a loss. The actual proportion of wins was matched across the two groups. Subjects in group 1 were seen to play for significantly longer than subjects in group 2. More recent studies have begun to systematically manipulate the frequencies of near-misses. Cote et al. (2003) assigned two groups of subjects to play a slot machine that either delivered no near-misses or a moderate (27%) frequency of near-misses. Subjects in the near-miss condition played significantly more trials on the game. A similar study compared three machines with 15, 30 and 45 per cent frequencies of near-misses, and reported an ‘inverted U’ effect with maximal persistence in the intermediate group (Kassinove & Schare 2001). Clearly, the potency of near-misses is diminished if they are over-represented, rather like ‘crying wolf’.

(c) Summary
The cognitive approach argues that gambling behaviour is maintained by erroneous beliefs and cognitive
distortions about the true chances of winning, such that gamblers perceive the expected value of gambling as positive, when in fact, the objective expected value is negative. The approach is not without its critics, who have argued that the think-aloud procedure is overly intrusive, that flippant verbalizations do not necessarily reflect cognitions held with conviction, and that there are only a limited number of ways that subjects can express accurate cognitions about chance and randomness during a period of gambling play (Dickerson & O’Connor 2006). Nevertheless, the cognitive approach has considerable explanatory power: this framework can capably explain the general prevalence of gambling as erroneous cognitions and inaccurate perceptions of randomness are common in infrequent gamblers. The cognitive framework can also explain the process by which gambling becomes pathological as problem gamblers are hypothesized to make more erroneous cognitions (or to have greater conviction in those beliefs, or to be more inclined to use their faulty beliefs to justify continued gambling). There is some evidence for this hypothesis using the think-aloud procedure (Walker 1992; Griffiths 1994; Baboushkin et al. 2001), although there is minimal work specifically comparing personal control or near-miss effects between problem and non-problem gamblers. In testing these ideas, one complexity is that cognitive distortions in regular gamblers can be highly idiosyncratic (Delfabbro 2004), such that a gambler may view many outcomes as ‘near-misses’ that would appear ‘full-misses’ to a non-gambler.

3. THE PSYCHOBIOLOGICAL APPROACH

The psychobiological approach attempts to identify differences in aspects of brain function between groups of individuals with and without gambling problems. Studies can be divided into those measuring neurotransmitter function, and those measuring the activity or integrity of different brain areas. The latter approach can be subdivided into neuropsychological studies, which measure brain function indirectly using tasks validated in patients with brain injury, and functional imaging studies, which measure brain activity directly during task performance, typically with functional magnetic resonance imaging (fMRI).

(a) Neurochemical studies

Studies of neurotransmitter function in gamblers have focussed on the monoamines, dopamine, serotonin and noradrenaline, which are known to play key roles in arousal, motivation and higher cognitive functions (see Robbins 2000 for a review). It is difficult to measure neurotransmitter levels directly in the human brain. Instead, a number of studies have measured peripheral markers in urine, plasma or cerebrospinal fluid (CSF). These studies reported increases in markers of noradrenaline function (Roy et al. 1988; Bergh et al. 1997), reductions in markers of serotonin function (Nordin & Eklandh 1999) and alterations in dopamine function (Bergh et al. 1997; Meyer et al. 2004). The study by Bergh et al. (1997) reported a decrease in CSF dopamine, coupled with increased levels of the dopamine metabolite, homovanillic acid, from CSF samples obtained in the clinic. The study by Meyer et al. (2004) measured dopamine and noradrenaline levels in plasma during a period of casino gambling in problem and non-problem gamblers. Problem gamblers showed greater increases in both noradrenaline and dopamine levels during casino gambling for real money, compared with a laboratory gambling session for points reward. Thus, the direction of effect—for dopamine changes in particular—remains unclear, and findings from peripheral markers must be treated with caution as their relationship with central activity is complex.

Another indirect approach has been to study genetic variants that are thought to affect neurotransmitter function. For example, the dopamine D2 receptor gene displays a common polymorphism (TaqIA, occurring in A1 and A2 alleles) that influences D2 receptor density in the brain, and is linked to the prevalence of alcohol and stimulant addictions (Noble 2000). Studies by Comings et al. (1996, 1999) reported changes in DRD2 and DRD4 polymorphism frequencies in groups of pathological gamblers, compared with the age, gender and race-matched non-gamblers. The reported TaqIA association (increased prevalence of the A1 allele) is consistent with reduced D2 receptor binding in the striatum in pathological gamblers (Pohjalainen et al. 1998). Genetic studies have also indicated effects on other genotypes affecting serotonin and noradrenaline function (Comings et al. 2001). However, this field has been plagued by failures of replication, and a recent study in siblings discordant for pathological gambling (140 pairs) indicated a significant association with the DRD1 gene but failed to support the DRD2 association (da Silva Lobo et al. 2007).

At least two other lines of evidence converge on the finding that dopamine transmission is altered in problem gambling. A number of case reports have described impulse control disorders, including problem gambling, in patients with Parkinson’s disease, where the primary neuropathology is degeneration of the dopamine system. The emergence of these impulse control disorders appears to be linked to treatment with dopamine agonist medications (Weintraub et al. 2006), and in particular, to two drugs, pramipexole and ropinirole, that have a relatively high affinity for the dopamine D3 receptor (Dodd et al. 2005). The emergence of pathological gambling has been linked to earlier age of onset of the Parkinson’s Disease, comorbid or familial alcoholism, and elevated trait impulsivity and sensation-seeking scores (Voon et al. 2007). However, it is unclear how the primary pathology in Parkinson’s Disease interacts with the action of the medication.

Second, studies by Zack & Poulos (2004, 2007) have looked at the effects of dopamine challenge in problem gamblers, on aspects of gambling behaviour. Their first experiment used amphetamine, an indirect dopamine agonist that also increases noradrenaline transmission. Amphetamine increased motivation to gamble and facilitated the reading of gambling-relevant words in problem gamblers. Their follow-up study used the more selective dopamine D2 receptor antagonist haloperidol, but unexpectedly reported
similar effects to amphetamine: haloperidol increased motivation to gamble and primed gambling-relevant words as well as increasing heart rate responses during a period of slot-machine play. While this study supports the role of the dopamine D2 receptor in gambling behaviour, the direction of effect is problematic from a treatment perspective, as both an indirect agonist (amphetamine) and a selective antagonist (haloperidol) increased gambling tendencies. It is possible that low doses of a dopamine receptor antagonist act preferentially on presynaptic autoreceptors to increase dopamine function (Moghaddam & Bunney 1990; Frank & O’Reilly 2006), and that higher doses of the antagonist would be needed to reduce dopamine transmission.

In summary, neurochemical studies of problem gambling have taken a number of indirect approaches to the measurement of neurotransmitter function. There are preliminary indications of changes in serotonin and noradrenaline function (see also Potenza 2008), and indeed, much reason to think that other transmitters like glutamate may be dysregulated (Grant et al. 2007). The most consistent finding at the current time is for dysregulation of dopamine function in problem gamblers, although the direction and precise mechanisms of this effect remain unclear.

(b) Neuropsychological studies
In a comprehensive review of studies that used clinical neuropsychological tests, Goudriaan et al. (2004) concluded that there was little evidence for impairment in language, perception, intellectual function, and memory in problem gamblers. In contrast, several studies have detected impairments on traditional tests of frontal lobe function; namely, the Wisconsin card sort test, which requires the subject to perform abstract rule shifts, and the Stroop test, which requires the subject to override the automatic tendency to read colour words in order to name the colour of the ink that the word is printed in (Goudriaan et al. 2006a; Kaleschtein et al. 2007; Forbush et al. 2008; Marazziti et al. 2008). At an anatomical level, these tasks are reasonably coarse, and performance on the Wisconsin card sort test may also be disrupted by posterior cortical lesions (Anderson et al. 1991). Neuropsychological probes that are more selectively associated with the dorsal aspects of the prefrontal cortex, like self-ordered (strategic) working memory tests, are not reliably disrupted in problem gamblers (Goudriaan et al. 2006a; Leiserson & Pihl 2007; Lawrence et al. 2009). Pathophysiology in the dorsal frontal region may only be present in the most severe pathological gamblers, such as Blassczynski & Nower (2002) ‘antisocial impulsive’ gamblers.

Neuropsychological measures of impulsive or risky decision-making have revealed more consistent deficits, resembling the effects seen in patients with damage to the ventromedial prefrontal cortex (vmPFC), who often display real-life difficulties with financial decision-making. This syndrome was initially measured using the Iowa gambling task (IGT; Bechara et al. 1994), where subjects make a series of card choices from four decks (A, B, C, D) that win and lose sums of hypothetical money. Unbeknownst to the subject, decks A and B are ‘risky’, associated with large wins but larger losses that incur gradual debt. Decks C and D are safe decks that yield smaller wins but with negligible losses. While healthy subjects develop a preference for the safe decks over 100 trials, patients with vmPFC damage maintain a preference for the risky decks, accumulating considerable debt. Similar performance has been reported in at least five studies of pathological gamblers to date (Petry 2001b; Cavedini et al. 2002; Goudriaan et al. 2006a; Forbush et al. 2008; Roca et al. 2008).

These findings have been corroborated using other tasks of risky decision-making (Brand et al. 2005; Lawrence et al. 2009) and delay discounting (Petry 2001a), which are also linked to vmPFC integrity (Mobini et al. 2002; Clark et al. 2008). The studies by Cavedini et al. (2002) and Lawrence et al. (2009) report impairment in risky decision-making in problem gamblers in the presence of intact executive ability (on the Wisconsin card sort test and spatial working memory, respectively), supporting the assertion that vmPFC pathophysiology is a more consistent marker in problem gambling. There is a concern that gamblers’ performance on these tasks of risk-taking and decision-making may be distorted by their extensive experience with monetary rewards, judging probabilities, and by their erroneous cognitions related to gambling. This would compromise a strict neuropsychological account of their deficits in terms of underlying brain dysfunction. However, these concerns are mitigated in studies showing comparable neurocognitive effects across problem gamblers and substance addictions (Petry 2001a; Lawrence et al. 2009); notably, patients with alcohol dependence were also impaired on tests of working memory function that were spared in the problem gamblers (Lawrence et al. 2009)). Nonetheless, there is a real need for studies looking at the impact of cognitive distortions upon these simplified neuropsychological tests of gambling behaviour, and to corroborate findings with psychophysiological measures of emotion and motivation, such as skin conductance responses (Goudriaan et al. 2006b).

(c) Functional neuroimaging studies
In recent years, several studies have compared brain responses in groups of problem gamblers and healthy controls during various cognitive tasks. In the first studies of their kind, Potenza and colleagues scanned male pathological gamblers and male healthy controls during performance of the Stroop colour–word interference task (Potenza et al. 2003a) and during presentation of videos of an actor-narrated gambling scenario (Potenza et al. 2003b). This latter ‘cue-induction’ procedure reliably elicits cravings in drug users. In both studies, the gamblers displayed decreased activation in the vmPFC region compared with the controls. In the cue-induction study, the PG group showed additional decreases in the striatum and thalamus. This diminished neural response to cue-induction might be considered surprising, given the elevated subjective reports of craving in these subjects.
A subsequent cue-induction study comparing casino videos against nature videos found increases in brain activity in pathological gamblers, in several regions including the right dorsolateral PFC (Crockford et al. 2005). Differences in the exact cue-induction procedure or patient characteristics may underlie these discrepancies.

Blunted activity in the vmPFC and striatum has been reported in subsequent studies. Reuter et al. (2005) compared brain activity during a card-guessing task in male pathological gamblers and healthy controls. The contrast of monetary wins minus monetary losses revealed a robust response (detectable at the single-subject level) in the ventral striatum and vmPFC. This response was attenuated in the gamblers, and these reductions were significantly correlated with SOGS gambling severity. The authors interpret their finding as consistent with a ‘reward deficiency’ hypothesis that has been applied to drug addiction (Bowirrat & Oscar-Berman 2005): gamblers may be motivated to gamble to stimulate a developmentally underactive brain reward system. This kind of hypothesis assumes that the monetary wins are reinforcing in pathological gambling, and a positron emission tomography study in seven pathological gamblers confirmed increases in striatal glucose metabolism following blackjack play for real money compared with scans performed after a blackjack session for points only (Hollander et al. 2005). Unfortunately, this study did not include a healthy control group for comparison. Reduced vmPFC activity was also reported in a study of substance-dependent problem gamblers as well as in substance-dependent non-gamblers, performing the IGT in the scanner (Tanabe et al. 2007). Pathological gamblers also showed diminished activity in the lateral sector of the ventral PFC, in response to both monetary wins and losses in a reversal learning task, in a recent study by de Ruiter et al. (2009).

Thus, there is some consistency in the observation of blunted ventral frontal cortex and striatal activation, across tasks of reward processing and decision-making (see also Potenza 2008). However, these findings must be treated as preliminary due to the small sample sizes, ranging from seven gamblers in the Hollander et al. (2005) study, to 19 in the de Ruiter et al. (2009) study. Further targets for research in this area also represent issues for the neurochemical and neuropsychological studies. First, the psychobiological approach has predominantly used the case-control design to compare groups of severe pathological gamblers against healthy non-gamblers, but there is a large spectrum of gambling involvement (and gambling harm) that lies between these two groups, and it is necessary to systematically assess the impact of gambling severity on markers of brain function. Second, there has been minimal consideration of sources of variability such as gender, psychiatric comorbidities, or preferred forms of gambling. For example, motivations to gamble may differ between players of different games: casino and sports betting gamblers may be driven predominantly by the excitement of gambling (i.e. positive reinforcement) whereas slot-machine gamblers may play to alleviate negative mood states such as boredom, stress or depression (i.e. negative reinforcement; Cocchi et al. 1995). These differences are likely to moderate the neural correlates of reinforcement processing in problem gamblers.

4. ANOMALOUS RECRUITMENT OF THE BRAIN REWARD SYSTEM DURING COGNITIVE DISTORTIONS

The cognitive and psychobiological accounts are rarely linked in the research literature, partly because of some key differences in approach and methodology. Cognitive studies of gambling frequently use non-gamblers or infrequent players (often university students), and place considerable emphasis on testing in naturalistic settings (e.g. a casino). In contrast, the psychobiological studies derive from a medical model of problem gambling, and have compared pathological gamblers who are typically in treatment, against healthy non-gamblers. In neuropsychological and functional imaging studies, the testing procedures are inherently laboratory based, and some studies have called into question the ecological validity of laboratory gambling, particularly where hypothetical points are involved instead of real money (Anderson & Brown 1984; Meyer et al. 2004). Nevertheless, the two approaches are not mutually exclusive: cognitive distortions must be instantiated at the neural level, and individual differences in brain function or neurochemistry may plausibly influence one’s susceptibility to developing erroneous beliefs about gambling.

In linking the two positions, let us start by considering the role of money. At a psychological level, money is a potent reward. More precisely, money is a conditioned reinforcer, meaning that it is not innately rewarding, but that its value is acquired through extensive pairing with primary rewards and through vicarious, cultural learning. Neurobiological findings indicate the existence of a specialized brain reward system that processes reinforcers and uses reinforcement to guide future decision-making (‘reinforcement learning’). At an anatomical level, FMRI studies demonstrate the central roles of the ventral striatum and the mPFC in this brain reward system; these regions are activated by monetary wins (Delgado et al. 2000; Breiter et al. 2001; Knutson et al. 2003) as well as primary rewards like fruit juice (Berns et al. 2001) or chocolate (Rolls & McCabe 2007).

At a neurochemical level, the mesolimbic dopamine projection from the midbrain to the striatum and PFC is also central to neurobiological accounts of reward processing (Wise 2004). A dominant hypothesis is that dopamine cells code a reward prediction error: the difference between the obtained and the expected reward (Schultz 2002; Montague et al. 2004). Electrophysiological recording from non-human primates has shown phasic bursts of dopamine cell activity in response to unexpected rewards (a positive prediction error). As the monkey learns to associate a conditioned stimulus (CS; e.g. a light) with later reward delivery, dopamine firing shifts to the onset of the CS, and disappears at the time of reward itself; as the reward is now predicted, the prediction error is minimal.
Subsequently, if the CS is presented but the expected reward then withheld, the dopamine cells show a pause in firing at the expected time of reward delivery (i.e. a negative prediction error). These observations have fuelled sophisticated computational models of reinforcement learning and decision-making based on the calculation of prediction errors (e.g. McClure et al. 2003; Daw et al. 2006).

Real-world tasks such as gambling games are more complex than the Pavlovian and instrumental conditioning tasks performed by experimental animals. Recent work has begun to indicate that activity within the brain reward system is modulated by some of the psychological manipulations that affect gambling behaviour. Our own work has focussed on the near-miss effect, using a gambling task based on a two-reel slot machine (see figure 1; Clark et al. 2009). The right-hand reel is spun so that the volunteer can either win £0.50p (if the two reels align) or not win anything; there are no losses in the task. In a study in 15 healthy volunteers with minimal involvement in gambling, the fMRI contrast of wins minus non-wins identified brain responses across established parts of the brain reward system, including the ventral striatum, medial PFC, anterior insula, thalamus and the dopaminergic midbrain (see figure 2a(i),(ii)).

Critically, the non-win outcomes could be further distinguished as ‘near-misses’ (where the reel stopped one position either side of the payline) and ‘full-misses’ (where the reel stopped more than one position away from the payline). Within the network of winsensitive areas, the direct contrast of near-misses and full-misses revealed significant and bilateral activation of the ventral striatum and anterior insula by near-miss outcomes (see figure 2b). Thus, although the objective outcomes were identical on these trial types (i.e. both non-wins), the brain responded to the near-misses in a way that was comparable to the response to a monetary win. This ‘anomalous’ activation may underlie the invigorating effects of near-miss outcomes on gambling play in the studies discussed above by Cote et al. (2003) and Kassinove & Schare (2001).

The slot-machine task was also designed to elicit a second cognitive distortion, of personal control: on half the trials, the subject was required to choose one of six icons on the left-hand reel as a ‘play icon’. The subject won if the right-hand reel stopped on that chosen icon. On the remaining trials, the computer chose the play icon and the subjects made a motor response to confirm selection. Ratings data taken on a trial-by-trial basis revealed greater confidence (‘How do you rate your chances of winning?’) on subject-chosen trials compared with computer-chosen trials, consistent with an illusion of control.

Similar manipulations of personal control have been studied in previous neuroimaging experiments, and show a modulation of brain activity at the level of the dorsal striatum and medial PFC (O’Doherty et al. 2004; Tricomi et al. 2004; Walton et al. 2004; Yeung et al. 2005). Notably, the ventral striatum appears to respond to reward regardless of the level of control (O’Doherty et al. 2004). The experiment by Tricomi et al. (2004) used an oddball task, where in one condition, monetary wins and losses were delivered at a fixed delay after a predictive stimulus. In the second condition, the volunteer was told that a choice response (left or right) would influence whether they won or lost money (in fact, the outcomes were fixed). The dorsal striatum was selectively activated by monetary wins under the choice condition. The study by Yeung et al. (2005) measured event-related potentials during a similar task, and reported greater feedback negativities, which are thought to derive from a medial frontal locus, when outcomes appeared contingent upon the volunteer’s choices, compared with when outcomes required no active choice (see also Walton et al. 2004).

Figure 1. The slot-machine task uses two-reels, with the same six icons displayed on each reel, and a horizontal ‘payline’ across the centre of the screen. On trials with a white screen background, the volunteer selects one ‘play icon’ on the left reel, using two buttons to scroll through the icons, and one button to select. On trials with a black screen background, the computer selects the play icon. Following icon selection, the right-hand reel spins for a variable duration (2.8–6 s), and decelerates to a stall. During outcome (4 s), if the right reel stopped on the selected icon (i.e. matching icons displayed in the payline), the subject was awarded £0.50; all other outcomes won nothing. Following the outcome phase, there was an inter-trial interval of variable duration (2–7 s). In the fMRI version of the task, two ratings were taken on intermittent (1/3) trials: following selection, subjects were asked ‘How do you rate your chances of winning?’, and following outcome, subjects were asked ‘How much do you want to continue to play the game?’. Reprinted from Clark et al. (2009).
The dorsal striatum is also known to be involved in the formation of habits, and this role has generated considerable interest in the context of drug addiction. For example, studies in experimental animals have given rise to the hypothesis that the neural regulation of drug taking progresses from the ventral striatum to the dorsal striatum as the initial recreational consumption of drugs (e.g. for their hedonic effects) develops into habitual and compulsive usage (Everitt & Robbins 2005). As supporting evidence for this role of the dorsal striatum in drug addiction, rodent studies have shown that inactivation of the dorsolateral striatum (by infusion of gamma-aminobutyric acid agonists) prevented context-induced reinstatement of cocaine seeking in withdrawn animals (Fuchs et al. 2006). Changes in dopamine function in the dorsal striatum are observed after chronic, but not acute, cocaine self-administration (Porrino et al. 2004). Whether this progression would also occur in a form of ‘behavioural addiction’ like problem gambling, where there is no drug involved, is unknown. As such, processes of habit formation and dorsal striatal function in problem gamblers represent an important target for future research that may answer broader questions of relevance to drug addiction.

The neuroimaging findings reviewed above suggest that gambling games harness a brain reward system that has evolved to learn about skill-oriented behaviours: situations where response feedback can be used either to improve the precision of the motor response itself, or to improve the prediction of future outcomes. This system often responds inappropriately under conditions of chance. Using the example of the near-miss, in many real-world situations such as target practice or getting to the railway station two minutes late, it is advantageous for the brain to assign value to near-miss outcomes, as they are a valid and useful signal of future success. However, in gambling games, where winning outcomes are largely or purely determined by chance, near-misses provide no information on future success, and it is misleading for the brain to assign them value. Similarly, in the case of personal control, it is obviously adaptive for the brain to learn how to control its environment, and specialized and sophisticated processes have evolved to identify rewards that occur contingently upon behaviour. However, the random nature of gambling games means that the availability of personal control has no actual bearing on the likelihood of a win occurring.

These data showing modulation of striatal and medial PFC activity by near-misses and personal control are from studies in healthy volunteers, who had low levels of gambling involvement. The findings therefore suggest that the brain reward system is naturally susceptible to these cognitive distortions associated with gambling. Nonetheless, the neuropsychological and functional imaging data described in the previous sections indicate substantial changes in the functionality of this system in problem gamblers, along with alterations in dopamine transmission. By the reasoning I have outlined above, the observed reductions in ventral striatum and vmPFC activity (Potenza et al. 2003a; Reuter et al. 2005) may be only part of the story. Under conditions of cognitive distortion, it is hypothesized that these regions would be excessively recruited in pathological gamblers. We are testing this prediction in ongoing work.

In conclusion, the data outlined above suggest that two of the better-established cognitive distortions in gambling behaviour, the near-miss effect and the effect of personal control, are associated with anomalous
recruitment in components of the brain reward system. The term ‘anomalous’ is justified by the objective status of near-misses as loss events that do not signal future success, and the objective irrelevance of personal control to gambling success on games of chance. This mechanism is unlikely to represent the only interface between the cognitive and psychobiological approaches to gambling, and recent neuroimaging work has highlighted several other possible avenues. For example, there are emerging links between chasing behaviour, which is often viewed as the final common pathway in problem gambling, and impaired recruitment of cortical brain regions involved in conflict monitoring and inhibitory control (Campbell-Meiklejohn et al. 2008; de Ruiter et al. 2008). The perception of patterns (or ‘streaks’) within random sequences, fuelling a Gambler’s Fallacy, has received little attention in the neuroimaging field, but is also likely to involve interactions between the frontal lobes and the striatum (Elliott et al. 2000). There is a need to develop better tasks to capture these cognitive distortions in the scanner, and it is encouraging that studies in irregular and non-gambler samples seem able to detect variability in these distortions at a neural level (Campbell-Meiklejohn et al. 2008; Clark et al. 2009). The longer-term objective here is to understand how this neural circuitry changes in the transition from recreational gambling to problem gambling. In order to achieve this target, there is also an urgent need for longitudinal designs that follow gamblers as they move in and out of problematic levels of gambling involvement.

Supported by a project grant from the Economic and Social Research Council and Responsibility in Gambling Trust (RES-164-25-0010, L.C. and T.W. Robbins) and completed within the University of Cambridge Behavioural & Clinical Neuroscience Institute, supported by a consortium award from the Medical Research Council (UK) and the Wellcome Trust. I am grateful to Dr R. Cools for feedback on an earlier version of the manuscript.

REFERENCES


