Examining Belief and Confidence in Schizophrenia

Dan W. Joyce, B.Sc., MBBS, Ph.D.[1], **Bruno B. Averbeck, B.S., Ph.D.**[2], **Chris D. Frith, PhD FRS**[3], and **Sukhwinder S. Shergill, B.Sc., MBBS, Ph.D., FRCPsych**[1]

[1]Kings College London, Institute of Psychiatry, Cognition Schizophrenia and Imaging Laboratory, PO 96, Department of Psychosis Studies, De Crespigny Park, London SE5 8AF

[2]Laboratory of Neuropsychology, NIMH/NIH, Building 49 Room 1B80, 49 Convent Drive MSC 4415, Bethesda, MD 20892-4415

[3]Wellcome Centre for Imaging Neuroscience, University College London, 12 Queen's Square, London WC1N 3BG

Abstract

Background—People with psychoses often report fixed, delusional beliefs that are sustained even in the presence of unequivocal contrary evidence. Such delusional beliefs are the result of integrating new and old evidence inappropriately in forming a cognitive model. We propose and test a cognitive model of belief formation using experimental data from an interactive "Rock Paper Scissors" game.

Methods—Participants (33 controls and 27 people with schizophrenia) played a competitive, time-pressured interactive two-player game (Rock, Paper, Scissors). Participant's behavior was modeled by a generative computational model using leaky-integrator and temporal difference methods. This model describes how new and old evidence is integrated to form both a playing strategy to beat the opponent and provide a mechanism for reporting confidence in one's playing strategy to win against the opponent

Results—People with schizophrenia fail to appropriately model their opponent's play despite consistent (rather than random) patterns that can be exploited in the simulated opponent's play. This is manifest as a failure to weigh existing evidence appropriately against new evidence. Further, participants with schizophrenia show a 'jumping to conclusions' bias, reporting successful discovery of a winning strategy with insufficient evidence.

Conclusions—The model presented suggests two tentative mechanisms in delusional belief formation – i) one for modeling patterns in other's behavior, where people with schizophrenia fail to use old evidence appropriately and ii) a meta-cognitive mechanism for 'confidence' in such beliefs where people with schizophrenia overweight recent reward history in deciding on the value of beliefs about the opponent.

Introduction

The cardinal features of psychotic illness are the presence of hallucinations (perceptual experiences in the absence of an external stimulus) and delusions (fixed false beliefs held contrary to evidence and against the prevailing sociocultural milieu). After any novel subjective perceptual experience, there is higher order processing that amalgamates experience with extant beliefs, or leads to the initiation of a new belief. These beliefs are tested in the environment, leading to maintenance or extinction depending on their utility (e.g. a belief will be extinguished if it proves to be incorrect and has low utility). Therefore, to maintain *appropriate* (i.e. non-delusionary) beliefs, experienced events must be *temporally* integrated with internal models, tested against the environment and then discounted or retained according to feedback from the environment. When the evidence supporting a belief is inconsistent or contrary, the belief is highly resistant to revision and is accompanied by a subjective feeling of conviction, it becomes delusional. This represents a dysregulation of *metacognitive* processing, which refers to the evaluation of one's own internal cognitive processes, assigning confidence and utilizing this in modifying behavior (Metcalfe *et al*, 1994; Koriat, 2007) including inferences about the behavior and intentions of others (i.e. theory of mind). Subjective feelings of confidence in one's beliefs can be described in terms of theory-based (e.g. deliberative thought and reasoning) or experience-based (e.g. intuitive and unconscious) with the latter being the dominant model of human metacognition (Bruno *et al*, 2012) particularly in procedural learning. Studies examining metacognition in schizophrenia have focused on metamemory processes – for example, in (Bacon *et al*, 2009) participants were shown a string of consonants, and asked to *prospectively* rate the probability of accurately spotting this string among seven distractors after a short interval ("feeling of knowing" rating). Patients with schizophrenia persistenly performed worse on actual recall and provided lower "feeling of knowing" ratings. In contrast, *retrospective* judgements tend to show patients with schizophrenia rate higher than controls (Danion *et al*, 2001; Moritz *et al*, 2003; Mortiz *et al*, 2006). Recent evidence suggests that deficits in metacognitive ability, including theory of mind, are stable features of schizophrenia (Lysaker *et al*, 2011) showing little change over time. These deficits in schizophrenia may be underpinned by an inability to integrate new information into existing belief systems (Cohen *et al*, 1996) and the established aberrant sensitivity to reward to correct negative beliefs and guide decision making (Fletcher and Frith, 2009, Waltz and Gold, 2007).

Existing cognitive accounts of psychotic symptoms suggest that illness arises through a mixture of dysfunctional predictive models (Frith and Done, 1989), jumping to conclusions (Garety and Freeman, 1999) allied with altered reward processing and dopaminergic dysfunction (Kapur, 2003; O'Daly et al 2011). As a common theme in all of these cognitive models, the suggestion is that both perceptual and inferential biases contribute to the establishment of psychotic symptoms which are then held with both confidence and rigidity, so that these beliefs are difficult to "overcome" even in the presence of strong counter-evidence (Blackwood *et al*., 2001, Fletcher and Frith, 2009). This can be conceptualized in a parsimonious fashion within a Bayesian framework (Fletcher and Frith, 2009), where developing beliefs about the world is viewed as a probabilistic inference task where old evidence (*prior* belief) is updated according to new experiences or evidence (*likelihoods*)

allowing the derivation of a new "model" of the world (*posterior* beliefs). For example, when one's prediction about the world (prior belief) fails to explain or predict some new observation (evidence), there is a "surprise" generated by this dissonance (i.e. the event is assigned salience) that triggers updating of one's existing beliefs. This important process for learning the causes of sensory experience is expressed as the dopaminergic-dependent prediction error signal (Hollerman and Schultz, 1998) within the striatum; reflecting a mismatch between the expectation and incoming sensory input. An additional factor which is often neglected is the questions of our *confidence* in our beliefs (meta-cognition) about the world. There is consistent but counterintuitive data showing the considerable discrepancy between the actual probability of certain events occurring, and people's confidence in the occurrence of the same event.

In brief, the missing aspect of contemporary models of psychotic beliefs is the process of belief maintenance and failure of abolition. Is there an optimal paradigm to examine this in an ecologically valid manner? Previous studies examining the role of confirmatory and contradictory evidence in reasoning and metacognition have tended to focus on high level "deliberative" reasoning tasks (Monestes *et al.*, 2008, Sellen *et al.*, 2005, Woodward *et al.*, 2006, Woodward *et al.*, 2008). However, evidence for a consistent deficit in people with schizophrenia using such reasoning tasks is variable (Fletcher and Frith, 2009). Those that attempt a more probabilistic explanation (e.g. the "jump to conclusion" phenomena) have focused on the beads-counting task, where draws of beads from one of two jars each containing a majority of yellow or black beads are undertaken sequentially. Participants decide when they have enough evidence to decide the jar being drawn from is the "majority yellow" or "majority black" jar (Garety and Freeman, 1999, Freeman *et al.*, 2006, Freeman *et al.*, 2008, Garety *et al.*, 1991, Huq *et al.*, 1988, Startup *et al.*, 2008), Such tasks have limited ecological validity, by virtue of the non-interactive, non-goal directed nature of the experiments (where, for example, the utility of beliefs is not crucial to the execution of the task). In order to answer this question, we propose an alternative experimental approach based on *active interactions* akin to those we routinely encounter in everyday life. This approach is exemplified in the behavioral economics and game theory literature (Camerer, 1999, Camerer, 2003, Fehr and Camerer, 2007, Fett et al 2012; King-Casas *et al.*, 2005, Rangel *et al.*, 2008) especially in iterated competitive games. Since we are concerned with the participant's ability to detect, model and use regularities in the opponent's plays we implement a simulated opponent which gives the illusion of playing like a "real" opponent, but in fact, presents a statistically defined frequency of plays. We use a modified "rock paper scissors" (RPS) game, where new evidence obtained after each trial must be selectively integrated with existing evidence in order to update belief and form the basis for future actions. We suggest this represents a "middle ground" between probabilistic inference tasks (e.g. bead counting) and the high-level reasoning tasks.

We hypothesize that patients with schizophrenia will differ from healthy controls in i) Failure to appropriately integrate new evidence with existing beliefs and ii) when evaluating their performance, patients will have excessive confidence in their beliefs.

Methods

Participants

Twenty seven participants with schizophrenia and thirty three control participants were recruited from the South London and Maudsley NHS Foundation Trust. Patients had a diagnosis of schizophrenia based on the DSMIV (APA 1994) and were selected on the basis of having current positive symptoms of hallucinations and delusions (greater than 3 on respective PANNS items) or commensurate levels of positive symptoms documented in their clinical records during exacerbations of their illness over the last 5 years. The average chlorpromazine equivalent was 219.8 (+/− 178.6). Control and patient groups where matched on years of formal education. Full demographics data are presented in Table 1. All subjects were required to give informed consent. This study was approved by the South London and Maudsley and the Institute of Psychiatry Research Ethics Committee.

Experimental Design

Participants were told they would play six RPS games against an opponent via a computer interface. Unknown to the participants, the opponent was a computer program (Gallagher *et al.*, 2002, Paulus *et al*, 2005, Paulus *et al.*, 2004) but with one important distinction – in our design, the distribution of computer opponent's moves are governed by a parameterized Multinomial distribution with three different parameter sets that define an easy, medium or hard game.

In any given game, the computer played randomly (i.e. with no pattern of favored plays) for 20 trials then began playing favoring one play for the 40 subsequent trials. In an easy game, the computer switched to an obvious distribution, favoring the same move (stochastically) on 80% of the trials with the other two moves being played on the remaining 10% of trials respectively. In a medium game, the computer behaves similarly, but favors one move on 60% of the trials, and the other two moves are played on 20% of trials. Finally, a "hard" game is one where the favored move is played on 40% of the trials, and the other two moves are played on 30% of the remaining trials respectively. The multinomial distribution generating the opponent's play assumes independent trials, so on each trial, the opponent's play is not dependent on preceding trials or participant's previous plays. Participants played two easy, medium and hard games, resulting in a total of six games.

During games, we also probed for participant's confidence in their beliefs that they had found and were able to exploit a "winning streak". Participants were told that on each trial, the winner gained one point and the loser incurred the loss of one point (in accordance with a symmetrical zero-sum game). A draw (i.e. both play rock) results in no points for either. If, at *any* trial they felt sure they were on a 'winning streak', they could instruct the experimenter to press a button that would double both their wins and losses. We refer to this as "increasing the payoff" – see section on metacognition and confidence below. Participants were told this was an irreversible, one-off decision to encourages a conservative approach to doubling their wins and losses. Participants were given *no* explicit feedback of their current total score or their performance from previous games, forcing them to rely entirely on their own estimates of performance.

On each trial, a count down from three to one preceded a "go" signal. Participants then played their move within one second (using the keyboard) and then simultaneously, the computer reveals its move (a photograph of rock, paper or scissors) and the outcome for the participant; whether they won, drew or lost. If participants did not play within a second of the countdown ending, they were instructed they were too slow and the trial restarted. Each trial had a total duration of 4250 milliseconds.

Prior to commencing the experiment, participants were first trained to use the keyboard to indicate their play on each trial. Each possible play was presented randomly until participants response time (e.g. pressing the correct button for rock, paper or scissors) decreased below a threshold. Once participants demonstrated a clear understanding they began playing the six experimental games. In addition, to monitor participants engagement with the task (to prevent inattention / distraction), the experimenter sat next to and observed the participants. The experimenter was also responsible for pressing a button to increase payoffs on verbal instruction from the participants.

Strategy; action selection

In the RPS games, participants are expected to build a model of their opponent that informs their play. To do this, they must balance the evidence available to them from the history of previous plays as well as weighting new evidence. In "hard" games, evidence from previous plays is practically redundant, as the opponent plays almost randomly throughout the game. In "easy" games, previous evidence is a reliable predictor of the opponent's strategy, as the opponent will play the same move on a high proportion of the trials.

We sought to test whether patients have difficulty correctly balancing or weighting existing evidence against new evidence. The implication of this is that they fail to detect and model meaningful regularities in the frequency of the opponent's play, resulting in a poor strategy for winning against the opponent. This was modeled by a combined leaky-integrator and temporal difference model (Sutton, 1998). Leaky integration is a key feature of neuronal mechanisms for coincidence detection. An incoming excitatory signal stimulates a neuronal population driving its activity upwards. If another excitatory signal arrives within a short time window, the integration process enables the two events to be accumulated (reflected by a sustaining or increasing in the activity of the population) but if the second or subsequent excitatory signals are too far apart, the "leaky" component effectively dissociates them (i.e. the population's activity falls). Similar techniques, such as the diffusion-drift model and Ornstein-Uhlenbeck process, have been used to understand reaction time and accuracy tradeoffs in two-alternative forced choice experiments; see (Bogacz *et al*, 2006) for review. In our study, the participant's playing strategy is modeled by parallel leaky integrators (one for each of rock, paper and scissors), each competing in a competitive winner-takes-all arrangement (a softmax function of the activity of three leaky-integrator components) enabling the derivation of probabilities and the selection of one of the three plays to be executed on subsequent trials. The model is updated on a trial-by-trial basis, using two different pieces of information (cf. incoming signals) i) new evidence about the current strategy with respect to the opponent's play – this signal represents the temporal difference between the *predicted* outcome of playing an action (i.e. the current state of the strategy

suggests playing rock will result in a win on the next trial) and *actual* outcome for a given action (i.e. rock was played, but the reward was a loss) ii) a decaying (i.e. leaking) of prior expectations based on previous evidence. Rock, paper and scissors all have an accumulated history of their utility against the opponent, but this knowledge will decay over time unless it is reinforced by continued new evidence in its favour. These two factors are modeled explicitly by two parameters – α models the leaky decaying of expectations about utility, and β is the weight given to new evidence. Figure 1 shows the theoretical parameter space and the corresponding playing strategies for the model. As α tends to zero, the model emphasizes the value of expected utility based on previous evidence (that is, it discards less prior evidence so one win with a particular move will be carried forward for some time). Conversely, if α tends to unity, then a subject would be ignoring all prior expectations from the history of plays (i.e. their play would be retrograde amnesic). One extreme model is represented by $\alpha = 1$, $\beta = 0$ where the participant ignores all prior expectations, and gives no weight to new evidence which would result in random play. Further, if $\alpha = 0$, $\beta = 1$, then the model reduces to frequency counting, and would approximate a Bayesian model with multinomial likelihood and Dirichlet conjugate prior.

Importantly, this model *does not* assume that players explicitly use the payoff matrix rationally as would, for example, a model based on fictitious play (Fudenberg, 1995, Fudenberg, 1998) or that participants have a statistical model of the *a posteriori* distribution of the strategy as a formal Bayesian model might (Bernardo and Smith, 2000).

Meta-cognition; confidence in decision making

In Figure, we propose that a point-estimate of confidence is derived from the action selection / strategy update model (the leaky integrator system described). Figure 2A shows how the decision to increase the payoff may be a function of the output of a running history of prediction errors (i.e. the model's expected reward on a trial given a strategy). In this case confidence becomes high when the prediction error becomes sufficiently small. This would imply an internal subjective estimate of the expected value of actions is being used. Alternatively, Figure 2B shows how the "absolute" reward /payoff may be used (i.e. similar to A, but where the absolute reward (−1, 0 or 1) is used instead of the model-derived expected reward). This would suggest a more objective evaluation of rewards (outcomes) received rather than an internal, subjective evaluation based on expectation. This is similar to Actor-Critic models (Sutton, 1998) – while one system updates the strategy upon which action selection takes place, another evaluates the success of the strategy.

Analogous to the action-selection model described earlier, the behavior of participants is modeled using another leaky integrator model with parameters η and κ being the weights associated with decaying the previous payoff history and accumulating new payoffs respectively (see Supplementary Information for details of implementation). These parameters are analogous to α and β, where they can vary between extremes of throwing away all new information and using only new evidence, and the converse of using only old information and ignoring new payoffs.

Results

All data analysis was performed using MATLAB 7.3 (MathWorks Inc., Sherbon, MA).

Overall Performance on Games Between Groups

Patients tended to perform worse than controls in terms of total cumulative score (wins minus losses) at the end of a game (Figure 3), across all three levels of difficulty (analysis of variance; game difficulty, total score, group; df = 1, F = 12.22, $p < 0.0005$).

This difference in performance is not due to poor engagement with the task, as if this were the case, patients would perform at the same poor standard irrespective of the game difficulty (see Supplementary Information for more detail). All participants learn the pattern quickly in easy games, less so in medium games and as expected, very little learning occurs in the hard games.

Model fitting

Each fitted model was run with the estimated parameters to generate a *predicted* sequence of play for each game, compared to what *should* have been played to win on each trial and averaged. The model is accurate in predicting the behavior of participants and the model fitted both controls and patients equally well (mean log likelihood for controls across all games = 0.898; SD = 0.171; mean log likelihood for patients across all games = 0.936; SD = 0.180). In terms of the fitted model predicting the participant's actual actions on every trial, across every game, the fitted models performed best on the easy games (as did the participants) with mean correct trial-by-trial model prediction of 0.707 (SD = 0.149) and 0.600 (SD = 0.203) in controls and patients respectively (where a score of 1.0 would indicate each model correctly predicted every trial of every game). On medium games, the model for controls yielded a mean correct prediction of 0.583 (SD = 0.144) and for patients 0.603 (SD = 0.156). On hard games (i.e. close to random) control models performed at 0.495 (SD = 0.140) and patient models 0.521 (SD = 0.142). Further analysis of the performance of the model is given in the Supplementary Information.

Between-Group Differences in Strategy

To assess how participants integrate new evidence with existing beliefs, the parameters α and β were averaged within group (controls, patients) as a function of game difficulty. An analysis of variance showed that for α, there was an effect of group (patient, control; df = 1, F = 10.35, $p < 0.002$) and also game difficulty (easy, medium, hard; df = 2, F = 12.96, $p < 0.0001$) but no interactions of group by difficulty (df = 2, F = 2.55, $p = 0.08$) (Figure 4). However, there were no significant differences for β. This result suggests that controls and patients use *new* evidence to a similar degree, but *differ* in how they use existing evidence to influence their strategy.

Patients place *less* emphasis on expectations based on previous evidence (i.e. α tends to one, and they discard previous accumulated evidence more quickly) than controls in the easy (one tailed t-test; t = −3.44; $p < 0.0004$; patients mean $\alpha = 0.44$; controls mean $\alpha = 0.23$) and medium games (one tailed t-test; t = −2.05; $p < 0.022$; patients mean $\alpha = 0.55$; controls

mean $a = 0.43$) but not on hard games (where outcomes are most unpredictable). This suggests that when meaningful regularities in the frequency of plays *are evident* in the opponent's play, the higher value of a causes the estimated utility of each play to fall more quickly in patients than in controls. Thus, patients are unable to temporally "link together" events that represent reliable predictors of an opponent's play. Further analysis is presented in the Supplementary Information.

Correlation of Parameters with PANSS Items

For easy games (where the pattern of play was obvious and exploitable), there was a modest correlation between higher values of a and greater scores on the Delusions item of the PANSS questionnaire for patients : Spearman's $rho = 0.273$, $p = 0.045$.

Between-Group Differences in Meta-cognition

To qualitatively explore the decision to gamble on doubling payoffs, games where no decisions were made were discarded, leaving 116 and 100 games for controls and patients respectively. Patients exhibited increased confidence (i.e. by doubling payoffs) earlier in the games than controls (Kolmogorov-Smirnoff test; $p < 0.0001$; controls mean decision at trial 31.46, SD=13.62; patients mean decision at trial 23.77, SD = 16.03).

By averaging the history of payoffs (wins, losses and draws) over the trials directly *preceding* the decision to double the payoff, it was found that participants generally experience an "upswing" of around ten trials (Figure 5) of positive reward before the trial on which they make the decision to double the payoffs (see Supplementary Information for detailed analysis).

In examining whether internally-derived measures (i.e. prediction errors) or actual absolute rewards influence metacognitive assessment, control subjects showed no correlation between the trial at which the decision to double-bets was made and measures of the average prediction error or average absolute reward in the ten trials preceding the decision (r^2 =0.068, p>0.05 and $r^2 = 0.131$, p>0.05). In patients, only the ten-trial average of the mean absolute reward showed a correlation with the trial at which the decision was made (r^2 =0.447, p<0.0001), but not the ten-trial mean prediction error. This suggests that a simple correlation-based explanation, where an increasing trend in reward rather than punishment predicts confidence judgments will not suffice.

A leaky-integrator model was used, and the parameters η and κ found by fitting the model to the data by minimizing a quadratic objective function of the time the participant made their decision and that predicted by the model for a given estimate of η and κ (see supplementary information). Using the absolute (rather than prediction error) payoff accurately predicted the decision to double payoffs with a mean error of ± 1.4 trials in controls, and ± 1.6 trials in patients. The model using the derived prediction error was much less accurate (± 9.0 trials in controls, and ± 15.1 trials in patients).

Interestingly, patients and controls did *not* differ in the amount by which they decay previous evidence (η) for the decision to double payoffs but patients gave significantly more weight

to new rewards (κ) (t-test; one tailed, patients > controls; $p<0.0008$; patients mean $\kappa = 0.51$, controls mean $\kappa = 0.36$).

Discussion

We have shown two dissociable mechanisms at work during decision making in schizophrenia; one strategically evaluating evidence for and deciding on a specific action, and another metacognitive mechanism that acts to assign confidence in the selected strategy. The results demonstrate that patients weight new evidence similarly to controls in deciding on their strategy, but they 'leak' prior evidence (the alpha parameter) preventing efficient action selection, as the temporal patterns from previous plays are not incorporated in decision making and they are less able to detect and exploit meaningful regularities in the opponent's play (Gray *et al.*, 1991). This was particularly pronounced during the easy games, where the pattern of the opponent's play was more obvious.

Despite their less efficient opponent modeling and action selection strategy, patients still exhibit over-confidence when assessing their confidence in their strategy – choosing to increase the stakes in the game (they "increase payoff" earlier) – in the face of less objective evidence and this is driven by an overweighting of *new* evidence (the kappa parameter) in the temporal sequence of absolute rewards.

In these patients with schizophrenia, these factors could explain why psychotic beliefs are maintained and not extinguished in the face of contrary evidence. In everyday interactions with the world, there are a succession of incoming signals, from the environment (i.e. rewards, feedback on performance or observations about other agents) – some of which are noise (i.e. random fluctuations) while others represent meaningful regularities or associations between events. These incoming signals must be evaluated for their utility (expected value) and temporally integrated when appropriate (for example, when these utilities are congruent with the consequences of actions in the environment) and discarded (leaked) when they represent meaningless coincidences. Furthermore, the meta-cognitive task of evaluating confidence in one's beliefs about strategy is skewed in favour of more *recent* events and their confidence model fails to filter out any sporadic random runs of success.

Theories of Belief

Our model frames these cognitive processes (modeling the environment by detecting meaningful regularities in events, action selection and confidence) in a simple, parsimonious model which is computationally plausible, and driven by the mechanisms by which neuronal populations integrate incoming signals in the temporal and spatial domains. The RPS game in this study naturally lends itself to a theory of belief representation in terms of mapping observable stimuli to actions via an internal representational scheme based on probabilistic representations. This contrasts to formal epistemological theories of belief representation (Hintikka *et al*, 2005) where the beliefs are defined in terms of theorem-based manipulations over symbolic propositions or predicates in the "language of thought" (Fodor, 1998). For example using Bratman's (1987) theory of practical reasoning, playing the RPS game would be formulated as inferences over sets such as $(C,O) \rightarrow A$, where O enumerates the most

recent play by the opponent, C is a finite set of contexts (e.g. enumerations over the set of opponent players such as an "easy" or "difficult" opponent) and A enumerates the plays available to the participant. This latter approach to modeling belief, particularly as applied to the dynamic processes underpinning belief and delusions, has yet to be evaluated. It can be argued that formal epistemological theories capture the explicitly "linguistic" and propositional nature of belief (e.g. "The opponent is a cheat") whereas our probabilistic approach represents an implicit, action-oriented and bounded rational interpretation of "belief". We propose that the dynamic, adaptive cognitive processes underpinning belief formation are best studied using such implicit action-directed approaches.

Our model posits that this failure of modulation by context is a function of the 'leaky' component of our model. If discrete units of evidence are allowed to 'leak-out' too quickly, then information about meaningful temporal sequences will never be correctly associated together and no reliable evidence will be available for the higher levels of the hierarchy (i.e. those responsible for maintaining or abolishing beliefs, and meta-cognitive systems that evaluate performance).

Theories of Metacognition

Theories of metacognition span two axes: "Monitoring/Control" or "Control/Monitoring" models (Koriat and Ackerman, 2009) and the "information/theory based" or "experience-based" (Koriat, 1997). A third position in metacognition is represented by Theory of Mind (Koriat and Ackerman, 2009) where stored representations of mental state (combined with rules of inference relating these stored representations to observable behavior) allow an individual to predict others intentions. Our model and experimental results suggest two parallel processes; the judgment of how well one is performing (evidenced by the metacognitive act of "doubling" bets when confidence reaches a threshold) appears to be better predicted by temporal changes in absolute reward rather than depending on measures derived from the "control" (action selection) process.

When confidence is directly *derived from* the internal action-selecting mechanism (suggesting one integrated system) we are unable to predict control's and patient's performance in increasing the stakes of the game by doubling payoff. But, for both groups, this performance *can* be predicted from *absolute* reward/payoff signals. The confidence process can be viewed within the standard actor-critic model, where one implicit mechanism is "fast and dirty" for driving trial-by-trial behavior and another "critic" evaluates the performance of this system cf. the proposed sequential Monitor/Control or Control/Monitor (Koriat and Ackerman, 2009). Patients consistently made decisions earlier than controls, with more limited information, and indeed, in our model, this was reflected by a higher weighting for *new* (more recent) rewards rather than previous history. This is analogous to the over-attribution of evidence observed in the 'jumping to conclusion' bias in the beads task and similar experiments (Freeman *et al.*, 2008, Speechley *et al.*, Woodward *et al.*, 2009), and is consistent with the observation that people with schizophrenia have poor self-assessment of their own performance and functional status (Bowie *et al.*, 2007). In our task, we attempted to study the dynamic process *underlying* the decision rather than manipulating experimental conditions that define probabilities of the likelihood (new evidence) and prior

(old or accumulated evidence). This probe of meta-cognitive ability emphasizes "output bound" performance (Koren *et al.*, 2006, Koren *et al.*, 2004) where self-monitoring, evaluation and commitment to one's own behavior is based on a model of the world (i.e. beliefs about interactions with other agents) and directed toward behavior. This is in contrast to what Koren and colleagues (2006) describe as "input-bound" measures (such as in beads counting tasks) where the problem is framed such that participants make assessments of input probabilities, forcing participants to commit to a response.

Relationship to Monoamine Theories of Schizophrenia

The metacognitive results of our study and model are explained by the pre-synaptic hyper-dopaminergic state found in schizophrenia (Howes *et al*, 2012) – with increased response to positive feedback (Pessiglione *et al.*, 2006) that may drive abnormal salience responses in schizophrenia (Kapur, 2003, Kapur *et al.*, 2005). However, striatal hyper-dopaminergia alone cannot explain the findings in the patient's action-selection strategy. Our model predicts poorer temporal integration in people with schizophrenia. This can be explained by "context-processing" deficits in schizophrenia (Barch and Ceaser, 2012). Here, "context" refers to appropriate online maintenance of representation of probable opponent play to enable action selection of counter-plays. It also requires filtering of irrelevant stimuli (i.e. moves by the opponent which do not concord with the emerging dominance of their preferred move). Prefrontal D1 and D2 neurons have been proposed to operate in dual-state networks - when these networks are driven by D1 activity, e.g. by experimental D2 blockade (Mehta *et al*, 2004), stable working memory formation dominates, with irrelevant stimuli being filtered but with poor context-switching and response flexibility (Cools and D'Esposito, 2011). In contrast, D2 dominated activity favours flexible response selection while sacrificing filtering of temporally intervening irrelevant stimuli (Durstewitz and Seamans 2008, Cools and D'Esposito, 2011).

Friston *et al* (2002) have suggested that higher levels of a cortical hierarchy provide contextual guidance to lower levels of processing based on a prediction of inputs and that these are modified in the presence of a mismatch – this does not occur in patients with schizophrenia. This failure to modify prior belief in the presence of new evidence is supported by an extensive literature demonstrating this for example, in visual processing of hollow mask illusions (Schneider *et al.*, 2002), and their neural correlates (Dima *et al.*, 2009). Other modalities such as event related potentials in processing discrepant information (Debruille *et al.*, 2007), impaired stimulus evaluation (Doege *et al.*, 2009) and our own work on predictive models distinguishing self and other (Shergill *et al.*, 2005; Simons et al 2010) suggest that top-down modulation in the integration of evidence is dysfunctional in schizophrenia.

In conclusion, patients with schizophrenia demonstrate metacognitive changes which lead them to a jumping to conclusions bias, making decisions on the basis of insufficient evidence, driven by selective over-weighting of recent rewarding events, rather than a carefully balanced assessment of recent successes and failures over time. This data supports a model of psychotic symptoms, concordant with a hyperdopaminergic state, which gives rise to the salience of aberrant perceptions linked to abnormal beliefs. These may occur

transiently in a significant proportion of the general population (Smeets *et al.* 2012). However, the reason these beliefs are not extinguished in the face of contrary information is because there is a failure of the normal mechanism for integrating evidence in the presence of meaningful temporally ordered events; this deficit is compounded by the changes in metacognitive processing, giving an inappropriately higher weighting to absolute rewards from the environment. Further work is required to disassemble the relative contribution of neural networks responsible for aberrant "leaking" of evidence. This data tentatively supports current therapeutic approaches that encourage efficient decision making through cognitive remediation where patients are encouraged to make explicit judgments during stepwise reasoning. Given that improving both positive symptoms and cognitive dysfunction independently have good predictive value for long term outcomes (Bowie *et al.*, 2006), further translation of experimental cognitive approaches focusing on serial decision making is warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Bacon E, Izaute M. Metacognition in schizophrenia: Processes underlying patients' reflections on their own episodic memory. Biological Psychiatry. 2009; 66:1031–1037. [PubMed: 19726032]

Bernardo, JM., Smith, A. Bayesian Theory. John Wiley and Sons; 2000.

Blackwood NJ, Howard RJ, Bentall RP, Murray RM. Cognitive neuropsychiatric models of persecutory delusions. American Journal of Psychiatry. 2001; 158(4):527–539. [PubMed: 11282685]

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. Psychological Review. 2006; 113(4):700–765. [PubMed: 17014301]

Bowie CR, Twamley EW, Anderson H, Halpern B, Patterson TL, Harvey PD. Self-assessment of functional status in schizophrenia. Journal of Psychiatric Research. 2007; 41(12):1012–1018. [PubMed: 17014866]

Bowie CR, Reichenberg A, Patterson TL, Heaton RK, Harvey PD. Determinants of real-world functional performance in schizophrenia subjects: correlations with cognition, functional capacity, and symptoms. American Journal of Psychiatry. 2006; 163(3):418–425. [PubMed: 16513862]

Bratman, ME. Intention, Plans and Practical Reason. CSLI Publications, University of Chicago Press; 1987.

Bruno N, Sachs N, Demily C, Frank N, Pacherie E. Delusion and Metacognition in patients with Schizophrenia. Cognitive Neuropsychiatry. 2012; 17(1):1–18. [PubMed: 22216943]

Camerer C. Behavioral economics: reunifying psychology and economics. Proceedings of the National Academy Sciences USA. 1999; 96(19):10575–10577.

Camerer C. Behavioural studies of strategic thinking in games. Trends in Cognitive Science. 2003; 7(5):225–231.

Cohen JD, Braver TS, O'Reilly RC. A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. Proceedings of the Royal Society of London, B. 1996; 351:1515–27.

Cools R, D'Esposito M. Inverted-U–Shaped Dopamine Actions on Human Working Memory and Cognitive Control. Biological Psychiatry. 2011; 69:113–125. [PubMed: 20678756]

Danion JM, Gokalsing E, Robert P, Massin-Krauss M, Bacon E. Defective relationship between subjective experience and behavior in schizophrenia. American Journal of Psychiatry. 2001; 158:2064–2066. [PubMed: 11729027]

Debruille JB, Kumar N, Saheb D, Chintoh A, Gharghi D, Lionnet C. Delusions and processing of discrepant information: an event-related brain potential study. Schizophrenia Research. 2007; 89(1–3):261–277. [PubMed: 16945505]

Dima D, Roiser JP, Dietrich DE, Bonnemann C, Lanfermann H, Emrich HM. Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modelling. NeuroImage. 2009; 46(4):1180–1186. [PubMed: 19327402]

Doege K, Bates AT, White TP, Das D, Boks MP, Liddle PF. Reduced event-related low frequency EEG activity in schizophrenia during an auditory oddball task. Psychophysiology. 2009; 46(3):566–577. [PubMed: 19298628]

Durstewitz D, Seamans JK. The Dual-State Theory of Prefrontal Cortex Dopamine Function with Relevance to Catechol-OMethyltransferase Genotypes and Schizophrenia. Biological Psychiatry. 2008; 64:739–749. [PubMed: 18620336]

Fehr E, Camerer C. Social neuroeconomics: the neural circuitry of social preferences. Trends in Cognitive Science. 2007; 11(10):419–427.

Fett A-KJ, Shergill SS, Joyce DW, Riedl A, Strobel M, Gromann PM, Krabbendam L. To trust or not to trust: the dynamics of social interaction in psychosis. Brain. 2012; 135(3):976–984. [PubMed: 22366802]

Fletcher PC, Frith CD. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nature Reviews Neuroscience. 2009; 10(1):48–58. [PubMed: 19050712]

Fodor, JA. Concepts. Oxford University Press; UK: 1998.

Freeman D, Garety P, Kuipers E, Colbert S, Jolley S, Fowler D. Delusions and decision-making style: use of the Need for Closure Scale. Behavioural Research and Therapy. 2006; 44(8):1147–1158.

Freeman D, Pugh K, Garety P. Jumping to conclusions and paranoid ideation in the general population. Schizophrenia Research. 2008; 102(1–3):254–260. [PubMed: 18442898]

Friston K. Functional integration and inference in the brain. Progress in Neurobiology. 2002; 68(2):113–143. [PubMed: 12450490]

Frith CD, Done DJ. Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. Psychological Medicine. 1989; 19(2):359–363. [PubMed: 2762440]

Frith C. Explaining delusions of control: The comparator model 20 years on. Conscious and Cognition. 2011; 21(1):52–54.

Fudenberg D, Levine DK. Consistency and cautious fictitious play. Journal of Economic Dynamics and Control. 1995; 19:1065–1089.

Fudenberg, D., Levine, DK. The Theory of Learning in Games. MIT Press; Cambridge: 1998.

Gallagher HL, Jack AI, Roepstorff A, Frith CD. Imaging the intentional stance in a competitive game. NeuroImage. 2002; 16:814–821. [PubMed: 12169265]

Garety PA, Hemsley DR, Wessely S. Reasoning in deluded schizophrenic and paranoid patients. Biases in performance on a probabilistic inference task. Journal of Nervous and Mental Disease. 1991; 179(4):194–201. [PubMed: 2007889]

Garety PA, Freeman D. Cognitive approaches to delusions: a critical review of theories and evidence. British Journal of Clinical Psychology. 1999; 38:113–154. [PubMed: 10389596]

Gray JA, Feldon J, Rawlins JNP, Smith AD, Hemsley DR. The Neuropsychology of Schizophrenia. Behavioral and Brain Sciences. 1991; 14(1):1–19.

Hintikka, J., Hendricks, VF., Symons, J. Knowledge and Belief: An Introduction to the Logic of the Two Notions. King's College Publications; 2005.

Hollerman JR, Schultz W. Dopamine neurons report an error in the temporal prediction of reward during learning. Nature Neuroscience. 1998; 1(4):304–309. [PubMed: 10195164]

Howes OD, Kambeitz J, Kim E, Stahl D, Slifstein M, Abi-Dargham A, Kapur S. The Nature of Dopamine Dysfunction in Schizophrenia and What This Means for Treatment: Meta-analysis of Imaging Studies. Archives of General Psychiatry. 2012; 69(8):776–786. [PubMed: 22474070]

Huq SF, Garety PA, Hemsley DR. Probabilistic judgements in deluded and non-deluded subjects. Quarterly Journal of Experimental Psychology A. 1988; 40(4):801–812.

Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. American Journal of Psychiatry. 2003; 160(1):13–23. [PubMed: 12505794]

Kapur S, Mizrahi R, Li M. From dopamine to salience to psychosis-linking biology, pharmacology and phenomenology of psychosis. Schizophrenia Research. 2005; 79(1):59–68. [PubMed: 16005191]

King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR. Getting to know you: reputation and trust in a two-person economic exchange. Science. 2005; 308(5718):78–83. [PubMed: 15802598]

Koren D, Seidman LJ, Goldsmith M, Harvey PD. Real-world cognitive--and metacognitive-dysfunction in schizophrenia: a new approach for measuring (and remediating) more "right stuff". Schizophrenia Bulletin. 2006; 32(2):310–26. [PubMed: 16397202]

Koren D, Seidman LJ, Poyurovsky M, Goldsmith M, Viksman P, Zichel S. The neuropsychological basis of insight in first-episode schizophrenia: a pilot metacognitive study. Schizophrenia Research. 2004; 70(2–3):195–202. [PubMed: 15329296]

Koriat A, Ackerman R. Metacognition and mindreading: Judgements of learning for Self or Other during self-paced study. Consciousness and Cognition. 2009; 19:251–264.

Koriat, A. Metacognition and Consciousness. In: Zelazo, PD.Moscovitch, M., Thompson, E., editors. Cambridge handbook of consciousness. Cambridge University Press; New York, USA: 2007. p. 289-326.

Lysaker PH, Olesek KL, Warman DM, Martin JM, Salzman AK, Nicolo G, Salvatore G, Dimaggio G. Meta-cognition in schizophrenia: Correlates and stability of deficits in theory of mind and self reflectivity. Psychiatry Research. 2011; 190:18–22. [PubMed: 20696482]

Mehta MA, Manes FF, Magnolfi G, Sahakian BJ, Robbins TW. Impaired set-shifting and dissociable effects on tests of spatial working memory following the dopamine D2 receptor antagonist sulpiride in human volunteers. Psychopharmacology. 2004; 176:331–342. [PubMed: 15114435]

Metcalfe, J., Shimamura, AP. Metacognition: knowing about knowing. MIT Press; Cambridge: 1994.

Monestes JL, Villatte M, Moore A, Yon V, Loas G. Decisions in conditional situation and theory of mind in schizotypy. Encephale. 2008; 34(2):116–22. [PubMed: 18597718]

Moritz S, Woodward TS, Ruff CC. Source monitoring and memory confidence in schizophrenia. Psychological Medicine. 2003; 33:131–139. [PubMed: 12537044]

Moritz S, Woodward TS. The contribution of metamemory deficits to schizophrenia. Journal of Abnormal Psychology. 2006; 115:15–25. [PubMed: 16492092]

O'Daly OG, Joyce D, Stephan KE, Murray RM, Shergill SS. Functional magnetic resonance imaging investigation of the amphetamine sensitization model of schizophrenia in healthy male volunteers. Archives of General Psychiatry. 2011; 68(6):545–554. [PubMed: 21300940]

Paulus MP, Feinstein JS, Leland D, Simmons AN. Superior temporal gyrus and insula provide response and outcome-dependent information during assessment and action selection in a decision-making situation. NeuroImage. 2005; 25(2):607–615. [PubMed: 15784440]

Paulus MP, Feinstein JS, Tapert SF, Liu TT. Trend detection via temporal difference model predicts inferior prefrontal cortex activation during acquisition of advantageous action selection. NeuroImage. 2004; 21(2):733–743. [PubMed: 14980576]

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature. 2006; 442(7106):1042–1045. [PubMed: 16929307]

Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making. Nature Reviews Neuroscience. 2008; 9(7):545–556. [PubMed: 18545266]

Schneider U, Borsutzky M, Seifert J, Leweke FM, Huber TJ, Rollnik JD. Reduced binocular depth inversion in schizophrenic patients. Schizophrenia Research. 2002; 53(1–2):101–108. [PubMed: 11728843]

Sellen JL, Oaksford M, Gray NS. Schizotypy and conditional reasoning. Schizophrenia Bulletin. 2005; 31(1):105–116. [PubMed: 15888430]

Shergill SS, Samson G, Bays PM, Frith CD, Wolpert DM. Evidence for sensory prediction deficits in schizophrenia. American Journal of Psychiatry. 2005; 162(12):2384–2386. [PubMed: 16330607]

Simons CJ, Tracy DK, Sanghera KK, O'Daly O, Gilleen J, Dominguez MD, Krabbendam L, Shergill SS. Functional magnetic resonance imaging of inner speech in schizophrenia. Biological Psychiatry. 2010; 67(3):232–237. [PubMed: 19846064]

Smeets F, Lataster T, Dominguez MD, Hommes J, Lieb R, Wittchen HU. Evidence That Onset of Psychosis in the Population Reflects Early Hallucinatory Experiences That Through Environmental Risks and Affective Dysregulation Become Complicated by Delusions. Schizophrenia Bulletin. 2012; 38(3):531–542. [PubMed: 21030456]

Speechley WJ, Whitman JC, Woodward TS. The contribution of hypersalience to the "jumping to conclusions" bias associated with delusions in schizophrenia. Journal of Psychiatry and Neuroscience. 2010; 35(1):7–17. [PubMed: 20040242]

Startup H, Freeman D, Garety PA. Jumping to conclusions and persecutory delusions. European Psychiatry. 2008; 23(6):457–459. [PubMed: 18585008]

Sutton, R., Barto, A. Reinforcement Learning. MIT Press; Cambridge: 1998.

Waltz JA, Gold JM. Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. Schizophrenia Research. 2007; 93:296–303. [PubMed: 17482797]

Woodward TS, Moritz S, Cuttler C, Whitman JC. The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions in schizophrenia. Journal of Clinical Experimental Neuropsychology. 2006; 28(4):605–617. [PubMed: 16624787]

Woodward TS, Moritz S, Menon M, Klinge R. Belief inflexibility in schizophrenia. Cognitive Neuropsychiatry. 2008; 13(3):267–277. [PubMed: 18484291]

Woodward TS, Munz M, LeClerc C, Lecomte T. Change in delusions is associated with change in "jumping to conclusions". Psychiatry Research. 2009; 170(2–3):124–127. [PubMed: 19906443]
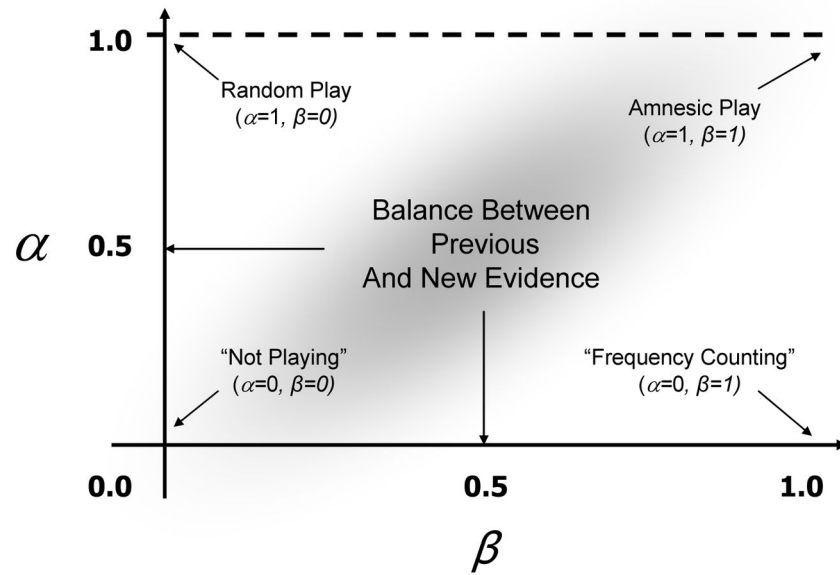
Figure 1. Parameter space for the action selection model

This diagram illustrates how strategy is updated based on differences between predicted and actual outcomes during different types of play where α models the decaying of prior evidence, and β is the weight given to new evidence.
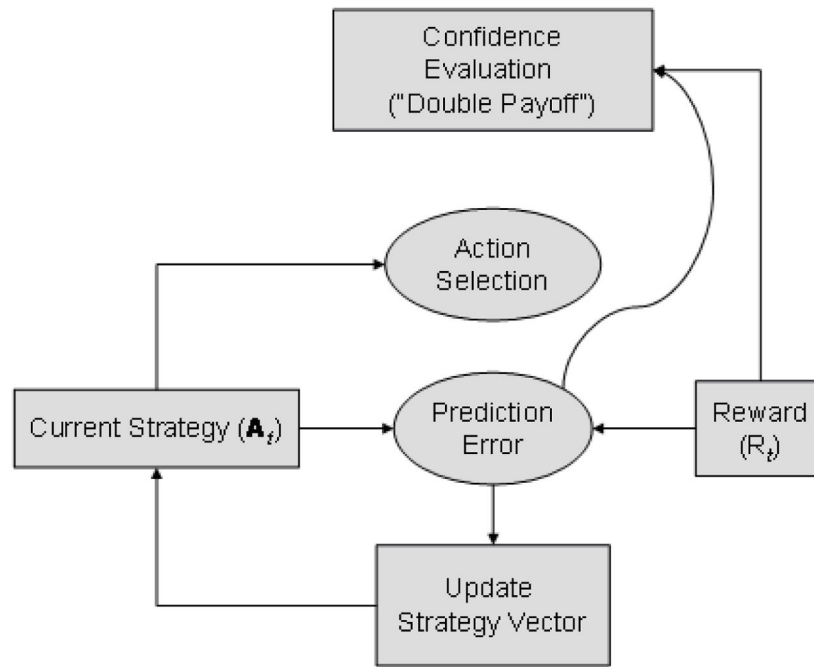
Figure 2. Performance evaluation models
This model demonstrates the strategy for each player over successive trials, where evidence for participants to confidently double payoff is derived from (**a**) prediction errors or separately, (**b**) from the absolute reward/payoff.
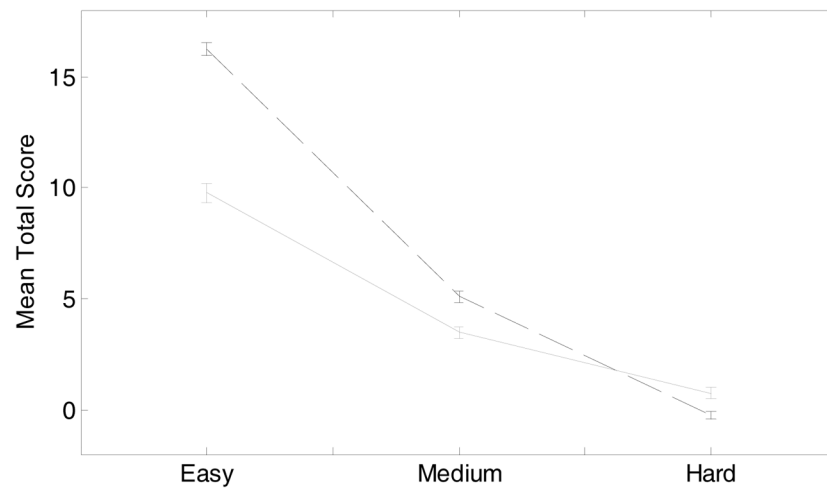
Figure 3. Average final scores on easy, medium and hard games

For each game, one point was gained when a subject won a trial or deducted if they lost the trial; no points were deducted for draws. This plot shows the final average scores across the difficulty level of each game, e.g. easy, medium and hard. Controls are shown with the dark dashed line and patients with the grey solid line. Error bars +/− 1 standard error.
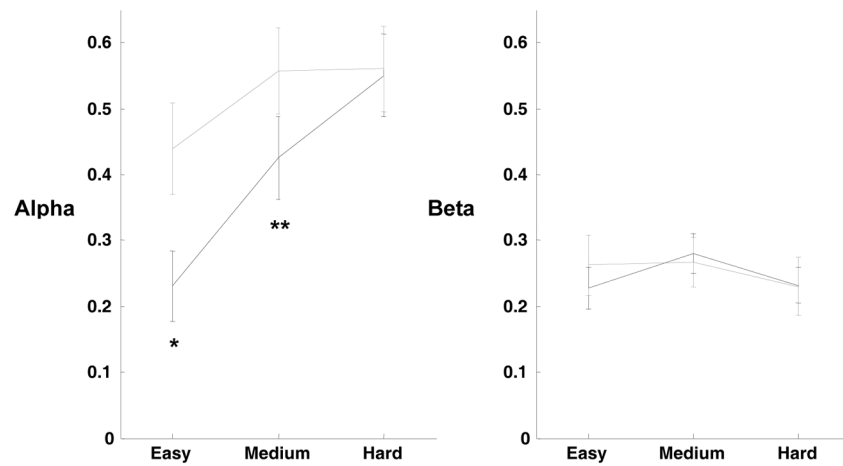
Figure 4. Model parameters across easy, medium and hard games

For each level of difficulty, easy, medium and hard, alpha models the decaying of prior evidence, and beta represents the weight given to new evidence by each subject. Controls are shown with the dark dashed line and patients with the grey solid line. Error bars are +/− 1 standard error. * significant diff. $p < 0.0004$; ** significant diff. $p < 0.022$.
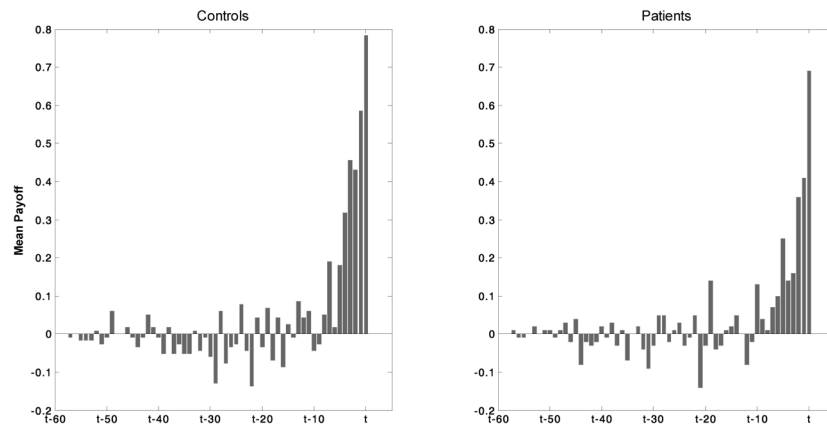
Figure 5. Mean outcome (reward/payoff) in trials preceding the decision to double payoffs (confidence);

The two histograms show a characteristic "upswing" where mean payoffs are increasing positive (i.e. wins greater than losses) over a time window of approximately 10 trials before the decision was made to "double bets".

Table 1

Patient and Control Demographics

	Patient group (N = 29, Male = 25)		Control group (N = 33, Male = 23)	
	Mean (s.d.)	Range	Mean (s.d.)	Range
Age (years)	40.8 (+/− 8.4)	24–60	34.9 (+/− 14)	18–59
Years in Full Time Education	13.3 (+/− 3.3)	10–21	11.95 (+/− 2.4)	9–20
Total PANSS score	54.1 (+/− 11.9)	37–80		
Positive	13.7 (+/− 5.4)	7–27		
Negative	13.6 (+/− 4.0)	7–22		
General	27.2 (+/− 6.0)	18–41		
Medication (Chlorpromazine equivalent mg)	219.8 (+/− 178.6)	0–600		