Cognitive Neuropsychiatry
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pcnp20

Bayesian modelling of Jumping-to-Conclusions bias in delusional patients

Michael Moutoussis [a], Richard P. Bentall [b], Wael El-Deredy [a] & Peter Dayan [c]

[a] School of Psychological Sciences, University of Manchester, Manchester, UK

[b] Department of Psychology, University of Bangor, Bangor, UK

[c] Gatsby Institute for Cognitive Neuroscience, University College London, London, UK

Available online: 07 Apr 2011

To cite this article: Michael Moutoussis, Richard P. Bentall, Wael El-Deredy & Peter Dayan (2011): Bayesian modelling of Jumping-to-Conclusions bias in delusional patients, Cognitive Neuropsychiatry, DOI:10.1080/13546805.2010.548678

To link to this article: http://dx.doi.org/10.1080/13546805.2010.548678

COGNITIVE NEUROPSYCHIATRY
0000, 00 (00), 1–26

Psychology Press
Taylor & Francis Group

Bayesian modelling of Jumping-to-Conclusions bias in delusional patients

Michael Moutoussis[1], Richard P. Bentall[2], Wael El-Deredy[1], and Peter Dayan[3]

[1]School of Psychological Sciences, University of Manchester, Manchester, UK
[2]Department of Psychology, University of Bangor, Bangor, UK
[3]Gatsby Institute for Cognitive Neuroscience, University College London, London, UK

Introduction. When deciding about the cause underlying serially presented events, patients with delusions utilise fewer events than controls, showing a "Jumping-to-Conclusions" bias. This has been widely hypothesised to be because patients expect to incur higher costs if they sample more information. This hypothesis is, however, unconfirmed.
Methods. The hypothesis was tested by analysing patient and control data using two models. The models provided explicit, quantitative variables characterising decision making. One model was based on calculating the potential costs of making a decision; the other compared a measure of certainty to a fixed threshold.
Results. Differences between paranoid participants and controls were found, but not in the way that was previously hypothesised. A greater "noise" in decision making (relative to the effective motivation to get the task right), rather than greater perceived costs, best accounted for group differences. Paranoid participants also deviated from ideal Bayesian reasoning more than healthy controls.
Conclusions. The Jumping-to-Conclusions Bias is unlikely to be due to an overestimation of the cost of gathering more information. The analytic approach we used, involving a Bayesian model to estimate the parameters characterising different participant populations, is well suited to testing hypotheses regarding "hidden" variables underpinning observed behaviours.

Keywords: Bayesian reasoning; Jumping-to-Conclusions; Paranoia; Psychosis; Sequential probability ratio test.

Correspondence should be addressed to Michael Moutoussis, 30 Melrose Avenue, Potters Bar, EN6 1TA, UK. E-mail: fzsemmo@gn.apc.org

We would like to thank Quentin Huys for his thoughtful comments. Funding was from the Gatsby Charitable Foundation (PD).

http://www.psypress.com/cogneuropsychiatry DOI: 10.1080/13546805.2010.548678

INTRODUCTION

Delusions are beliefs that appear fixed but are unwarranted on the basis of the available evidence (American Psychiatric Association, 2000). It has therefore been hypothesised that probabilistic inference is defective in deluded patients. A key test used to assess such inference is the "beads in a jar" task (Huq, Garety, & Hemsley, 1988), which abstracts the personal (including threat-related) content of delusions away from the inferencing process. Here, study participants are told that a sequence of blue (b) and green (g) beads will be drawn from one of two jars. One jar, B, has a majority of blue beads; the other jar, G, has the same majority of green beads. Participants are presented with beads, one by one. They are asked to think whether B or G is the underlying cause of the bead sequence and asked to declare which it is when they themselves are sure. The main outcome is the number of draws that participants take to decide.

A robust finding of such studies is that participants with paranoid beliefs or other kinds of delusions take a smaller number of draws to decide than controls (Corcoran et al., 2008; Fear & Healy, 1997; Garety et al., 2005; Garety, Hemsley, & Wessely, 1991). This is usually interpreted as a "Jumping to Conclusions (JTC) bias" and thus claimed to reflect an aspect of delusion formation. The unwarranted derivation of beliefs in paranoid thinking, as well as the precipitous decisions in probabilistic-reasoning tasks, are often attributed to increased motivation to draw conclusions, for example as part of a cognitive style geared towards detecting threats (Dudley, John, Young, & Over, 1997).

Further studies, however, showed that paranoid reasoning in the "beads task" shows modest abnormalities compared to what one might expect if delusions were largely due to jumping to conclusions (Fine, Gardner, Craigie, & Gold, 2007). This challenged the validity of the standard interpretation of precipitous decisions in the "beads task" in terms of the cognitive processes involved in paranoid inference. In one variant of the task, participants are asked to give serial estimates of the chance that the beads are drawn from one specific jar, rather than to decide when to stop drawing. In this version of the task paranoid subjects appear to shift their certainty estimates more than controls when presented with beads favouring the jar opposite to their currently preferred one (Fear & Healy, 1997; Young & Bentall, 1997). This would be surprising if developing a preference for a jar involved jumping to a delusion-like conclusion, as delusions are by definition resistant to contrary evidence (American Psychiatric Association, 2000).

We thus sought to take advantage of advances in the theory of decision making to develop a more refined view of the processes and mechanisms associated with the beads task. Our intent was to realise the constructs underlying decision making in a rigorous and testable ideal-observer, Bayesian framework that could potentially reveal even subtle biases and inferential flaws in performance on the existing task.

The ideal-observer Bayesian approach

Models of optimal decision making explicitly parametrise all the factors, including prior expectations, costs, and noise, that should control choice (Dayan & Daw, 2008; Green & Swets, 2008; Kording, 2007). They can be used to ask how subjects' behaviour approximates this optimal solution, and what might be the nature of any deviations from optimality. Decision making in ideal observer models is based on two considerations: the posterior probabilities of various scenarios, given what has been observed, and the values of different decisions (action values) for each scenario.

In tasks such as "beads-in-a-jar", the likelihoods can be estimated exactly using Bayes' theorem (Appendix, Equations 2–3). Fear and Healy (1997) compared the exact Bayesian likelihoods with estimates that patients and healthy controls made in the serial-estimates version of the task. No clear difference was found to support the hypothesis that delusion formation is based on a deficit of estimating such probabilities. Analysing the factors contributing to action values may therefore be quite important.

In the version of task where paranoid patients do differ from controls, actual choices or commitments must be made, rather than estimates of probabilities. After each sample, three choices are available: D_B and D_G, which decide on jars B and G respectively, and D_S, i.e., "sample again". In the decision-theory literature, these problems are called optimal stopping problems (Bertsekas, 1995; Puterman, 1994). The task instructions do not completely specify the costs or benefits of different actions. Participants are only told to sample beads until they are sure which jar they come from, up to a maximum $n_{max} = 20$. These instructions may suggest that the subjects should weigh heavily the cost of deciding wrongly (C^W). The cost that normally afflicts decision-making problems under uncertainty, namely the cost of sampling a further bead (C^S), is not mentioned.

However, C^S may nevertheless be important. One factor that has often been put forward to explain delusional thinking is inflated personal cost (or value) associated with the collection of information under uncertainty. Different instances of this "high-cost hypothesis" in the psychology literature include the "need for closure" reflecting a high subjective cost of uncertainty (Bentall & Swarbrick, 2003), the cost to self-esteem (Bentall, 2003), or the cost of cognitive dissonance experienced when intensely salient experiences strain the patient's explanatory theories about the world (Freeman & Garety, 2004; Kapur, 2003). A common theme is that delusional patients may experience sampling costs to be greater than healthy people, and adopt cognitive strategies to minimise them.

Our primary hypothesis was *that paranoid participants tend to assume higher costs of gathering more data in serial probabilistic inferencing tasks, which explains their early decisions* (the high-sampling-cost hypothesis). Our secondary hypothesis rendered more rigorous and testable the long-standing psychological hypothesis that paranoid participants may show deficiencies in Bayesian reasoning (Hemsley & Garety, 1986). We hypothesised that *paranoid probabilistic reasoning may deviate away from the Bayesian ideal (towards a simpler model) more than that of healthy subjects.* We thus built two models for the task. The first was an implementation of ideal observer Bayesian analysis, including Costs (the CB model). It is detailed in the Appendix; here we just note that its calculation of the cost of gathering more information (sampling again) involves a broad and deep consideration of all future outcomes (Appendix, Equations 6–8). We expect such calculations to be highly challenging for patients, and indeed controls, so that in vivo approximations may be used. Our simpler model was based on the Sequential Probability Ratio Test (SPRT, see later). This treats costs in a less direct way, and is much simpler in practice. We estimated all model parameters by using the Expectation-Maximisation (EM) algorithm (Dempster, Laird, & Rubin, 1977).

One critical deviation from ideal that we can expect from both controls and patients is behavioural noise (e.g., Pleskac, Dougherty, Rivadeneira, & Wallsten, 2009). Consider data from 99 trials of 33 healthy people performing the beads task that we will analyse later (Corcoran et al., 2008). For seven trials (from five participants), a decision was taken after the second sample, which was always discordant with the first. Thus, at that stage participants had no information at all as to the underlying jar, and yet they were so far from the maximum possible number of sampled beads (20, in this case), that this limit would be unlikely to be exerting an effect. Some type of process error or extraneous influence that we can subsume in the concept of noise must be having an effect. A standard manoeuvre to encompass such choices is to introduce behavioural noise by having subjects choose randomly between the three possible actions with probabilities depending on their relative action values. We assume that the impact of the "noise" is controlled by a temperature-like parameter, τ, via the sort of Softmax or Luce choice rule employed by a bulk of other models of human and animal decision making (e.g., O'Doherty et al., 2004).

The CB model would then have three parameters, τ, C^W, and C^S. However if τ, C^W, and C^S are all scaled by the same factor, the same probabilities will ensue. In other words, we need to choose the "currency" by which to measure these three parameters arbitrarily. We thus set the cost $C^W = 100$ for all participants; C^W can thus be seen as the "unit of internal cost" for each person, relative to which all other quantities are measured.

Comparison with the sequential probability ratio test

Optimal stopping problems frequently arise in psychological studies of decision making (Gold & Shadlen, 2001; Laming, 1968; Link, 1992; Ratcliff & Smith, 2004; Smith & Ratcliff, 2004; Usher & McClelland, 2001). Much of this work has been organised around the SPRT. In the SPRT one maintains the log-likelihood ratio, $l(n_d, n_g)$, that the beads come from jar G rather than B, if n_g green beads have been drawn out of n_d samples. Decisions are taken when $l(n_d, n_g)$ of two thresholds, θ_G or θ_B:

$$D = \begin{cases} D_G \; if \; l(n_d, n_g) > \theta_G \\ D_B \; if \; l(n_d, n_g) < \theta_B \\ D_S \; otherwise \end{cases} \quad (1)$$

We used thresholds equidistant from $l(n_d, n_g) = 0$ (the point of maximum uncertainty) i.e., $\theta_G = -\theta_B = \theta$. The thresholds encode the level of certainty (excess beads of one colour) a participant may demand to decide. In keeping with the previous observations, we also used an additive noise term (parameters: mean 0, variance τ^2) to account for behavioural noise (see Figure 1 and Appendix, Equations 9–11).

Given appropriate thresholds, the (noise-free) SPRT would produce exactly the same choices as the (noise-free) CB model under the assumption that there is no limit on the number of possible samples (Wald, 1945). This is a remarkable result, since the SPRT is computationally a vastly more straightforward implementation of an optimal policy than is known for almost any other case for CB, not requiring the tree of possible future states to be explicitly enumerated. However, it is not quite statistically optimal for the beads task, since n_{max} is not infinite. Noise may also perturb it in a different way. We used the SPRT to test whether some subjects use a decision-making strategy considerably simpler than the optimal.

METHODS

In order to test our hypotheses we applied the CB and SPRT models to reanalyse experimental results previously obtained by Corcoran et al. (2008).

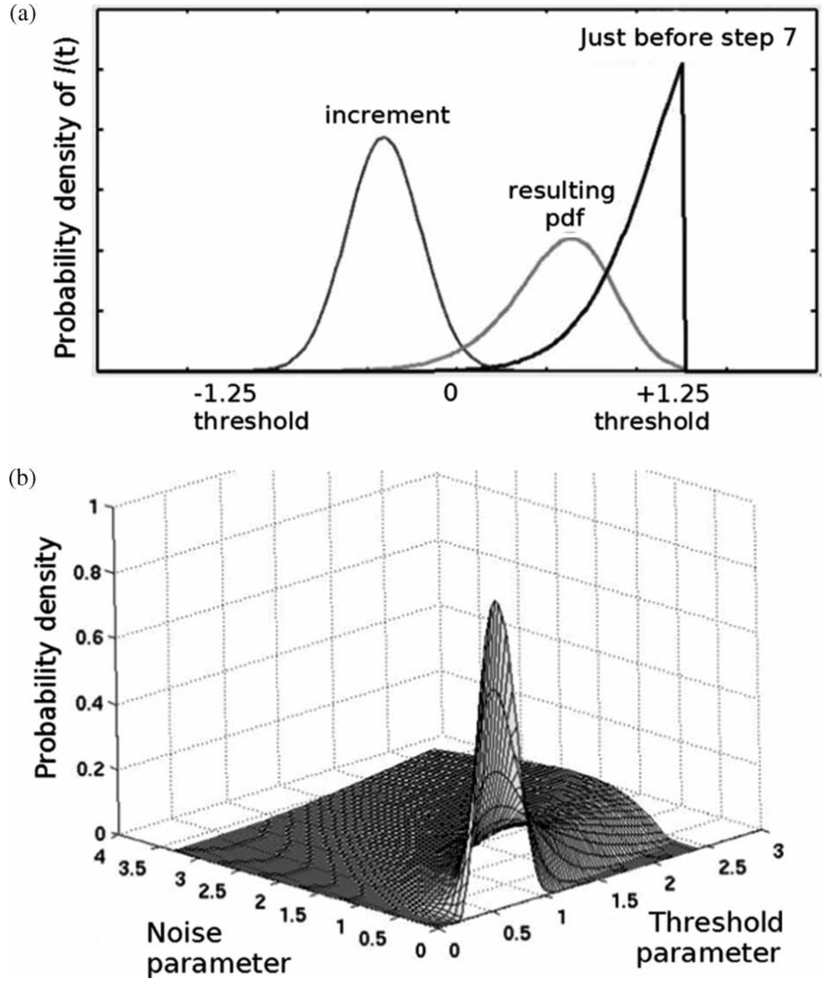

Figure 1. (a) SPRT simulation, draw 7 of sequence 1011110 ... (sequence 3). In this case, by draw 6 most participants have declared; only the lower tail of a Gaussian-like distribution remains, truncated at the high threshold $\theta_G = +1.25$ and renormalised before considering draw 7. This draw is discordant with the ones before (B), so that the increment in $l(t)$ is distributed as a Gaussian with a negative mean. The sum of the two random variables gives a left-skewed probability density function (pdf) relatively removed from both thresholds. (b) Probability of deciding at draw 4 as a function of threshold and noise in the estimation of the increment of $l(t)$. The distribution peaks at near-zero error. There it still has a finite width with respect to threshold values as the initial condition $l(0)$ has been set to contain a small noise term.

Participants

The study of Corcoran et al. (2008) included three diagnostic groups of participants aged over 65, and five groups under 65. Our hypotheses pertained to the group with persecutory delusions who were under 65 (original $N = 39$; we excluded three participants with very limited beads-task data, giving $N = 36$). We therefore analysed their data together with those of the under-65 healthy controls ($N = 33$). Persecutory delusions were judged to be present on the basis of endorsement of the question "Does anyone seem to be trying to harm you (trying to poison you or kill you?)" (World Health Organization, 1997), examination of case notes, and the answer to the question "Do you ever feel as if you are being persecuted in some way?" from the Peters et al. Delusions Inventory (Peters, Joseph, & Garety, 1999). We also performed limited analyses on depressed-without-delusions ($N = 26$), depressed-with-delusions ($N = 20$), and remitted-paranoid ($N = 29$) under-65 groups. The reader is referred to the original study for further details of the participants.

Stimuli

Participants were tested using the original version of the beads task (Garety et al., 1991), and a formally equivalent version of the task developed by Dudley et al. (1997) which uses valenced words rather than coloured beads. In the latter, participants had to choose between two surveys, each containing good and bad comments about an individual. They were told that each jar (or survey) contained 60% of the dominant colour (comment type). Each version used three particular sequences, (1) 01000010001011110111; (2) 01000100101000011001; and (3) 10111101110100001000 (using "0" and "1" to stand in for a particular colour or valence). Each participant thus provided six values of the number of beads or social words viewed before deciding, also called "draws-to-decision" (2 task versions, 3 sequences).

Analysis

We considered a pair of statistical models: one generative, and its statistical inverse, the recognition model. The generative model parametrises the process by which the experimental data are considered to have been generated. The recognition model takes the actual data from the participants and infers the parameters of the generative model that are likely to be responsible. We used the EM algorithm to fit model parameters to the experimental data. It is the "E" phase of the EM algorithm that involves the recognition model (Dempster et al., 1977).

In detail, for the generative model, we assumed that each experimental group was described by its own parametrised statistical prior distribution. We called these group-level descriptive statistics "macroparameters". These are the mean C^S_μ, standard deviation C^S_σ etc. The specific "microparameters" C^S_i (or, for the SPRT, $\theta_{i)}$ and τ_i characterising participant i are considered to be sampled from the distributions associated with the group of that participant. These microparameters act through the CB or SPRT model of the task to determine the distribution over possible experimental choices of participant i. As costs are unlikely to assume positive values (i.e., mistakes or slowness will not be positively rewarding) and uncertainty cannot be negative, we assumed that the "microparameters" are sampled from independent gamma distributions (in the case of costs with the sign "flipped" to negative values). We used the EM algorithm to find the values of the macroparameters that maximise the log-likelihood that the experimental data for each group could have been created by each model (CB or SPRT).

We assessed how well the models accounted for the data in several ways. First we used a parametric bootstrap resampling technique (Efron & Tibshirani, 1993). Here we used the best-fit parameter values to simulate the experiment of Corcoran et al. (2008) many times. The bootstrap tested if different models of decision making produced outcomes resembling the real data. The main outcome we examined was the difference between the mean draws-to-decision in the two participant groups. Second, in order to examine the null hypothesis (H_0) that the different groups could be described equally well by a single set of parameters, we merged the group data and fitted parameters to the combined set. We then compared the "merged" versus "separate" models using the Bayesian Information Criterion (BIC; Raftery, 1995; Schwarz, 1978). This is a measure based on model likelihood that takes into account the number of parameters a model uses (see Appendix). Third, we derived distributions of key parameter models through parametric bootstrap resampling. We thus estimated "bootstrap" confidence intervals (bCI) for the parameters fitted to each group. All programs used in the analyses are freely available from the authors on request and under General Public Licence (GNU GPL).

RESULTS

In general, both the SPRT and CB models produced good fits to the data. That is, when the fitted parameters were used to generate artificial data sets and these were reanalysed with the methods that we used for the original data, the sum-of-log-likelihoods for the actual data fell well within the distribution of values of the artificial data. Therefore the observed data could be a typical output of our generative model.

Most critically, the results were contrary to our high-sampling-cost hypothesis. In the beads task analysed with the CB model, the mean and variance of the cost of sampling converged to near-zero values for both healthy and paranoid groups (Table 1), consistent with the experimental instructions for the task. This finding challenged somewhat our use of a gamma probability distribution to fit the population distribution of the sampling-cost parameter, as the range of this distribution does not include zero itself. More importantly, it implied that we could simplify our CB models by removing the sampling cost variable (or by always setting $C^S = 0$), with a negligible reduction in model fit. The BIC value for this simplified model improved, as the model has two fewer parameters. Note that as the SPRT has no strictly equivalent or separate cost parameter, removing or fusing parameters to obtain a simpler model would be arbitrary.

TABLE 1
Best-fit parameters and BIC values for JTC tasks

			Beads task version		*BIC*
CB ($k = 4$)	τ_μ	τ_σ	C^S_μ	C^S_σ	
Healthy	5.09	3.63	−0.06	0.05	468.4
Remitted	5.75	4.06	−0.01	0.01	464.0
Paranoid	12.07	7.18	−0.05	0.03	423.9
CB ($k = 2$)	τ_μ	τ_σ			
Healthy	5.29	3.82	—	—	**459.3**
Remitted	5.74	4.06	—	—	457.3
Paranoid	12.09	7.09	—	—	416.0
SPRT ($k = 4$)	τ_μ	τ_σ	θ_μ	θ_σ	
Healthy	0.76	0.60	1.34	0.18	492.6
Remitted	0.74	0.64	1.49	0.76	***401.1***
Paranoid	1.53	0.90	1.20	0.38	***393.0***
			Words task version		
CB ($k = 4$)	τ_μ	τ_σ	C^S_μ	C^S_σ	
Healthy	5.91	3.83	−0.04	0.03	461.2
Remitted	8.05	7.45	−0.11	0.16	418.1
Paranoid	9.83	5.36	−0.04	0.03	423.8
CB ($k = 2$)	τ_μ	τ_σ			
Healthy	5.92	3.83	—	—	**454.8**
Remitted	8.25	7.51	—	—	423.2
Paranoid	9.83	5.36	—	—	416.6
SPRT ($k = 4$)	τ_μ	τ_σ	θ_μ	θ_σ	
Healthy	0.81	0.68	1.35	0.44	459.8
Remitted	0.87	0.53	1.41	0.81	**375.2**
Paranoid	1.39	0.62	1.4	0.73	**410.9**

The second important finding was that the paranoid group had a higher mean cognitive noise (parameter τ_μ) than both the healthy and clinical controls and for both versions of the task (Figure 2). This finding was statistically robust (see later). In the SPRT model the noise parameter was again larger for the paranoid than the healthy group (Table 1). The effect of the higher noise parameter τ_μ on decisions in the CB model is illustrated in Figure 3. This shows how action values for the three possible actions change through the experiment. Initially the "sample again" action is quite advantageous for the healthy group, but less so for the paranoid group. Two samples later only the healthy group, but not the paranoid one, still perceives an advantage in sampling again.

We used three different approaches to judge the statistical significance of the increased cognitive noise for paranoid group. First, we assumed that the all five groups were sampled from populations with similar noise structures (a null hypothesis, H_0). Hence, we merged the data-sets for all groups and fitted a merged-group set of CB parameters. We compared the ability of these merged-group-parameters to describe the data, as compared to the separate-group parameter fit shown in Table 1 (H_1).

We therefore simulated the experiments using merged versus separate parameters, thus creating large sets of simulated data. We compared three key descriptive statistics for the beads task under H_0 versus H_1: the grand mean difference in draws-to-decision between groups, averaged over the three experimental sequences used; and the within-sequence variances in draws-to-decisions in each group. We found that the merged-data model was quite unlikely ($p < .02$ for words, $p < .001$ for beads, two-tailed) to give rise to the experimentally observed difference in draws-to-decision. The latter was near the modes of the simulated distribution under H_1. Therefore H_0 is rejected in favour of H_1. Similarly, if the remitted and paranoid groups are merged and best-fit parameters derived, hypothesis H_0' that the difference in draws-to-decision between them arose by chance is again rejected for the beads task ($p \sim .0002$) and the words task ($p \sim .02$). We repeated this last analysis using the SPRT rather than the CB as generative model, and got essentially identical results. The beads-task results remain significant for the beads task under Bonferroni correction for multiple comparisons ($p < .001$), but the words-task results are reduced to (just) trend significance (both $p \sim .06$). If the same analysis is repeated for the difference between healthy and remitted groups, the null hypothesis that they are sampled from the same distribution cannot be rejected. The experimentally observed variance in draws-to-decision of the healthy group was accounted for equally well under H_0 or H_1. The variance of the paranoid group was also consistent between the H_1 model and experiment, if a single outlier result was excluded. The decisions of this participant (2, 20 and 20 draws-to-decision), are indeed extremely unlikely to be produced by a stable Bayesian or SPRT model.

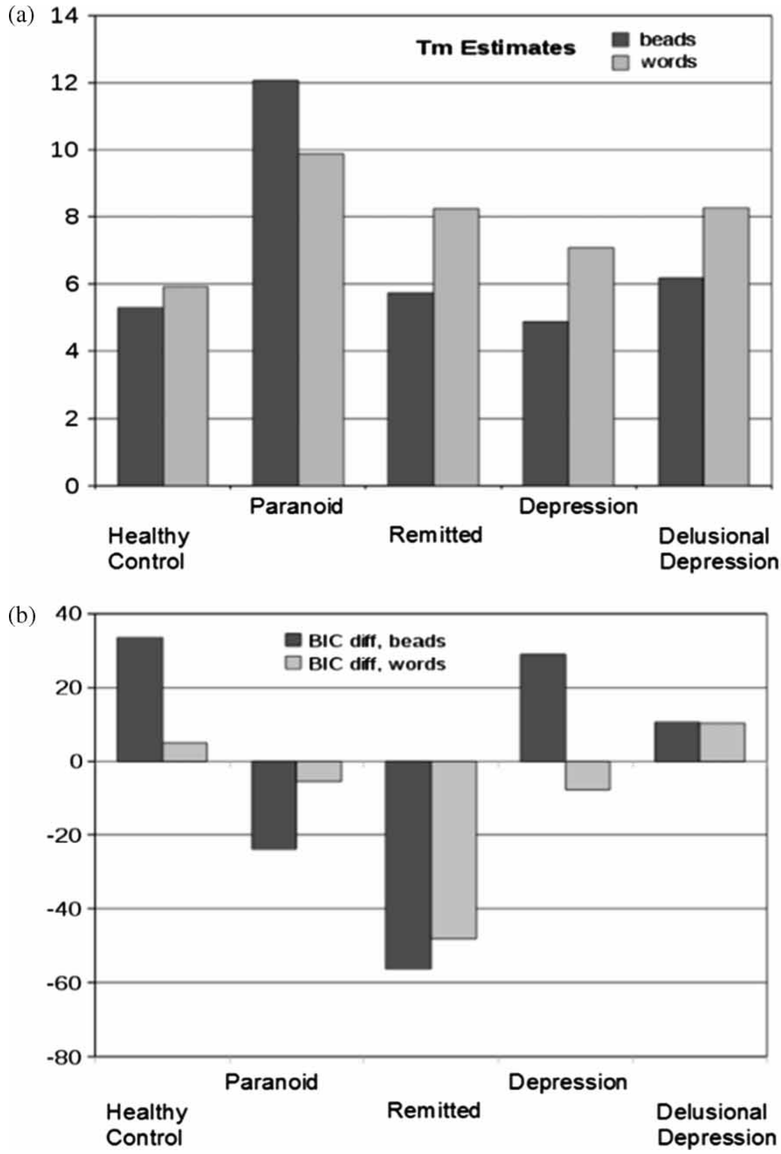

Figure 2. Group comparisons: (a) Best-fit noise parameters for the Bayesian model. The analogous noise parameter plot for the SPRT is almost identical (not shown). The paranoid group has larger mean noise, especially for the beads task. (b) Model fit according to the BIC. Positive values favour the Bayesian model, negative the SPRT. A difference of 10 is conventionally considered "very strong" evidence in favour of a model (Raftery, 1995). In the beads task, never-psychotic groups are closer to the Bayesian norm, whereas schizophrenia-spectrum-disorder groups are closer to the SPRT. The words task brings out differences less clearly, probably due to the complex social thinking it invites.

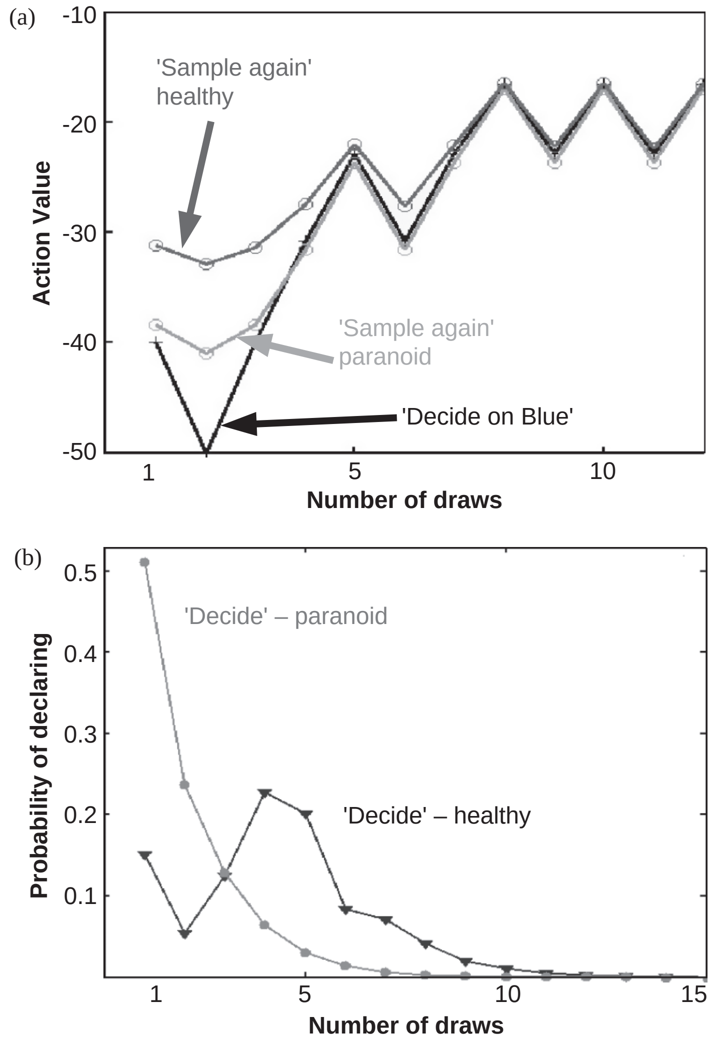

Figure 3. Action values and resulting probabilities-to-declare corresponding to the best-fit mean noise for the healthy and paranoid groups. A "representative participant" corresponding to the mean for each group is shown, for reasons of clarity. The sequence presented is bgbbbbgb ... (a) Action values curves for "sample again" and "decide on blue" for the two groups. The "decide on blue" curve is the same for both groups. (b) Resulting probabilities of deciding at specific stages. The peak is at 1 for the paranoid group but much later, at 4, for the controls.

Exclusion of this individual resulted in similar parameter estimates for the paranoid group but improved the model fit as might be expected. The other analyses reported here were not materially affected by excluding this outlier.

Second, we used the BIC to compare a CB model of the paranoid and healthy groups fitted separately, with one where these two groups were merged. The BIC penalises extra parameters substantially, but still slightly favoured fitting separate parameters to each group (BIC = 898 vs. 902 when τ_μ, τ_σ, C^S_μ, and C^S_σ are fitted). Third, we applied EM to obtain bootstrap-confidence-intervals (bCI) for the parameters. We found that the noise parameter τ_μ estimate for the healthy-beads group as well as the corresponding estimate for the remitted group fell outside the 0.01 bCI for the paranoid-beads group (correcting for multiple comparisons). The 0.01 bCI for τ_μ for the beads-version of the task for each group, however, each included the estimate for the words version for the paranoid group. Compared with the beads-version, the words-version looks as if it makes levels of cognitive noise more similar across groups, rendering differences nonsignificant by the measure of bCI.

Finally, we sought to compare which model, CB versus SPRT, best fitted the data of the healthy, paranoid, and remitted groups. Our hypothesis was that paranoid participants deviated from the Bayesian model more than healthy ones, and that this difference would be related to the paranoid state itself. For the healthy controls, the BIC favoured the τ_μ, τ_σ CB model over the τ_μ, τ_σ, θ_μ, θ_σ SPRT model for both (but especially the beads) versions of the task. For the paranoid group the SPRT did better (beads) or somewhat better (words) than the CB model (Table 1), in support of our hypothesis. However, the remitted group was clearly nearer to the SPRT (Figure 2), contrary to expectation.

DISCUSSION

We used a Bayesian approach to analyse the "beads in a jar" task, a popular tool used to assess probabilistic reasoning in patients with psychiatric disorders (Fear & Healy, 1997; Garety et al., 1991). It is important to elucidate the mechanisms that the task assesses, as the "jumping to conclusions" bias seen in this task has been postulated to have aetiological importance in delusions (Freeman et al., 1998). Such is its perceived importance that specific therapeutic procedures have been designed to correct this bias (Moritz & Woodward, 2007). Our work is the first to quantify subjective motivational factors (perceived cost) in this task, and to test the common assumption that cost considerations account for jumping to conclusions.

High-noise processing versus the high-sampling-cost hypothesis

Contrary to the "high sampling cost hypothesis", we found that increased noise in decision-making accounts robustly for the JTC bias. We measured noise relative to the subjective cost of making a wrong decision; an alternative interpretation of this result might be that paranoid patients utilise reduced effective costs of making wrong decisions. Distorted effective salience (Kapur, 2003) might conceivably make it difficult for paranoid patients to put cost estimates to good use. However, we consider this interpretation unlikely as paranoid participants tend be highly avoidant (Freeman, Garety, & Kuipers, 2001) and sensitive to failure experiences (Bentall & Kaney, 2005). In addition, patients with marked "negative" symptoms were excluded from this study. While the possibility of "reduced effective motivation" cannot be ruled out on the basis of this study, it can be expected that paranoid patients would be highly motivated to avoid failure experiences and thus would be unlikely to have a reduced cost of making the wrong decision. Our interpretation implies that this task should be compared with a control one where paranoid participants demonstrate equal motivation not to "get it wrong" as controls, irrespective of ability. Most importantly, experimentally manipulated (e.g., monetary) cost-of-sampling and cost-of-wrong-decision should be examined. We predict that increasing C^S relative to C^W would not, as might be expected from the high sampling-cost hypothesis, make healthy control data delusion-like, but inducing "noise" in selecting one of the three actions would. Modelling could allow the influence of control tasks and externally manipulated costs to be used to infer the relative value of the "personal" cost of error that we used as a comparator here.

We found no evidence that paranoid participants perceive increased costs in this task when given socially salient stimuli. This is consistent with other work (Warman, Lysaker, Martin, Davis, & Haudenschield, 2007). It may still be that the anticipation of high personal costs specifically contributes to the fixity of the self-referent ideas that paranoid participants hold. Future research should therefore examine probabilistic reasoning relevant to specific delusional beliefs. It could test how such beliefs may (or may not) shift in the face of different types of personally salient evidence. Applying a Bayesian approach would allow estimation of (1) prior probabilities of harm, (2) accuracy of derivation of posterior probabilities, (3) "internal/social" costs such as "if this belief is false, I must be mad", and (4) "external" costs such as "if I get it wrong, my persecutors will get me". A related direction for future research is the examination of asymmetric costs. In the case of paranoia, deciding that someone is trustworthy when they are not may incur a much greater immediate cost than the opposite error. Deciding that people are ill-disposed when they are not may be more costly in the long run.

Paranoid decision making and noise

Decisions may be affected by noise in two key ways which our Bayesian modelling helps to clarify. First, noise directly reduces the impact of a given difference between the values of the actions on choice. "Noisy" participants would declare more often even if faced with similar differences in action values favouring sampling again (Appendix, Equation 7). Second, early decisions in the beads task reflect smaller differences in action values favouring different actions, as per Figure 3. This is because the calculation of the action value for sampling relies on values of future states being taken into account accurately (Equation 7 feeds into Equation 8; this feeds into Equation 6 for the previous step). Note the assumption under the CB model that paranoid participants still perform optimal Bayesian reasoning given their view of future outcomes.

Of course, calculating the values of actions based on a search through a forward model is challenging. Humans probably carry out such searches to solve simpler tasks such the "Tower of London" (Marczewski, Linden, & Laroi, 2001), while rats may engage in forward searching in the course of goal-directed decision making (Daw, Niv, & Dayan, 2005; Dickinson & Balleine, 2002). In both cases, there is a critical role for areas of prefrontal cortex, and specific regions of the striatum (Balleine, Liljeholm, & Ostlund, 2009; Unterrainer & Owen, 2006). The schizophrenia-spectrum diagnoses associated with our paranoid group are thought to involve a relative hypofrontality, with a predisposing and/or consequent limbic hyperdopaminergia (Langdon, McKay, & Coltheart, 2008; Laruelle, 2008) and such pathological processes may contribute to our findings. Emotional factors could also contribute to the process substantially, if they involved a sense of greater proximity of threat. There is evidence that the latter shifts information processing away from frontal areas (Mobbs et al., 2007).

Here the research implication of the present study is that the psychological mechanisms causing higher "noise", including perception of threat, need to be elucidated. Furthermore, a beads-in-a-jar task could be used to separate deluded patients with respect to their level of cognitive "noise". This would require more trials-per-participant so as to enable accurate determination of each participant's individual noise level. Cognitive mechanisms underlying delusions in the presence of low noise may differ from those in the high-noise case. High noise in itself makes cognition inefficient, or even biased, as cognitively more distant alternatives cannot be taken into account well. Therefore, factors eventually found to increase this noise may be a target of therapeutic interventions. The high-noise explanation of the JTC phenomenon also suggests that "high-noise" paranoid subjects should be compared to nonparanoid participants with similar cognitive impairments in probabilistic reasoning.

Bayesian versus threshold-driven decisions in paranoia

We found some support for the hypothesis that paranoid inference deviates from the Bayesian ideal more than that of healthy subjects. People with paranoia may employ more often non-Bayesian reasoning, where the estimated likelihood of the cause of an event is simply compared to a threshold. Overall, the costed-Bayesian model fits the healthy subjects better, whereas the data from paranoid subjects (and remitted) were better explained by the simpler sequential probability ratio test model, which does not involve consideration of possible future outcomes. It could be that SPRT-type reasoning is a lowest common denominator mechanism, on which people improve by using a more Bayesian-like approximation. This would be consistent with evidence that paranoid subjects tend to revert more easily to simple heuristics (Glockner & Moritz, 2008). Bayesian reasoning, however, requires considerable cognitive resources.

The structure of decision making for the two models can be compared by using an "urgency plot". This shows an effective threshold for the Bayesian model, allowing clear comparison with the fixed threshold of the SPRT (Figure 4). The Bayesian and SPRT methods of estimation diverge most markedly for the last five draws, when the Bayesian model decides with greater urgency (fewer excess beads of one colour). We therefore suggest that future studies seeking to differentiate the types of human reasoning may employ shorter sequences, of only about 10 pieces of information, so as to bring the last few draws within the range actually chosen by participants.

Methodological advances

Assessment procedures should ideally measure accurately those factors which contribute substantially to pathological processes. Such assessment procedures would highlight in each individual patient causal factors that would make good targets for therapeutic intervention. Unfortunately assessment of paranoid ideation has not yet reached this stage. The specific model-based analysis that we have developed here is not as yet intended for clinical practice but for research into the cognitive biases and deficits contributing to paranoia.

Our study sits comfortably within the current trend in studies of decision making. These studies utilise models that quantitatively capture observable

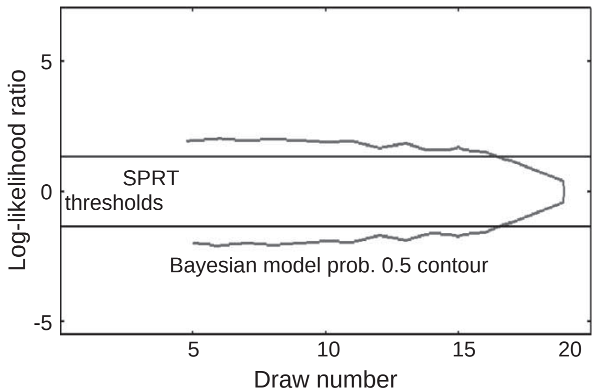

Figure 4. Comparison of SPRT and Bayesian models using the mean parameters for the healthy group doing the beads task. The SPRT model gives a decision when the estimated log-likelihood ratio crosses one of the two constant, symmetric boundaries. The Bayesian model does not have such fixed boundaries; an effective threshold, i.e., the log-likelihood ratio corresponding to a probability of deciding of .5, is plotted for comparison. This curve is not defined for 4 or fewer draws, as the probability of sampling again is always greater than .5 for these states. The Bayesian model estimates backwards, starting from the last draw, whereas the SPRT does not take account of the approaching end and this allows the greater "urgency" of the Bayesian model in the last few draws.

behaviour by postulating hidden psychological variables such as subjective beliefs and values that obey near-normative dynamics. Such variables are frequently the target of functional neuroimaging studies (Doya, 2002; O'Doherty et al., 2004) and offer accounts of neural activity in animals (e.g., Morris, Nevet, Arkadir, Vaadia, & Bergman, 2006; Schultz, Dayan, & Montague, 1997). These models are sufficiently precise to test and rule out important hypotheses such as the standard view of the motivational factors in JTC presented here. Similar approaches have been used in other psychiatric and neurological patient populations (Batchelder & Riefer, 2007; Busemeyer & Stout, 2002; Dayan & Huys, 2008; Frank, 2005; Kumar et al., 2008).

Building normative models allows for examination of different types of null hypothesis pertaining to the underlying variables *in silico*. It also allows checking that the "best-fit" model is likely, in absolute terms, to have produced the experimental results. This process showed that our model clearly separated the experimental groups. One further useful property of this form of modelling is that it provides a signature that the groups under study may not be well described by the unimodal distributions over model microparameters that we assumed as a starting point. Once the best-fit parameters for a group have been found, the posterior parameter distributions for each of the participants can be calculated and accumulated to produce what we called the "experimental Bayesian" distribution. This is sometimes called the marginal posterior density (Gelman, 2002) (Appendix, Equation 14). The form of this distribution provides hints as to extra structure in the data that is missed in the current model. As an illustration, we found evidence (Figure 5) that there is a subgroup of healthy subjects with very low error rates (and, on inspection, late decisions) who may have applied different heuristics, or indeed interpreted the rather vague instructions in a different way. It would be interesting to test task variants that probe these characteristics.

Limitations of the present study

Explicit, sequential Bayesian calculations are a competence rather than a performance model (Marr, 1982), and we have only been able to speculate in rather coarse terms about the (prefrontal) processes involved. The same is true for the many Bayesian models in modern computational cognitive science (Chater, Tenenbaum, & Yuille, 2006; Xu & Tenenbaum, 2007). There is a pressing need to study the approximations that biological systems may use to estimate posterior probabilities and related variables (Yu & Dayan, 2005). Like other ideal observer accounts, the costed-Bayesian model thus serves first and foremost as a point of reference.

It may be argued that the high-noise explanation is inconsistent with the finding that paranoid subjects underestimate their own uncertainties (Fine et al., 2007; Warman et al., 2007). Such a metacognitive deficit may, however, be quite consistent with schizophrenia being characterised by a poor perception of one's own mental function (Fletcher & Frith, 2009).

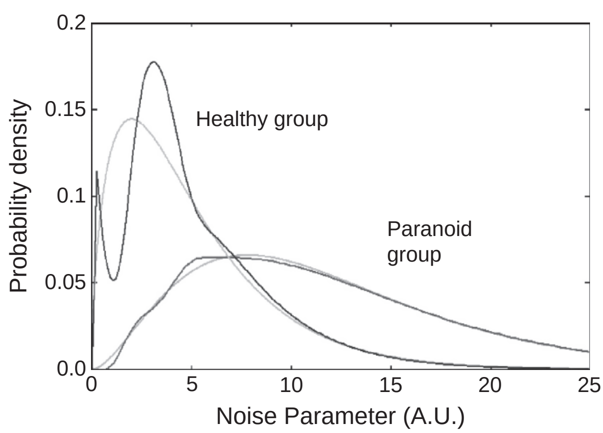

Figure 5. Best-fit gamma distributions (grey) and experimental Bayesian distributions (black) for the Beads data, for the healthy and paranoid groups. Bringing the data to bear does not alter the curve for the paranoid group much, but the control group appears to contain at least one subgroup characterised by a very small noise parameter.

Our study did not allow for explicit comparison with some systematic deviations from the Bayesian norm, such as the primacy effect. Future studies could include specific, psychologically motivated models of human heuristics as comparators to our noisy-Bayesian model. Our "high-noise" model predicts that delusional participants would make choices discordant with the most likely cause of sequences of information more often than controls. This is consistent with the literature (e.g., Fear & Healy, 1997) but the Corcoran et al. (2008) data did not allow a relevant analysis.

Methodologically, we note that the "bootstrap confidence intervals" that we used are not true confidence intervals, i.e., intervals such that if the true value of the parameter falls within the interval then the parameter estimate actually obtained would not be too improbable. We also note that the gamma distribution may not be optimal for describing cost parameters, as it can be poorly behaved very near zero.

It will be important to replicate the current analysis with other datasets and test our "high noise interpretation of JTC" with new data. Analyses could also include other groups, such as older participants, participants suffering from obsessive-compulsive disorder, etc.

SUMMARY AND CONCLUSIONS

We have introduced explicit models of how cost considerations and noise may skew probabilistic judgements, and we compared two types of model, the (optimal) costed-Bayesian and the sequential-probability-ratio test models. We applied an expectation-maximisation algorithm and analysis of synthetic data to estimate best-fit model parameters and choose the preferred model. We compared healthy and paranoid people as to their probabilistic reasoning. We found that the costed-Bayesian model gave a better account overall of the performance of healthy participants, whereas the SPRT fitted paranoid (and remitted) participants better. The commonly held hypothesis that paranoid people make early decisions through assuming a higher cost of gathering information was rejected. The most striking finding in both models was the much higher noise parameter for the paranoid group. Therefore, the "beads task" may best be seen as assessing not "jumping to conclusions" but executive functions subserving probabilistic reasoning. We suggest several new directions for methodological, computational, and experimental research. Based on modelling, we suggest that the "beads task" should be used with shorter sequences (e.g., maximum of 10 draws), more trials-per-participant, and experimentally manipulated rewarding and aversive returns. Interpretation of task results in terms of underlying decision mechanisms (costs, noise, and decision thresholds) has the potential to increase the construct validity of the task and eventually even to render it more relevant to assessing decision making in clinical situations.

Manuscript received 19 January 2010
Revised manuscript received 9 November 2010
First published online month/year

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Rev.; DSM-IV-TR). Washington, DC: Author.

Balleine, B. W., Liljeholm, M., & Ostlund, S. B. (2009). The integrative function of the basal ganglia in instrumental conditioning. *Behavioural Brain Research*, *199*(1), 43–52.

Batchelder, W. H., & Riefer, D. M. (2007). Using multinomial processing tree models to measure cognitive deficits in clinical populations. In R. Neufeld (Ed.), *Advances in clinical cognitive science* (pp. 19–50). Washington, DC: American Psychiatric Association.

Bentall, R. P. (2003). *Madness explained: Psychosis and human nature*. London, UK: Penguin.

Bentall, R. P., & Kaney, S. (2005). Attributional lability in paranoia and depression. *British Journal of Clinical Psychology*, *44*, 475–488.

Bentall, R. P., & Swarbrick, R. (2003). The best laid schemas of paranoid patients: Autonomy, sociotropy and need for closure. *Psychology and Psychotherapy*, *76*(2), 163–171.

Bertsekas, D. (1995). *Dynamic programming and optimal control*. Belmont, MA: Athena Scientific.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment*, *14*(3), 253–262.

Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences*, *10*(7), 287–291.

Corcoran, R., Rowse, G., Moore, R., Blackwood, N., Kinderman, P., & Howard, R. (2008). A transdiagnostic investigation of "theory of mind" and "jumping to conclusions" in patients with persecutory delusions. *Psychological Medicine*, *38*(11), 1577–1583.

Daw, N., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Reviews Neuroscience*, *8*(12), 1704–1711.

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, and Behavioral Neuroscience*, *8*(4), 429–453.

Dayan, P., & Huys, Q. J. (2008). Serotonin, inhibition, and negative mood. *PLoS Computational Biology*, *4*(2), e4.

Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, *39B*(1), 1–38.

Dickinson, A., & Balleine, B. (2002). The role of learning in motivation. In C. Gallistel (Ed.), *Steven's handbook of experimental psychology: Vol. 3. Learning, motivation and emotion*. New York, NY: Wiley.

Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks*, *15*(4–6), 495–506.

Dudley, R. E., John, C. H., Young, A. W., & Over, D. E. (1997). The effect of self-referent material on the reasoning of people with delusions. *British Journal of Clinical Psychology*, *36*(4), 575–584.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Boca Raton, FL: CRC Press.

Fear, C. F., & Healy, D. (1997). Probabilistic reasoning in obsessive-compulsive and delusional disorders. *Psychological Medicine*, *27*(1), 199–208.

Fine, C., Gardner, M., Craigie, J., & Gold, I. (2007). Hopping, skipping or jumping to conclusions? Clarifying the role of the JTC bias in delusions. *Cognitive Neuropsychiatry*, *12*(1), 46–77.
Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, *10*(1), 48–58.
Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, *17*(1), 51–72.
Freeman, D., & Garety, P. A. (2004). Delusions and disconfirmatory evidence. In *Paranoia: The psychology of persecutory delusions* (pp. 61–75). Hove, UK: Psychology Press.
Freeman, D., Garety, P., Fowler, D., Kuipers, E., Dunn, G., & Bebbington, P. (1998). The London-East Anglia randomized controlled trial of cognitive-behaviour therapy for psychosis. IV: Self-esteem and persecutory delusions. *British Journal of Clinical Psychology*, *37*(4), 415–430.
Freeman, D., Garety, P. A., & Kuipers, E. (2001). Persecutory delusions: Developing the understanding of belief maintenance and emotional distress. *Psychological Medicine*, *31*(7), 1293–1306.
Garety, P. A., Freeman, D., Jolley, S., Dunn, G., Bebbington, P. E., & Fowler, D. G. (2005). Reasoning, emotions, and delusional conviction in psychosis. *Journal of Abnormal Psychology*, *114*(3), 373–384.
Garety, P. A., Hemsley, D. R., & Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients: Biases in performance on a probabilistic inference task. *Journal of Nervous and Mental Disease*, *179*(4), 194–201.
Gelman, A. (2002). Posterior distribution. In A. El-Shaarawi & W. Piegorsch (Eds.), *Encyclopedia of environmetrics* (Vol. 3, pp. 1627–1628). Chichester, UK: Wiley.
Glockner, A., & Moritz, S. (2008). *A fine-grained analysis of the jumping to conclusions bias in schizophrenia: Data-gathering, response confidence and information integration* (Preprints of the Max Planck Institute for Research on Collective Goods). Bonn, Germany: Max Planck Institute for Research on Collective Goods.
Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, *5*(1), 10–16.
Green, D., & Swets, J. (2008). *Signal detection theory and psychophysics*. Los Altos, CA: Peninsula Publishing.
Hemsley, D. R., & Garety, P. A. (1986). The formation of maintenance of delusions: A Bayesian analysis. *British Journal of Psychiatry*, *149*, 51–56.
Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *Quarterly Journal of Experimental Psychology*, *40A*(4), 801–812.
Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, *160*(1), 13–23.
Kording, K. (2007). Decision theory: What "should" the nervous system do? *Science*, *318*(5850), 606–610.
Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., & Steele, J. D. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain*, *131*(8), 2084–2093.
Laming, D. (1968). *Information theory of choice reaction time*. New York, NY: Wiley.
Langdon, R., McKay, R., & Coltheart, M. (2008). The cognitive neuropsychological understanding of persecutory delusions. In D. Freeman, R. Bentall, & P. Garety (Eds.), *Persecutory delusions: Assessment, theory and treatment* (pp. 221–236). Oxford, UK: Oxford University Press.
Laruelle, M. (2008). Dopamine and persecutory delusions. In D. Freeman, R. Bentall, & P. Garety (Eds.), *Persecutory delusions: Assessment, theory and treatment* (pp. 239–266). Oxford, UK: Oxford University Press.

Link, S. (1992). *The wave theory of difference and similarity*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Marczewski, P., van der Linden, M., & Laroi, F. (2001). Further investigation of the supervisory attentional system in schizophrenia: Planning, inhibition, and rule abstraction. *Cognitive Neuropsychiatry*, *6*(3), 175–192.

Marr, D. (1982). *Vision*. New York, NY: Freeman.

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., & Seymour, B. (2007). When fear is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science*, *317*(5841), 1079–1083.

Moritz, S., & Woodward, T. S. (2007). Metacognitive training in schizophrenia: From basic research to knowledge translation and intervention. *Current Opinions in Psychiatry*, *20*(6), 619–625.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nature Neuroscience*, *9*(8), 1057–1063.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*, 452–454.

Peters, E., Joseph, S., & Garety, P. (1999). Measurement of delusional ideation in the normal population: Introducing the PDI. *Schizophrenia Bulletin*, *25*, 553–576.

Pleskac, T. J., Dougherty, M. P., Rivadeneira, A. W., & Wallsten, T. S. (2009). Random error in judgment: The contribution of encoding and retrieval processes. *Journal of Memory and Language*, *60*, 165–179.

Puterman, M. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. New York, NY: Wiley.

Raftery, A. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333–367.

Schultz, W., Dayan, P., & Montague, P. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neuroscience*, *27*(3), 161–168.

Unterrainer, J. M., & Owen, A. M. (2006). Planning and problem solving: From neuropsychology to functional neuroimaging. *Journal of Physiology, Paris*, *99*(4–6), 308–317.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592.

Wald, A. (1945). Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, *16*, 117–186.

Warman, D. M., Lysaker, P. H., Martin, J. M., Davis, L., & Haudenschield, S. L. (2007). Jumping to conclusions and the continuum of delusional beliefs. *Behavior Research and Therapy*, *45*(6), 1255–1269.

World Health Organization. (1997). *Schedules for clinical assessment in neuropsychiatry* (SCAN version, 2nd ed.). Geneva, Switzerland: Author.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272.

Young, H. F., & Bentall, R. P. (1997). Probabilistic reasoning in deluded, depressed and normal subjects: Effects of task difficulty and meaningful versus non-meaningful material. *Psychological Medicine*, *27*(2), 455–465.

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692.

APPENDIX

Let $q > .5$ be the proportion of blue beads in jar B. In other words, the probability of drawing a blue ball given that the jar of origin is B is P (b|B) = q . Similarly for the "Green" jar, P(g|G) = q . The prior probabilities that the beads come from either jar before any beads are drawn are the same, P(B|0,0) = P(G|0,0) = .5 (in the notation "|0,0" the first 0 means that $n_d = 0$ draws have taken place, the second 0 that $n_g = 0$ g[reen] balls have been drawn").

Costed Bayesian model

By Bayes theorem, the posterior probability of jar G being the underlying cause if n_d beads have been drawn, of which n_g were of type g, is:

$$P(G|n_d, n_g) = \frac{P(n_d, n_g|G)P(G|0,0)}{P(n_d, n_g|G)P(G|0,0) + P(n_d, n_g|B)P(B|0,0)} \quad (2)$$

Given the earlier notation and conventions, this becomes

$$P(G|n_d, n_g) = \frac{1}{1 + \left(\frac{q}{1-q}\right)^{(n_d - 2n_g)}} \quad (3)$$

The possible actions to take are D_B, D_G, and D_S, which are respectively *D*eciding on Blue, Green, or Sampling. At the last step, $n_d = 20$, decisions D_G and D_G are the only options. In all other steps D_S is also possible. We define the action value $Q(a\,;s)$ as the average long-term return of taking action a in state s, and $m(a\,;\,s)$ as the probability with which the participant will take action a in state s. Note that the symbol "|" denotes conditional probability, whereas expressions following ";" *denote a known state. The set of all* $m(a\,;s)$ defines the current policy that the agent pursues. If we take the return of choosing the correct decision to be zero and the cost of deciding erroneously to be C^W, we have:

$$Q(D_B; n_d, n_g) = C^W \times P(G|n_d, n_g) \quad (4)$$

$$Q(D_G; n_d, n_g) = C^W \times (1 - P(G|n_d, n_g)) \quad (5)$$

The immediate cost of taking a sample is C^S. However, since sampling leads the participant to further choices at further states, the full cost is a function of the values, $V(s)$, of those states s, weighted by the probabilities of getting to those states. The value here is defined as the average long-term return to be expected if one found oneself in state s and followed the current policy. The possible outcomes of sampling are that either b or g will turn up. In case g turns up, the participant will find themselves in state ($n_d + 1$, $n_g +$

1). Otherwise the state will be $(n_d + 1, n_g)$. Let $V(n_d + 1, n_g + 1)$ and $V(n_d + 1, n_g)$ be the (as yet unknown) values of these states. If the true underlying cause is G, the latter value will obtain with probability $P(b|G) = 1 - q$; the former value, with probability $P(g|G) = q$. In addition, the underlying cause is G with probability $P(G \mid n_d, n_g)$. Adding all the contributions together,

$$\begin{aligned} Q(D_S; n_d, n_g) = C^S + C^W \times P(G|n_d, n_g) \\ P(G|n_d, n_g) \times [V(n_d + 1, n_g + 1)P(g|G) + V(n_d + 1, n_g)P(b|G)] + \\ P(B|n_d, n_g) \times [V(n_d + 1, n_g + 1)P(g|B) + V(n_d + 1, n_g)P(b|B)] \end{aligned} \tag{6}$$

An idealised participant would always choose the decision with the lowest cost in each state. However, people do not choose so deterministically. To introduce behavioural uncertainty one can allow the model to choose randomly amongst the actions, but weighted by their action values. The Softmax choice function defines the probability of choosing action a at state $s = (n_d, n_g)$ as:

$$m(a\ ;\ s) = \frac{e^{Q(a,s)/\tau}}{\sum_{b \in \{D_G, D_B, D_S\}} e^{Q(b,s)/\tau}} \tag{7}$$

(except for $n_d = 20$, when $a,b \in \{D_B, D_G\}$). This includes the temperature-like parameter τ. A large value of τ means that given differences in action values have less impact on choice probabilities (more noisy decisions). Then the average value of each state is:

$$\begin{aligned} V(n_d, n_g) = m(D_S; n_d, n_g) \times Q(D_S; n_d, n_g) \\ + m(D_G; n_d, n_g) \times Q(D_G; n_d, n_g) \\ + m(D_B; n_d, n_g) \times Q(D_B; n_d, n_g) \end{aligned} \tag{8}$$

At the last step, $n_d = n_{max} = 20$ the action "Sample again" is unavailable, and Equation 8 has no term associated with D_S. Once all the values $V(n_{max}, n_{max} - n_g)$ can been calculated (with the help of Equations 5, 7, and 8), these can be used in to obtain all action values for $n_d = 19$, and so on back to $n_d = 1$. For each possible sequence of beads, therefore, a number of draws n_{dec} can be found where D_S ceases to be the most rewarding action. If the same action values were kept but behavioural noise was eliminated, this would become the step where the model would "declare".

Sequential probability ratio test model

In the SPRT, the key computed quantity at time t is the log-likelihood ratio of the sequence of data that have been produced given one possible cause, over the equivalent expression given the other cause:

$$l(t) = \ln \frac{P(d_1, d_2, \ldots | cause = G)}{P(d_1, d_2, \ldots | cause = B)} \quad (9)$$

Decisions are taken by comparing $l(t)$ to two thresholds, θ_G and θ_B, as described in the main text. We used $\theta_G = -\theta_B$ for consistency with using a single cost of making the wrong decision (C^W above). Since the draws are independent (which amounts to assuming that the jars contain many more beads than the participant is allowed to draw), $l(t)$ accumulates additively. Following the rationale described for the Bayesian model, we also used an uncertainty or noise factor:

$$l(t) = l(t-1) + \ln \frac{P(d_t|G)}{P(d_t|B)} + \varepsilon \quad (10)$$

We assumed that the noise factor is normally distributed with mean zero and standard deviation τ. In our simulations we also included a very small amount of noise at the starting point $t = 0$. In the deterministic case it would be easy to increment $l(t)$ by adding the penultimate term of Equation 10. In the case involving noise we have to add to the random variable $l(t-1)$ the last two terms of Equation 10, which also form a random variable. We thus find the probability distribution of l by convolution:

$$p_t(l) = \int_{-\infty}^{\infty} p_{t-1}(v) p_t(l - v) dv \quad (11)$$

Where p_{t-1} is the prior pdf of l and p_t is the probability distribution of the increment for the current step, t (i.e., the pdf of the last two terms in Equation 10). At each step, the values of l that fall outside the thresholds result in a decision, so that the probability distribution is truncated at the thresholds. It is then renormalised to compute the distribution of l for which the decision was "sample again". The latter distribution forms the starting point for the next step (Figure 1). The case involving noise or error is more demanding to implement efficiently on a computer than the noise-free case.

Model comparison and the Bayesian Information Criterion (BIC)

No single way of selecting a preferred model is always best. One of our comparisons asked which model gave a better account of the data of a particular group for a particular condition. Another comparison involved different groupings of the data. One therefore needed a basic measure that would be applicable across models of different structure. Model likelihood is such a measure: This is the probability that the experimental data would arise if the model in question was the mechanism giving rise to them. As the data for each participant arise independently within each group, the likelihood of model M for a specific group is:

$$L(M,\boldsymbol{D}) = P(\boldsymbol{D},M)\prod_{alli} P(\mathbf{d}_i M) \tag{12}$$

Where **D** is the matrix of all the data of the group in question and $\mathbf{d}_i$ is the vector of data for subject i. However, a model with more free parameters, i.e., less parsimonious, should fit the data better. The BIC combines the model likelihood L with a penalty for the number of parameters k used, taking account also of the number of data points N that are to be explained (Schwarz, 1978):

$$BIC = -2\ln L + k\ln N \tag{13}$$

Given two models applied to the same data, the one with the lower value of BIC is to be preferred.

Experimental Bayesian distribution

Once a particular model M is chosen, including macroparameters that describe the experimental group j, the recognition distribution density $p(\mathbf{g}|\mathbf{d}_i; M)$ describes how likely it that the participant i, who has furnished data $\mathbf{d}_i$, is characterised by microparameters **g**. We can now average this recognition distribution over all participants i in group j to obtain an estimate of the probability density $p_{exp}(\mathbf{g}_j)$ of microparameters for a random member of this group. This is sometimes called a marginal posterior density (Gelman, 2002); but, for clarity, we call it an experimental (Bayesian) distribution:

$$p_{\text{exp}}(\mathbf{g}_j) = \frac{1}{N_j}\sum_{i=1}^{N_j} \frac{P(\mathbf{d}_i|\mathbf{g}_j; M)p(\mathbf{g}_j; M)}{\int_{all\,\mathbf{g}} P(\mathbf{d}_i|\mathbf{g}; M)p(\mathbf{g}; M)d\mathbf{g}} \tag{14}$$