

Probabilistic reasoning in schizophrenia is volatile but not biased

Gerit Pfuhl (gerit.pfuhl@uit.no)

Psychology Department, Teorifagbygget hus 6, Huginbakken 32
UiT The Arctic University of Norway, Tromsø, Norway

Håkon Tjelmeland (haakon.tjelmeland@ntnu.no)

Department of Mathematical Sciences, Sentralbygg 2, Alfred Getz vei 1
NTNU, Trondheim, Norway

Abstract

We update our beliefs based on evidence. Aberrant belief updating has been linked to schizophrenia and autism. It is not clear whether the faulty updating is due to reduced general cognitive abilities, overweighting of recent information, or lower thresholds for switching from one belief to another. A common task to assess belief updating is the beads task. Patients with schizophrenia show hasty decision-making.

We here present a model describing the deviations from an ideal Bayesian observer and apply the model to three independent datasets, totalling $n=176$ healthy controls and $n=128$ patients with schizophrenia. The parameters describe a) the number of beads considered (memory), b) systematic deviation and c) unsystematic deviations (volatility) from probability estimates.

We find that, on average, patients use fewer beads and or more volatile responding. However, patients have, on average, probability estimates that are closer to the true probabilities. Closer investigations yielded relevant differences among the datasets and sequences used. More challenging sequences improve the performance of patients.

Our model captures well the cognitive mechanisms proposed to contribute to the performance differences in the beads task.

Keywords: decision making; computational psychiatry; beads task; urn task; jumping to conclusion

Background

You see dark clouds, how probable is it that it will rain? Such probability estimates require integration of previous evidence, and appropriate updating of the odds that it will rain. A commonly used paradigm to assess probabilistic reasoning in the lab is the beads task Phillips and Edwards (1966). Here, participants are presented with jars that contain coloured beads. In the standard version two jars are presented, each containing opposite ratios of beads e.g. 80% black and 20% white and vice versa). One bead at a time is drawn with replacement. Participants make probability estimates, also referred to as graded estimates.

A variant of the task, draws to decision, is often used in clinical settings. A prominent finding is that patients with schizophrenia, particularly those with delusions, decide after a few beads from which jar the beads are coming e.g. Huq, Garety, and

Hemsley (1988). This has been named the jumping to conclusion bias. Performance in the beads task has been attributed to a) reduced general cognitive abilities, i.e. a lower working memory capacity, b) overweighting of the most recent information and underweighting of previous information, or c) lower thresholds for switching from one belief to another (Moritz & Woodward, 2005).

Without a mathematical model one cannot easily distinguish the contribution of those accounts on the jumping to conclusion bias. We here describe a model that uses the graded estimate version and asks whether patients show more aberrant probability estimates than healthy controls, and whether this is due to basing the inference on fewer evidence. This is an important question, as this allows identifying a cognitive mechanism for the performance difference found in the beads task.

Methods

We analysed three datasets that did assess probability estimates in the beads task. Dataset 1 was kindly provided by S. Moritz. They tested 62 patients with a diagnosis of schizophrenia and 30 healthy controls (S. Moritz et al., 2016). There were two test sessions, but we use only the first test session. We use here only the first two sequences, a symmetric 80:20 and an asymmetric 50:50 vs 80:20 sequence. Participants made both a draws to decision judgement, as well as providing probability estimates for each of the two jars.

Dataset 2 is from Peters and Garety (2006), referred to here as P&G, and kindly provided by Rick Adams. We used the first session, which had 23 patients with schizophrenia, and 35 healthy controls. We did not include the non-psychotic patient group here. P&G used a symmetrical 85:15 sequence and participants estimated the probabilities on a 0 to 100 scale.

Dataset 3 is from Adams, Napier, Roiser, Mathys, and Gilleen (2018), referred to here as Gilleen. They tested 56 patients with schizophrenia and 111 healthy controls, all tested only once. There are four sequences made of two identical pairs (ratio 80:20), and the order of presenting the four sequences was randomised. Probability estimates were given on a 7 point scale from sure it is jar A, indifferent whether it is jar A or jar B, to sure it is jar B; and only an estimate for one jar was made.

Common is that a total of 10 beads were shown sequentially, and beads were drawn with replacement, i.e. the proportion of beads in the jars did not change.



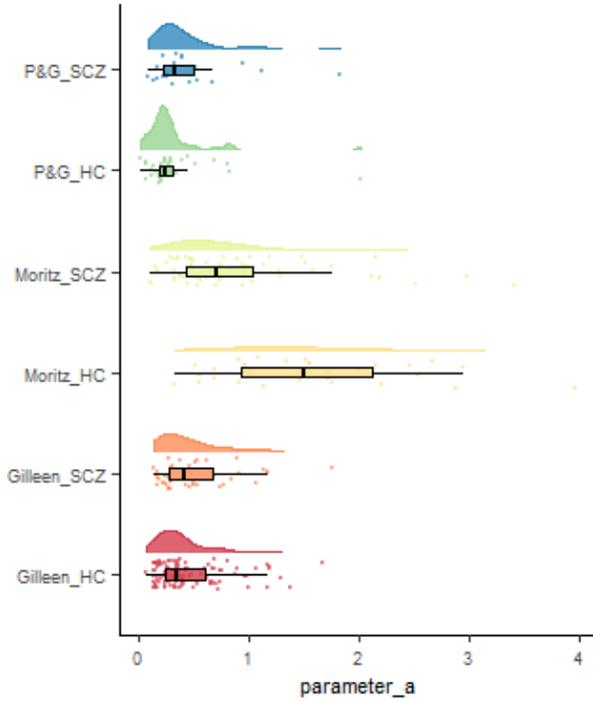


Figure 1: Systematic deviation α

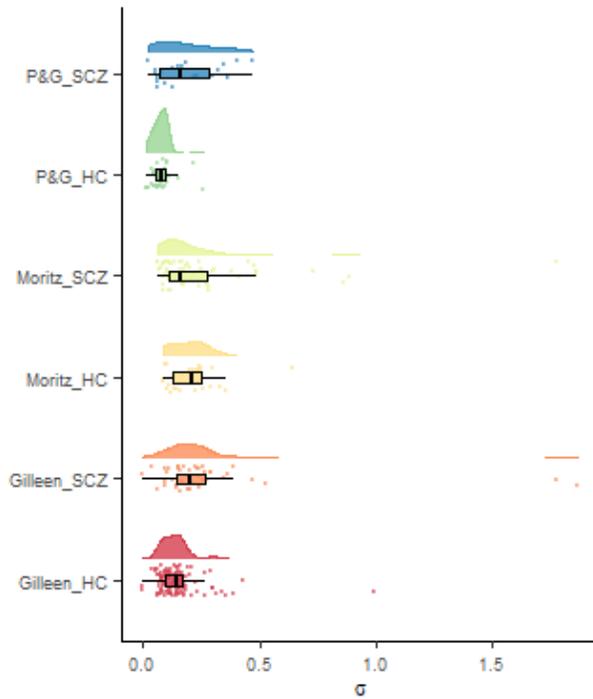


Figure 2: Random / stochastic deviation σ

The model

The probability that bead k comes from jar A in sequence j can be calculated using Bayes formula. We model here the deviations from an ideal Bayesian observer (IBO). We assume that a participant i may or may not use all available information, and model this with a parameter m , describing the number of beads included in one's probability estimates, i.e. the memory length. m ranges from 0 to 10, i.e. if $m_i = 0$ the participant totally ignores the sampled beads, if $m_i = 1$ one only considers the last bead sampled and so on, and if $k \leq m_i$ one considers all the k beads sampled so far.

We further include two parameters describing systematic and random deviations between the subjective probability estimates and those from the IBO. We define this probability as ρ_{ijk}^L for the left bag, and correspondingly the probability for the other bag is $1 - \rho$. Note, that in the Moritz dataset participants did indicate both probabilities, but to compare the 3 datasets we used this specification for all three datasets. Since the probability estimates range from 0 to 1, we use a truncated normal distribution which we associate with a normally distributed variable $\pi_{ijk}^L \in (-\infty, \infty)$ and define

$$\rho_{ijk}^L = \begin{cases} 0 & \text{if } \pi_{ijk}^L \leq 0, \\ \pi_{ijk}^L & \text{if } \pi_{ijk}^L \in (0, 1), \\ 1 & \text{if } \pi_{ijk}^L \geq 1. \end{cases} \quad (1)$$

The mean and standard deviation of π_{ijk}^L we define as

$$E[\pi_{ijk}^L | \theta_i] = \frac{1}{1 + \left(\frac{1 - p_{ijk}^L(m_i)}{p_{ijk}^L(m_i)} \right)^{a_i}} \quad \text{and} \quad \text{SD}[\pi_{ijk}^L | \theta_i] = \sigma_i, \quad (2)$$

respectively.

For $a_i = 1$ we have $E[\pi_{ijk}^L | \theta_i] = p_{ijk}^L(m_i)$. Parameter values $a_i < 1$ and $a_i > 1$ model tendencies for individual number i to specify values for ρ_{ijk}^L that are too close to a half and too close to zero or one, respectively, relative to the probability $p_{ijk}^L(m_i)$ used to decide from which jar the beads are coming. The standard deviation σ_i models to what degree the specified probability estimates ρ_{ijk}^L of individual number i deviate from $E[\pi_{ijk}^L | \theta_i]$. Note that a can take on any value, and a describes how well one can discriminate probabilities.

We adopt a hierarchical Bayesian setup and assign a prior distribution for the parameter vector of each individual, $\theta_i = (m_i, a_i, \sigma_i)$. Given a vector of hyper-parameters, φ , we assume the parameter vectors $\theta_i, i = 1, \dots, n$ to be a priori independent. Moreover, given φ we assume the components of θ_i to be independent. We assume all a_i and σ_i to be gamma distributed and parameterise these gamma distribution by their mean values and standard deviations. For m_i we just assume it has some discrete distribution over the possible values. We assign a vague prior to the vector of hyper-parameters φ . Further, we estimate the parameters for each group (patients and controls) and datasets separately.

Results

Participants with a diagnosis of schizophrenia use on average fewer beads, i.e. their m is lower than that for healthy control participants. The systematic deviation a is smaller in patients than in healthy controls (Fig 1). In P&G and Gilleen’s datasets both groups have $a < 1$ but the patient group has a smaller systematic deviation than the control group (see table 1). In the Moritz dataset, on average, the patient group shows nearly no systematic deviation, compared to the more sigmoid behaviour of the control group. However, closer inspection shows that this is due to the asymmetric sequence (see below). The unsystematic deviation was larger in patients, across sequences and consistent for all three datasets (Fig 2).

Table 1: Descriptive Statistics

	a		σ		m	
	HC	SCZ	HC	SCZ	HC	SCZ
P&G mean	0.32	0.45	0.08	0.19	8.45	5.32
P&G SD	0.35	0.4	0.05	0.13	2.65	2.08
Moritz mean	1.78	0.96	0.20	0.26	9.60	8.8
Moritz SD	0.93	0.66	0.11	0.28	1.78	2.25
Gilleen mean	0.44	0.52	0.15	0.30	7.94	4.78
Gilleen SD	0.31	0.36	0.11	0.37	3.59	4.12

In all three datasets σ was negatively correlated with m , i.e. the fewer beads considered the more random deviations the participant also showed, however the overall correlation was only $r = -.14$ with an 95% CI $[-.245 \text{ } -.032]$. Similarly, there was an overall positive correlation between systematic and random deviation, $r = .167$, 95% CI $[.059 \text{ } .271]$ however, the correlation was negative in the P&G dataset and close to zero in the Moritz dataset, i.e. this relationship was driven by the large number of controls in the Gilleen dataset.

Figure 3 illustrates the interaction of the memory parameter and stochastic behaviour for two cases with nearly perfectly linear probability estimates. In the left-hand of Fig 3 changes in observed beads have a high correlation with the specified probabilities; a low m is reasonable, as only the last bead is considered. When the low m explains most of the changes in the specified probabilities σ becomes low. In contrast, on the right-hand of Fig 3 the participant seems more or less to ignore the observed beads, then σ must become large.

Which sequence is most discriminative

Next, we looked at how the type of sequence affects performance. In the Gilleen dataset, we found that the systematic deviation was smallest for the sequences A and D than the sequences B and C. Also the unsystematic deviation σ was lowest for the first sequence participants saw. Healthy controls used the same number of beads in all four sequences whereas patients used around 2 beads more in the sequence B, indicating that they are aware that estimating from which jar the bead comes is harder to judge in a (de facto) 60:40 sequence.

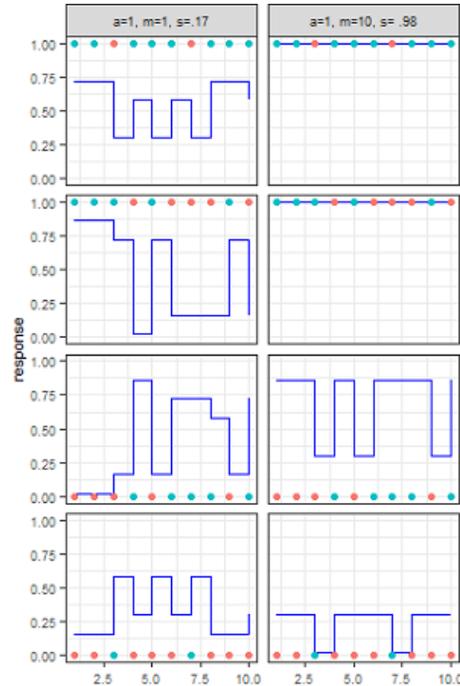


Figure 3: Responses from two participants with different σ (s) and m but same a , dots represent the beads shown. Top row is 80:20, then 60:40, 40:60 and last row is the 20:80 trial.

Moritz et al. used a symmetrical and an asymmetrical sequence. The asymmetrical sequence yielded in both groups a more sigmoid estimate of the probabilities, most extreme in healthy controls ($m = 1.91$, $sd = .89$) compared to patients ($m = 1.28$, $sd = .64$). This is in contrast to the symmetric sequence where the systematic deviation was very good for controls ($m = .81$, $sd = .29$), whereas patients showed a systematic deviation similar to that seen in P&G and Gilleen dataset ($m = .55$, $sd = .26$). The random deviation σ was similar in the symmetric condition but increased for both groups in the asymmetric sequence. Notably controls did use on average one more bead to make their probability estimates in the asymmetric condition.

Combining over all three datasets we found that the difficulty of a sequence mostly increases the unsystematic deviation in patients whereas such an increase in σ was not so profound for healthy controls.

Discussion

Our model captures two important aspects of the beads tasks: understanding probabilities and sequential updating of the information. We find clear differences between healthy controls and patients on all three parameters, and importantly also between the datasets. The number of beads considered is lower in patients than in healthy controls, but both groups evaluate more beads in the asymmetric sequence and in sequences with more even ratios, indicating that both groups are sensi-

tive to the cognitive effort required to estimate the probabilities. Reduced cognitive abilities have been linked to the jumping to conclusions bias in the beads task (Garety et al., 2013; Speechley, Whitman, & Woodward, 2010; Woodward, Munz, LeClerc, & Lecomte, 2009). Thus, our model captures this explicitly by modelling the number of beads used for making an inference. Importantly, fewer beads considered does not lead to a more biased probability estimate or always to more unsystematic responding (see Fig 3). We find a smaller systematic deviation in patients than in controls, indicating good discriminability of probabilities, in fact patients treat probabilities more linearly than healthy controls do. The larger systematic deviation in healthy controls might reflect the conservative bias (Phillips & Edwards, 1966), i.e. updating of the probability estimate is less than prescribed by Bayes theorem. Unsystematic or random deviation was larger in patients than in controls, but patients were also more heterogeneous than controls. The larger unsystematic deviation found here is in agreement with the results from Moutoussis, Bentall, El-Derey, and Dayan (2011) who modelled the Draws to Decision version of the beads task. The more stochastic responding may reflect a propensity to see a change in jar, particularly in those cases where the number of beads considered is large (Pfuhl, Sandvik, Biegler, & Tjelmeland, 2015). The model by Adams et al. (2018) and other paradigms support the notion that patients perceive the world as more volatile, i.e. attributing new evidence to a change in the environment (Deserno et al., 2017). In the Moritz dataset patients show on average very good understanding of probabilities but this was driven by the asymmetric sequence where controls show - contrary to the symmetric sequence - a sigmoid response, i.e. they prefer to dichotomize the probabilities into very sure / absolute sure it comes from jar A or it comes from jar B. In contrast, P&G's and Gilleen's datasets show that both patients and controls have a large range to which they respond with around 50% or could be any of the two jars. This might be driven by how the task was presented. Indeed, Moritz's dataset showed very good memory, i.e. m around 10 for most participants including patients. P&G and Gilleen, on the other hand, indicate that patients use fewer beads when judging probabilities. In Moritz we find a difference in subjective probability, not memory, whereas in P&G and Gilleen the main difference is one of memory, not of subjective probability estimates. Still, in all three datasets, patients on average respond more randomly (Fig 2). We also found that the first presentation is most discriminative between patients and controls. Patients become better on the second sequence, although that could be confounded by the difficulty of the sequences. Also, using asymmetric sequences reduces group differences. In sum, our model may resolve why some find hypersalience in patients whereas others find a lower decision threshold. Participants use different strategies, some respond bead by bead, seen in a low m whereas others are nearly perfectly Bayesian which can appear in comparison to the healthy control sample as having a lower decision threshold. Probabilities are not

treated linearly but, importantly, patients have a lower systematic bias in estimating probabilities than healthy controls.

Acknowledgments

This work is funded by a Norwegian Research council grant (FRIMEDBIO 262338)

References

- Adams, R. A., Napier, G., Roiser, J. P., Mathys, C., & Gilleen, J. (2018). Attractor-like dynamics in belief updating in schizophrenia. *Journal of Neuroscience*, *38*(44), 9471-9485.
- Deserno, L., Boehme, R., Mathys, C., Katthagen, T., Kaminiski, J., Stephan, K. E., ... Schlagenhaut, F. (2017). Overestimating environmental volatility increases switching behavior and is linked to activation of dorsolateral prefrontal cortex in schizophrenia. *bioRxiv*. doi: 10.1101/227967
- Garety, P., Joyce, E., Jolley, S., Emsley, R., Waller, H., Kuipers, E., ... Freeman, D. (2013). Neuropsychological functioning and jumping to conclusions in delusions. *Schizophrenia research*, *150*(2-3), 570-574. doi: 10.1016/j.schres.2013.08.035
- Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *The Quarterly J Exp Psy Section A*, *40*(4), 801-812. doi: 10.1080/14640748808402300
- Moritz, & Woodward. (2005). Jumping to conclusions in delusional and non-delusional schizophrenic patients. *British Journal of Clinical Psychology*, *44*, 193-207.
- Moritz, S., Scheu, F., Andreou, C., Pfueller, U., Weisbrod, M., & Roesch-Ely, D. (2016). Reasoning in psychosis: risky but not necessarily hasty. *Cognitive Neuropsychiatry*, *21*(2), 91-106. doi: 10.1080/13546805.2015.1136611
- Moutoussis, M., Bentall, R. P., El-Derey, W., & Dayan, P. (2011). Bayesian modelling of jumping-to-conclusions bias in delusional patients. *Cogn Neuropsychiatry*, *16*(5), 422-47. doi: 10.1080/13546805.2010.548678
- Peters, E., & Garety, P. (2006). Cognitive functioning in delusions: a longitudinal analysis. *Behaviour research and therapy*, *44*(4), 481-514.
- Pfuhl, G., Sandvik, K., Biegler, R., & Tjelmeland, H. (2015). Identifying the computational parameters gone awry in psychosis. In *Brain informatics and health* (p. 23-32).
- Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *J Exp Psychol*, *72*(3), 346-54.
- Speechley, W. J., Whitman, J. C., & Woodward, T. S. (2010). The contribution of hypersalience to the "jumping to conclusions" bias associated with delusions in schizophrenia. *Journal of Psychiatry and Neuroscience*, *35*(1), 7-17. doi: 10.1503/jpn.090025
- Woodward, Munz, LeClerc, & Lecomte. (2009). Change in delusions is associated with change in "jumping to conclusions". *Psychiatry Research*, *170*, 124-127.