Less Precise Hidden States in Perceptual Decision-Making show Lower Confidence

Bruno Carlos Pless

Vision and Control of Action Group, Cognition and Development Department, Universitat de Barcelona

Supervisor: Daniel Linares

22nd June, 2023

Master's Degree in Neuroscience

LINARES HERREROS DANIEL - 53066206P

Digitally signed by LINARES HERREROS DANIEL - 53066206P
Date: 2023.06.22 13:15:54 +02'00'

Abstract

Perceptual decision-making involves using sensory information to guide behaviour in various real-world scenarios, and while there is a significant body of research on the topic, the full understanding of its computational and neural basis is not completely understood.

A recent study by Ashwood et al. (2022) suggests that behaviour is best represented by transitions between different discrete internal states, as opposed to a single state, and this can be analysed using a combination of Generalized Linear Models and Hidden Markov Models known as GLMHMM.

The present study aimed at assessing this novel method of modelling with a data set from Rouault et al. (2018) that consisted of non-trained participants. The findings suggest that a two states model can better account for perceptual decision-making processes in a large segment of the population (200 out of 498 participants), although the better fit of the two states model over a single state model was found to be dependent on the number of trials, indicating the need for further research with a greater number of trials per participant.

There is a vast literature about the relation between confidence and perceptual decision making. However, it had not been assessed from a multiple internal states perspective. The present work revealed that states with lower precision are associated with lower confidence, suggesting that participants may be metacognitively aware of their internal state.

This work underscores the potential of the GLMHMM as a powerful tool in the study of perceptual decision-making and paves the way for understanding the cognitive and neural basis implicated in this complex phenomenon.

Identification and Reflection about the Sustainable Development Goals

The present work aimed to contribute to the United Nations' Sustainable Development Goal of Good Health and Well-being. It promoted health equity by supporting a novel method to assess perceptual decision-making, which is a critical function necessary for survival. Perceptual decision-making has been observed to be altered in various health conditions and disorders. Gaining a better understanding of these alterations could potentially lead to the development of innovative therapeutic treatments and assessing techniques.

The development of this new approach to analysing perceptual decision-making holds substantial potential for further research, especially in the fields of neuropsychology and cognitive neuroscience. Particularly, it may allow more precise diagnostic techniques for numerous pathological conditions.

Furthermore, our methodology has implications for the Sustainable Development Goal of Climate Action. By proposing a method that may eliminate the need for expensive and resource consuming equipment traditionally used in assessing perceptual decision-making, we contribute to more sustainable research. This aligns with global efforts to reduce material consumption and energy usage in scientific research.

In addition, our methodology has the potential to make diagnosis accessible for more people. In comparison with more expensive diagnostic tools such as neuropsychological tests, if proven to be useful, this novel method has the potential to become an inexpensive way of assessing symptoms. This could facilitate the access to evaluations for a broader population, irrespective of their socio-economic status or geographical location. Which can help us move closer to achieving health equity, which is a key component of the Good Health and Well-being goal. Ultimately, the development of this new method embodies the spirit of the Sustainable Development Goals, highlighting the mutual dependence of health, innovation, and sustainability.

List of Abbreviations

AIC: Akaike Information Criterion.

GLM: Generalized Linear Model.

GLMHMM: Generalized Linear Model with Hidden Markov Model.

Index

1.- Introduction and Objectives

Perceptual decision-making involves using sensory information to guide behaviour towards the external world such as driving, shopping, diagnosing illnesses, and playing sports. It encompasses processes such as gathering sensory input, evaluating, and integrating it based on the person's objectives and internal state, and generating motor responses. In neuroscience, one of the main aims is to uncover the computational mechanisms through which neural circuits encode, store, and analyse sensory signals, integrate them with other behaviourally relevant information, and resolve conflicts between competing motor plans (Hauser & Salinas, 2014). Despite the large literature about it and its basis (Heekeren et al., 2008; Kelly & O'Connell, 2015), perceptual decision-making is far from being completely understood.

To study perceptual decision-making in the lab while being able to greatly control the stimuli, there is a wide range of perceptual decision-making tasks such as discrimination between different stimuli, detecting whether a stimulus is present or not depending on its intensity, or categorizing different stimuli based on common features. As for the stimuli, several dimensions have been studied, such as contrast, size, and motion. As an example, let's consider a contrast detection task. A stimulus is presented on the right or on the left of the screen, and the task of the subject is to identify the location. The difficulty of the task is manipulated by presenting stimuli of different levels of contrast on each trial.

Data from perceptual decision-making tasks are often modelled using a psychometric function in which the x axis represents stimulus magnitude, while the y axis represents the probability of making one type of response. For example, for the contrast detection task described above, the psychometric function usually models the probability of choosing the right location as a function of the signed contrast. The signed contrast is the contrast of stimuli multiplied by -1 when the stimulus is on the left and by 1 when the stimulus is on the right.

Formally, the psychometric function is an instance of a Generalized Linear Model (GLM); it is composed by an intercept and a slope, which are the point at which the curve touches the y axis and the rate of change of the dependent variable respectively.

Modelling the responses of a subject in a perceptual decision-making task using a psychometric function assumes that only one single perceptual decision-making strategy (determined by an internal state) is used, but this view has been challenged. A recent study that has received great attention (Ashwood et al., 2022) found that behaviour is better described by transitions between different discrete internal states than by a single state. This paper proposed to use Hidden Markov Models in combination with the traditional Generalized Linear Model (together GLMHMM) to analyse it. This consists in building a model that generates different psychometric functions that represent different hidden internal states (they are hidden because they are only inferred by the data). This model also provides the probability to be in each state, on each trial, and to stay or transit between states (e.g., in a model with two hidden states this would be the probability for staying in state 1 and transitioning from state 1 to state 2).

In the aforementioned paper, GLMHMMs based on different quantities of hidden states were built and assessed using three data sets: two with mice and one with humans. It was found that in all cases GLMHMMs with more than one state proved to explain the data better than the single state model.

The first mice data set in Ashwood et al. (2022), which was obtained from Aguillon-Rodriguez et al. (2021), was the result of a contrast detection task in which each mouse had to choose a stimulus that appeared either on the left or the right side of a screen. The independent variable was the signed contrast of the stimulus, which ranged from -100% to 100%, negative numbers reflecting that the stimulus was on the left. The dependent variable was the probability to choose the right side. It was found that the three states model was the best at explaining the data without risking overfitting. These three states were: (i) an unbiased state, in which the sensory evidence was being taken into account and no significant bias played a role; (ii) a bias-left state in which there was a bigger tendency to respond left; and (iii) a bias-right, which was similar to the bias-left but on the right side. They also found that it was common to remain in each state for many trials in a row. For example, the animal tended to be in a right bias state for a time, then left bias, and then unbiased.

As for the other mice database, which was retrieved from Odoemene et al. (2018), a four state model composed by an engaged, bias-right, bias-left, and a win-stay state (in which the mouse would chose the same response as the last one if it was rewarded) showed to be marginally better than the three states model while also taking interpretability and simplicity into account. The fact that models with different numbers of states proved to be the best for each mice data set might result from a difference in the experimental protocols used in each experiment.

A human data was also assessed in this study. The data set was obtained from Urai et al. (2017) and was composed by 27 trained humans that performed a motion discrimination task, which consisted in being presented two different sets of moving dots and choosing the set in which the most dots were moving in the same direction (motion coherence). In this case, the model composed by two states (one biased toward perceiving more and the other toward perceiving less movement) was the one that best explained the observations. It was also observed that precision varied between states, suggesting different levels of engagement.

Even though Ashwood et al. (2022) demonstrated that the modelling with GLMHMMs was better at explaining three different data sets than the traditional modelling, it is unknown whether this result in humans is specific to this experiment, or it can be obtained with other data sets involving other tasks and with non-expert participants.

As we have described, different states are often associated with different levels of precision. It is unknown, however, whether participants are aware of their precision in each state (e.g., if they are engaged or disengaged).

Confidence in perceptual decisions is a complex subject to which there are still many debates taking place. One of them is about the relation between accuracy, confidence, and response times, for it seems that confidence is associated up to a certain point with both (Mamassian, 2016).

When it comes to GLMHMMs in perceptual decision making, the degree of metacognition on internal states can be measured by the degree of confidence that participants have in each trial's response, that is, how certain they are about having chosen the correct response. From this, we can then calculate each participant's overall confidence level as well as the confidence associated with each degree of the stimulus intensity. Which was not done neither in Urai et al. (2017) nor in Ashwood et al. (2022).

Shifts in the overall confidence level have been observed to be associated with self-reported psychiatric symptoms independently of the performance level (Rouault et al., 2018), underscoring the importance of studying confidence related to perceptual decision making.

Taking all the above into account, the objectives in the present work were: (i) to analyse whether the human results seen in Ashwood et al. (2022) can also be found using another perceptual task and with non-trained participants and (ii) to evaluate the level of confidence that each participant has about each of their answers in relation to each state, as a way of measuring how aware they are of their own internal state.

To give an answer to these questions, we decided to use the database from Rouault et al. (2018), which is a large data set that can be downloaded from the Confidence Database, which is itself an enormous database of confidence studies comprising a wide set of paradigms, participant populations, and fields of study (Rahnev et al., 2020).

The completion of the present work's objectives may help to better understand the mechanisms behind perceptual decision-making and support this novel way of modelling such mechanisms.

2.- Materials and Methods

In the present work we used the human data set from experiment 1 of Rouault et al. (2018) which consisted of 498 volunteers (mean age= 35.71; SD= 11.37; 237 women). Participants performed 210 trials divided in 5 blocks. In each trial (Figure 1), participants saw a fixation cross for one second, and then two black squares (left and right) containing different quantities of white dots which were set in random positions for 300 ms. One of the squares had half of all possible positions inside it filled with dots (313 out of 625), while the other had from 1 to 70 more dots (mean dot difference= 35.5; SD= 20.20). Then, the dots disappeared, and the black squares remained on the screen until the participants pressed a button on the keyboard. Then, participants had to choose the position (right/left) of the square that contained the highest number of dots; the position of this was pseudo-randomised across all trials within five difficulty bins. The selected square was highlighted for 500 ms; participants received no feedback.

Lastly, participants had to rate their confidence in their response from 1 to 11. Participants had no time limit for neither their response nor their confidence rating.

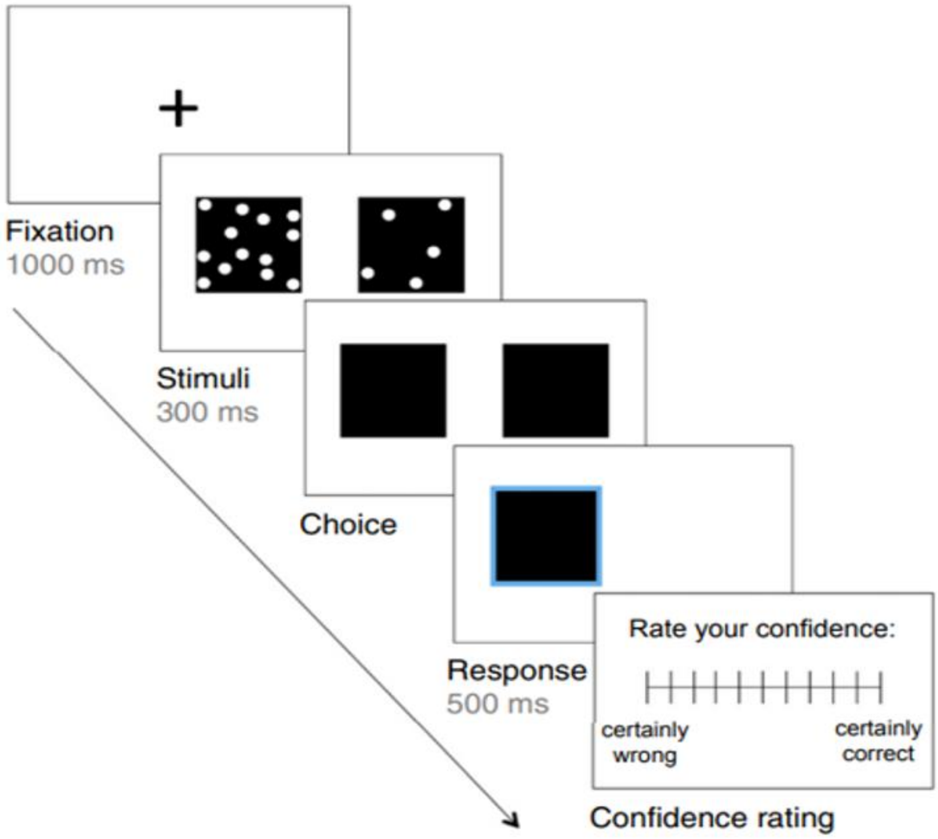

Figure 1. Illustration of the perceptual decision-making task in Rouault et al. (2018) and used in the present work. This image was obtained from “Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but NotTask Performance” by M. Rouault et al., 2018, *Biological Psychiatry*, *84*(6), 443–451.

To fit a GLMHMM model to each participant, we used the ssm library from Linderman (2020), which is a free Python software. The adjusting method consists in first obtaining the state transition matrix, which is the matrix of probabilities of staying in the current state of changing to another, calculating the set of slopes that influences choice in each state, and the initial state distribution, which is the state that is the most probable that the participant started in; each of these sets of initial parameters is known as Θ and is generated randomly with the use of seeds. We used 20 seeds per model in the present work.

Finally, the posterior probability of each set is calculated, that is, the probability of each set of parameters Θ given the observed data X (real data) is obtained; from it, the maximized log-posterior is obtained using the Expectation Maximization (EM) algorithm and the highest one is chosen.

The EM consists in generating the expected value of the log-likelihood function, with respect to the conditional distribution of the latent variables given the observed data under the current estimate of the parameters; this expected value is known as the Expected Complete data Log-Likelihood (ECLL). Then, it maximizes each ECLL with respect to the model parameters, Θ. The log-posterior is used instead of the posterior probability itself because it turns the products in the calculation into sums, which are computationally easier to handle.

3.- Results

Figure 2 shows the results for an example participant. The dots in Figure 2A (left) show the proportion of times that the participant responded that the square in the right had more dots than the square of the left as a function of the difference in dots between the two sides (negative numbers indicate more dots on the left and positive numbers more dots on the right). The proportion of responses was fitted with a psychometric function with only two parameters (intercept and slope). This is, indeed, a single state GLMHMM.

The degree of inclination of the curve (slope) represents the precision. The steeper the slope is, the more precise were the participant's decisions. The accuracy (bias) can be assessed by the point along the *x* axis in which the curve crosses the 0.50 probability of the *y* axis. For example, a curve shifted towards the right side indicates a left bias, which means that, while in that state, the individual had a preference for choosing left even if there was more than enough evidence that the rightward stimulus had more dots. A curve without bias or with perfect accuracy would be one in which the point in the *x* axis that corresponds to the 0.5 probability of the *y* axis is equal to 0.

In the particular case of the results of participant five, we can infer from the one state model's psychometric curve that the participant has a slight left bias, that is because it can be observed that the point along the *x* axis corresponding to 0.50 on the *y* axis is slightly shifted towards the right.

When it comes to the graph on Figure 2A right, the *x* axis represents the sequence of trials, while the *y* axis is the probability of being in each state. The black lines at the top and bottom of each figure represent choosing right or left correspondingly for each trial, while the coloured line represents the probability of being in state 1 (which was always 1 out of 1 because there is a single state for that model).

For the two states model (Figure 2B), it appears that state 1 has practically no bias, while state 2 is quite shifted towards the right end of *x* axis, which denotes a strong left bias, the curve also has a less steep slope than state 1, which indicates worse precision, which could mean that the participant was more disengaged in the task while in this state. In Figure 2B right we can observe that, for the two states model, the participant began the task in the second state, but at around the middle of the task he or she began shifting into state 1.

In the three states model, by the first graph we can deduce that the participant had a nearly perfect precision and very small right bias in the first state, a worse precision and a considerable left bias in state 2, and an even worse precision and bigger left bias in state 3. In the right graph it can be appreciated that the participant started the task while being in state 3 and around at the middle of the task he or she was never again in this state and instead switched between states 1 and 2 constantly.

Lastly, for the four states model, the precision and bias of state 1 was practically the same as in the three states model, state 2 was similar but with a strong left bias, state 3 practically corresponded to state 2 of the three states model, and state 4 to state 3 also of the three states model. The transition between states was again very similar to the three states model, just that in this case state 4 was the predominant one in the first half of the trials, while the other three intercalated in the second half.

The reader may notice that states are ordered in a descendent manner based on their slope; this was done to have states with similar precision arranged in the same way across individuals.

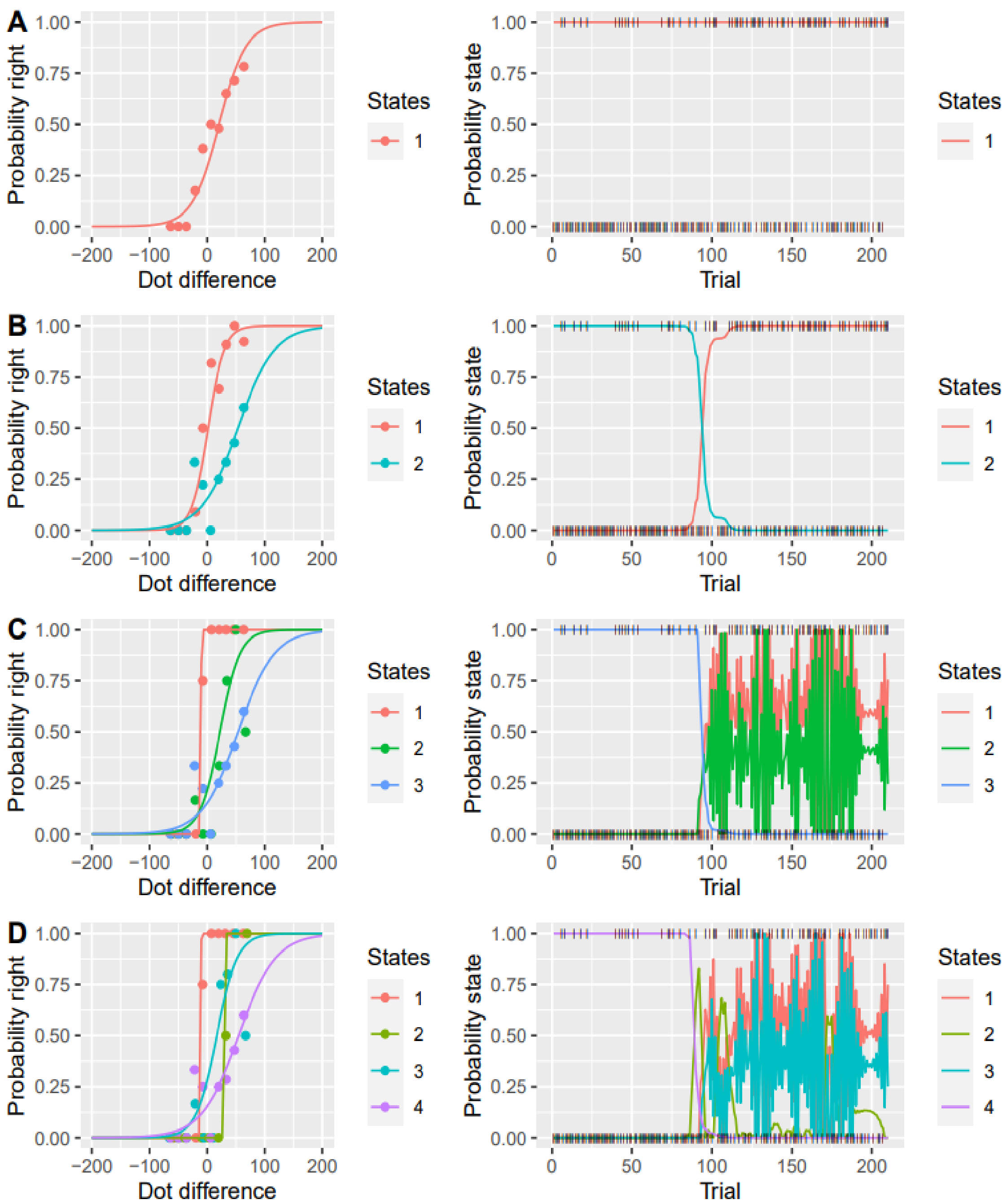

Figure 2. Graphical representation of participant number five models' psychometric curve and probability to be in each state for each trial for the single state model **(A)**, two states model **(B)**, three states model **(C)**, and four states model **(D)**.

Each model has a number of parameters determined by the following formula:

$$\boldsymbol{k = 2n + n^n - n} \tag{1}$$

Where:

$\boldsymbol{k}$ is the number of parameters.

$\boldsymbol{n}$ is the number of hidden states.

In this formula the first term ($2\boldsymbol{n}$) indicates the number of parameters of the fitted GLMs; it is multiplied by two because each state has two parameters (intercept and slope) The rest of the formula corresponds to the parameters of the transition matrix. The transition matrix has $\boldsymbol{n}$ times $\boldsymbol{n}$ cells, but to obtain the number of free parameters $\boldsymbol{n}$ needs to be subtracted. This is because each row should add up to 1 given that being a given state the probability to transition to the rest of states should add to 1. As an example, the transition matrix of participant 5 for the 3 states model can be seen at Table 1.

Table 1. Transition matrix of participant 5 for the 3 states model.

	State 1	State 2	State 3
State 1	0.297	0.702	0.172 × 10-33
State 2	0.999	7.80506 × 10-6	0.245 × 10-33
State 3	0.625 × 10-6	0.010	0.989

To assess which model describes better the data in this and all participants, we used the Akaike Information Criterion, which follows this formula:

$$AIC = 2k - 2ln(L) \tag{2}$$

Where:

k is the number of parameters in the statistical model.

L is the maximized value of the likelihood function for the estimated model.

As one can appreciate in the formula, this criterion takes into account both the logarithmic likelihood (how well each curve fits the data) and the number of parameters of each model to avoid overfitting. For this particular subject, the best model, that is, the model with the lowest AIC score, was the one with three states (AIC = 167.51).

We fitted the data as we described for participant 5, for all the participants in the study. Different participants had different biases and precision but given our first objective we will focus on assessing which model is the best at describing the data using the AIC.

We found that the one state model was the best at explaining the data for the biggest quantity of subjects (262), while the two states model did for 200 persons, the three states model for 34, and the model with four states for only 2 (Figure 3).

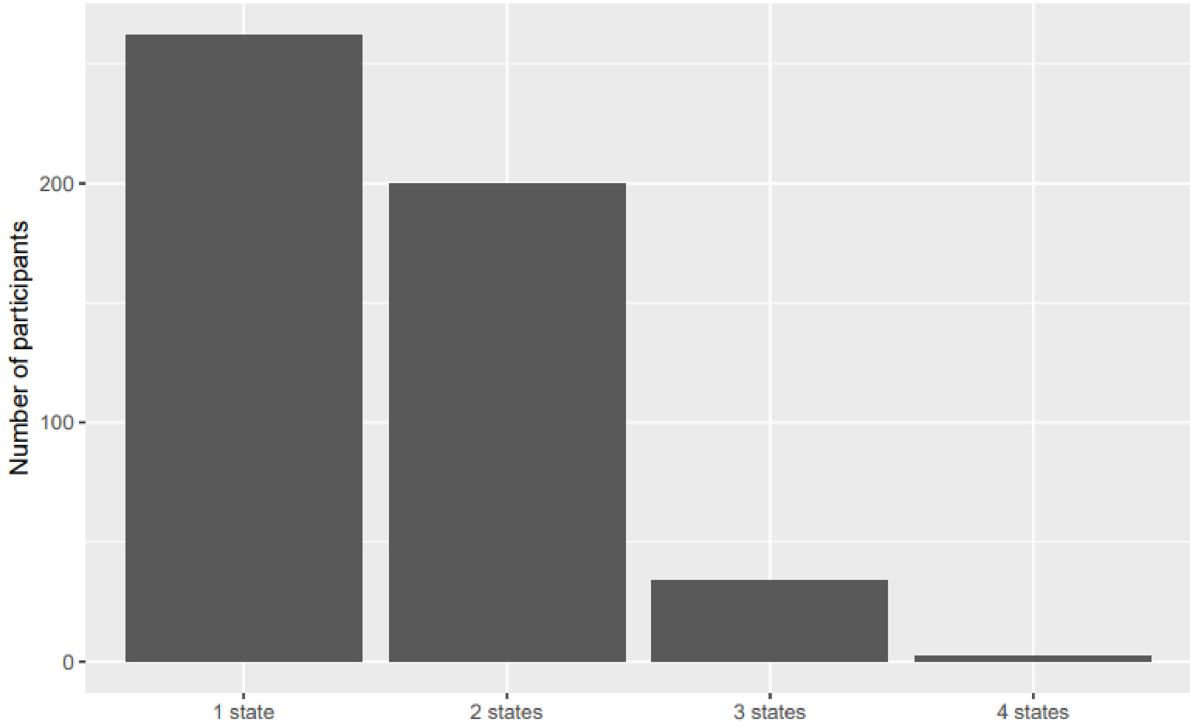

Figure 3. Number of participants for which each model fitted the best.

The fact that for so many participants the two states model was the best is consistent with the findings of Ashwood et al. (2022); however, what is apparently inconsistent is the fact that the single state model still best accounted for the most people. We thought that a possible explanation for this finding was the difference between the number of trials in the task used in our data set and the one used for the task in Ashwood et al. (2002), which were 210 and 500 trials respectively. To explore this idea, we decided to run the same analysis but with subsamples of the 52 first trials, the 104 first trials, and the156 first trials of each participant to see if the models with more states were favoured as the numbers of trials increased. The results can be seen in table 2 and Figure 4.

Table 2. Number of participants for which each model fitted the best for different numbers of initial trials.

	52 trials	104 trials	156 trials	210 trials
One state model	379	356	308	262
Two states model	104	118	161	200
Three states model	14	24	27	34
Four states model	1	0	2	2

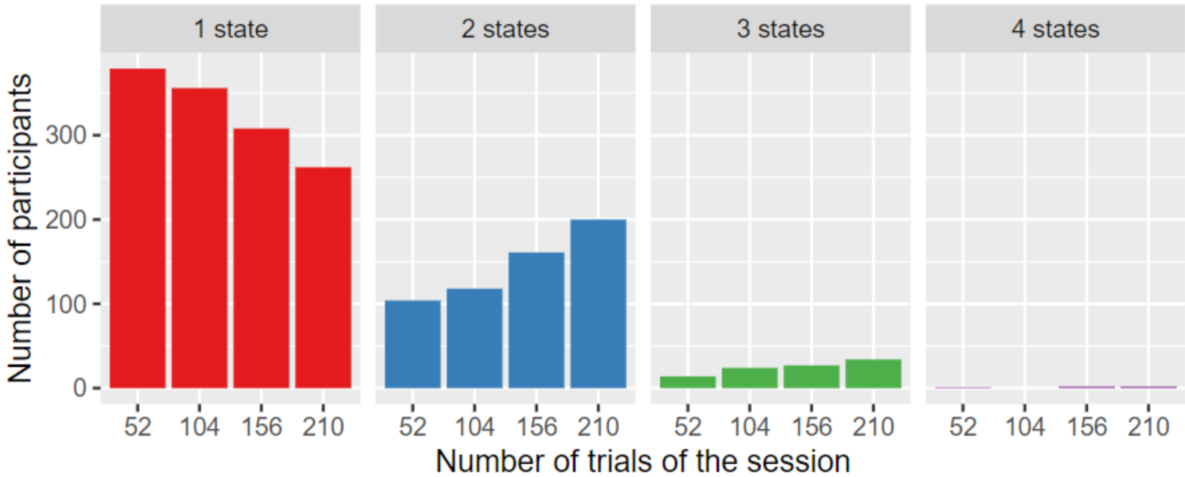

Figure 4. *Number of participants for which each model fitted the best for different numbers of initial trials.*

It can be seen that the number of participants for which the 3 and, especially, the 2 states model fit the best increases the more trials were included in the analysis, which suggests that less trials favour models with less states. This result indicates that if the data in Rouault et al. (2018) had more trials, the results will not be so different from Ashwood et al. (2022).

Therefore, we think that our results are consistent with Ashwood et al. (2022): 2 and 3 states are necessary to fit well the data of many participants in a perceptual decision-making task.

To fulfil our second objective, for each model of every participant, we assigned the state that had the highest likelihood to each trial and, so, obtained the mean confidence for each state of all models in every participant; then we calculated the overall mean confidence and its confidence intervals of each state of each model taking all participants into account.

We then compared the level of confidence associated with each state inside each model. We found statistically significant differences between state 1 and state 2 of the two states model ($t = 6.65$, $p = 7.38\times10\text{-}11$).

For the three and four states models we used repeated measures ANOVA. We obtained significant results for both the three and the four states models ($p = 2\times10\text{-}17$, $p = 2\times10\text{-}16$ respectively). Tukey comparisons between the individual states of these models can be seen in Table 2.

Table 3. Results of Tukey comparisons between states in the three and four states models with states ordered by precision

Model	Comparison	*t*	*p*
Three states model	S1 - S2	0.170	0.984
	S1- S3	10.633	<10-22*
	S2- S3	10.464	<10-22*
Four states model	S1 - S2	0.937	0.784
	S1 -S3	3.837	7×10-4*
	S1 - S4	11.578	<10-22*
	S2 - S3	2.9	0.019*
	S2 - S4	10.640	<10-22*
	S3 - S4	7.739	<10-22*

The graphic representations can be appreciated in Figure 5 where we can see the comparison of confidence distributions between states of each model represented by dots that symbolize the confidence ratings, while the highlighted line represents the mean and the coloured lines represent the confidence intervals. For example, we can clearly see the difference between states 2 and 3 of the 3 states model by noticing how the confidence intervals do not coincide at all on the *y* axis.

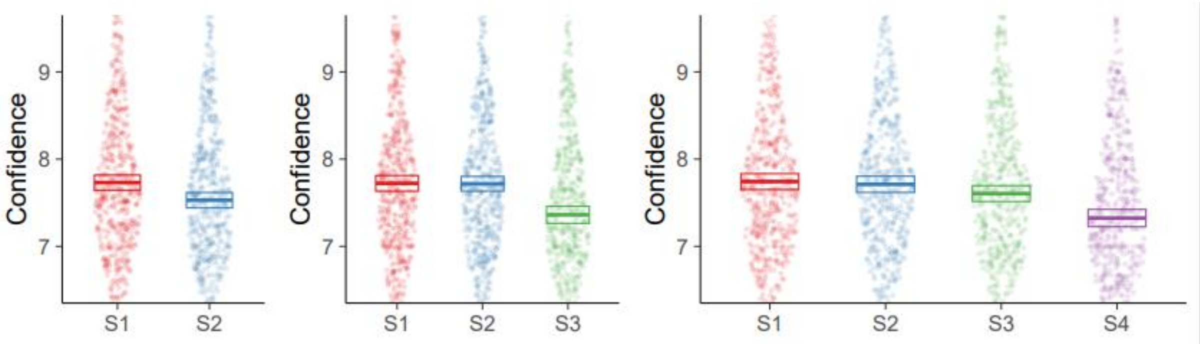

Figure 5. Differences in confidence between states ordered by precision in **A)** the two states model. **(B)** the three states model, and **(C)** the four states model. Coloured lines represent confidence intervals. Some points have been left outside of the graph to better view the confidence intervals.

Overall, these results indicate that states with less precision are related to lower confidence in one's own perceptual decisions.

4.- Discussion

In addressing our first objective, which was to determine whether the human results observed in Ashwood et al. (2022) could also be found using a different perceptual decision-making task and with non-trained participants, we discovered that the two states model fit best for a significant portion of all participants. This supports the findings of Ashwood et al. (2022). Even though we found that the single state model still accounted for the majority of participants' data, with the subsampling of initial trials we were able to observe an increase in the number of states used as the participant completed more trials.

Based on these results, despite having a large sample of participants, we identify the number of trials per participant as an unforeseen limitation. We believe further analysis involving more subjects than in Ashwood et al. (2022) and more trials per participant than in the current study will be necessary to assess the trend of increasing the number of states used as a participant completes more trials. Exploring the neural and cognitive basis of this trend will also be beneficial to determine if it occurs due to phenomena such as numerosity adaptation (Burr and Ross, 2008) or attentional fatigue (Boksem et al., 2005). Furthermore, we speculate that requiring participants to assess their own confidence for each trial's decision might encourage them to stay focused for longer, since they must evaluate their level of attention for each trial. We think that promoting attention to one's tasks may lead to a higher level of engagement and, subsequently, higher precision.

Ashwood et al. (2022) noted in their discussion that the ability to infer the condition or strategy used by an animal on a trial-by-trial basis could aid in distinguishing performance variations across different sessions and animals. They suggested that this could be a powerful tool for neuroscientists examining the neural mechanisms underlying decision-making. Finally, they proposed that varying strategies may rely on distinct neural circuits or patterns of neural activity.

A month after the publication of Ashwood et al. (2022), Bolkan et al. (2022) published a study based on experiments with mice in which they optically inhibited each Dorsomedial Striatum (DMS) pathway prior to a T-maze task. The mice had to accumulate sensory evidence before making a decision. Once again, they found the 3 states model to be the best fit. They found that the contribution of the DMS pathways was a consequence of the perceptual decision made. Unilateral inhibition of the indirect DMS pathway, which curbs involuntary movements, resulted in a bias towards contralateral choices, whereas inhibition of the direct DMS pathway led to an ipsilateral bias.

In this paper, significant variations across states in the weighting of sensory evidence, previous trial response, and notably, the inhibition of DMS pathways were also discovered. Specifically, behaviour in both states 1 and 2 was heavily influenced by sensory evidence, resulting in high performance, and was minimally influenced by choice history. However, the effects of DMS pathway inhibition were much more pronounced in state 2. As for state 3, considered a state where mice were disengaged with the task, behaviour was less influenced by sensory evidence and DMS pathway inhibition, and more affected by choice history.

While this experiment was conducted with mice, its conclusions, combined with our results and those in Ashwood et al. (2022), suggest that GLMHMM can be an extremely useful tool in neuroscience, particularly for studying neural and cognitive mechanisms in perceptual decision-making tasks. There is still much to explore with this novel technique, especially in humans, as to date, there has not been a study about the potential neural substrates.

It will also be crucial to evaluate the potential biases behind the characteristics of different states, the reasons for switching between states, and how these factors differ with the characteristics of the task at hand, such as whether correct answers are rewarded or not.

Regarding our second objective, which was to assess the confidence level that each participant feels about each of their answers in relation to each state as a measure of their awareness of their own internal state, we found significant differences in our confidence analysis. When ordering states by precision in a descending manner, we observed differences between states 1 and 2 in the 2 states model, both states 1 and 2 in comparison to state 3 in the 2 states model, and among all states except between states 1 and 2 in the 4 states model. This suggests that participants tend to have high awareness of their internal state, and that states with lower precision are associated with lower confidence.

It would then be pertinent to investigate how this metacognition of internal states is impacted by various pathologies. As mentioned in the introduction, Rouault et al. (2018) demonstrated that the overall confidence in one's own perceptual decisions was associated with different psychiatric symptom dimensions. Specifically, a symptom dimension related to anxiety and depression was associated with reduced confidence and increased metacognitive efficiency. In contrast, a dimension describing compulsive behaviour and persistent thoughts corresponded with enhanced confidence and reduced metacognitive efficiency.

In the same paper from which the dataset of the present study was sourced, psychiatric symptoms did not predict changes in accuracy. However, it has been repeatedly observed that various pathologies can impact performance in perceptual decision-making (Foryś et al., 2017; Jassim et al., 2022; Reckless et al., 2015). Additionally, confidence distributions have been shown to be very similar for tasks involving different sensory modalities but possessing similar structure, and these distributions tend to be stable over time (Ais et al., 2016). This suggests that confidence could serve as an efficient mean of studying various pathologies, even in individuals with sensory disabilities.

Regarding the neural basis of confidence in perceptual decision-making, Fleming et al. (2012) demonstrated that the activity in the right rostrolateral prefrontal cortex (rlPFC), as captured by fMRI, fulfils three prerequisites for playing a significant role in decision-making metacognition. The rlPFC exhibited heightened activity during self-reporting compared to a corresponding control scenario. The rlPFC activity correlated with the confidence levels reported by participants, and the strength of this correlation was indicative of individual metacognitive abilities. Additionally, they detected an enhanced functional interconnectivity between the right rlPFC, the contralateral PFC, and the visual cortex during the generation of metacognitive reports. Given that the rostrolateral prefrontal cortex is known to play a pivotal role in various situations and cognitive processes (Gilbert et al., 2006), it would be beneficial to analyse the activity of this region related to different states.

Indeed, improving methods for modelling both perceptual decision-making and its relationship with confidence is critical. It can shed light on both the perceptual and cognitive processes involved in this complex set of phenomena, and how they differ in states of health and illness. This paves the way for a better understanding of these processes, which could be crucial in developing new treatments and diagnostic protocols, as has been done in the past (Dully et al., 2018).

Furthermore, studying behaviour, including perceptual decision-making, is essential in neuroscience. Some even argue that in most cases, understanding the neural mechanisms underlying behaviour is best accomplished after conducting behavioural analysis. This underscores a broader perspective in neuroscience regarding the connection between brain and behaviour: studying behaviour enhances comprehension, while examining the brain reveals causal relationships (Krakauer et al., 2017).

5.- Conclusions

The main conclusions of the present work are: (i) using the classical psychometric function model (1 state model) to analyse behaviour in a perceptual task is not adequate for many human participants because it does not take into account the behavioural fluctuations that participants could experience with time; models that incorporate 2 or more states are needed. (ii) Selection across models with different numbers of states depends on the number of trials.

When it comes to the participants' metacognition, (iii) confidence across states ordered by precision does differ; states with a lower precision are associated with a lower confidence, which means that participants may be metacognitive aware about their internal state.

6.- Bibliography

Aguillon-Rodriguez, V., Angelaki, D., Bayer, H., Bonacchi, N., Carandini, M., Cazettes, F., Chapuis, G., Churchland, A. K., Dan, Y., Dewitt, E., Faulkner, M., Forrest, H., Haetzel, L., Häusser, M., Hofer, S. B., Hu, F., Khanal, A., Krasniak, C., Laranjeira, I., … Zador, A. M. (2021). Standardized and reproducible measurement of decision-making in mice. ELife, *10*, 1–28.

Ais, J., Zylberberg, A., Barttfeld, P., & Sigman, M. (2016). Individual consistency in the accuracy and distribution of confidence judgments. Cognition, 146, 377–386.

Ashwood, Z. C., Roy, N. A., Stone, I. R., Urai, A. E., Churchland, A. K., Pouget, A., & Pillow, J. W. (2022). Mice alternate between discrete strategies during perceptual decision-making. Nature Neuroscience, *25*(2), 201–212.

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An ERP study. Science Direct, 25, 107–116.

Bolkan, S. S., Stone, I. R., Pinto, L., Ashwood, Z. C., Iravedra Garcia, J. M., Herman, A. L., ... & Witten, I. B. (2022). Opponent control of behavior by dorsomedial striatal pathways depends on task demands and internal state. *Nature neuroscience*, 25(3), 345-357.

Burr, D., & Ross, J. (2008). A Visual Sense of Number. In *Current Biology* (Vol. 18, Issue 6).

Dully, J., McGovern, D. P., & O'Connell, R. G. (2018). The impact of natural aging on computational and neural indices of perceptual decision making: A review. Behavioural Brain Research, *355*(2018), 48–55.

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal Contributions to Metacognition in Perceptual Decision Making. The Journal of Neuroscience, 32(18), 6117–6125.

Foryś, U., Bielczyk, N. Z., Piskała, K., Płomecka, M., & Poleszczuk, J. (2017). Impact of time delay in perceptual decision-making: Neuronal population modeling approach. Complexity, *2017*.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional Specialization within Rostral Prefrontal Cortex (Area 10): A Meta-analysis. Journal of Cognitive Neuroscience, 18(6), 932–948.

Hauser, C.K., Salinas, E. (2014). Perceptual Decision Making. In: Jaeger, D., Jung, R. (eds). Encyclopedia of Computational Neuroscience. Springer, New York, NY.

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. Nature Reviews Neuroscience, 9(6), 467–479.

Jassim, N., Owen, A. M., Smith, P., Suckling, J., Lawson, R. P., Baron-Cohen, S., & Parsons, O. (2022). Perceptual decision-making in autism as assessed by “spot the difference” visual cognition tasks. Scientific Reports, *12*(1), 1–7.

Kelly, S. P., & O’Connell, R. G. (2015). The neural processes underlying perceptual decision making in humans: Recent progress and future directions. Journal of Physiology Paris, 109(1–3), 27–37.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron, 93(3), 480–490.

Lindermann (2020). ssm [Source code]. GitHUB. https://github.com/lindermanlab/ssm

Mamassian, P. (2016). Visual Confidence. Annual Review of Vision Science, 2, 459–481.

Odoemene, O., Pisupati, S., Nguyen, H., & Churchland, A. K. (2018). Visual evidence accumulation guides decision-making in unrestrained mice. Journal of Neuroscience, *38*(47), 10143–10155.

Rahnev, D., Desender, K., Lee, A. L. F., Adler, W. T., Aguilar-Lleyda, D., Akdoğan, B., Arbuzova, P., Atlas, L. Y., Balcı, F., Bang, J. W., Bègue, I., Birney, D. P., Brady, T. F., Calder-Travis, J., Chetverikov, A., Clark, T. K., Davranche, K., Denison, R. N., Dildine, T. C., Double, K. S., … Zylberberg, A. (2020). The Confidence Database. Nature human behaviour, 4(3), 317–325.

Reckless, G. E., Andreassen, O. A., Server, A., Estefjells, T., & Jensen, J. (2015). Negative symptoms in schizophrenia are associated with aberrant striato-cortical connectivity in a rewarded perceptual decision-making task. NeuroImage: Clinical, *8*, 290–297.

Rouault, M., Seow, T., Gillan, C. M., & Fleming, S. M. (2018). Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. Biological Psychiatry, *84*(6), 443–451.

Urai, A. E., Braun, A., & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nature Communications, *8*.