Measuring Behavioural Variability in Random Dot Motion Task

Hariharan Purohit*1, Anjali Sifar*2, Nisheeth Srivastava*3
*Department of Cognitive Science, IIT Kanpur
[1]hariharan22@iitk.ac.in , [2]sanjali@iitk.ac.in , [3]nsrivast@iitk.ac.in

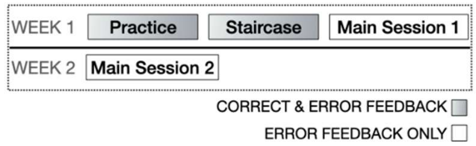
1 Introduction

Random dot motion discrimination task has a long history in its usage to understand perceptual decision-making (Britten et al., 1992; Gold et al., 2007; Ratcliff et al., 2018). Trial-by-trial fluctuations in the behavioural response have been measured with the use of a large number of trials, termed internal noise. On the other hand, the variation in response on account of the different instances of the same class of stimuli is termed external noise. Ratcliff et al. (2018) set out to dissociate the two sources of noise through a double-pass experimental manipulation - in which the same stimuli were presented and separated by a large number of trials. In the above manipulation, however, the dissociation between the two sources of noise was not complete on account of possibly multiple confounds like presentation order, short term fluctuations, and so on.

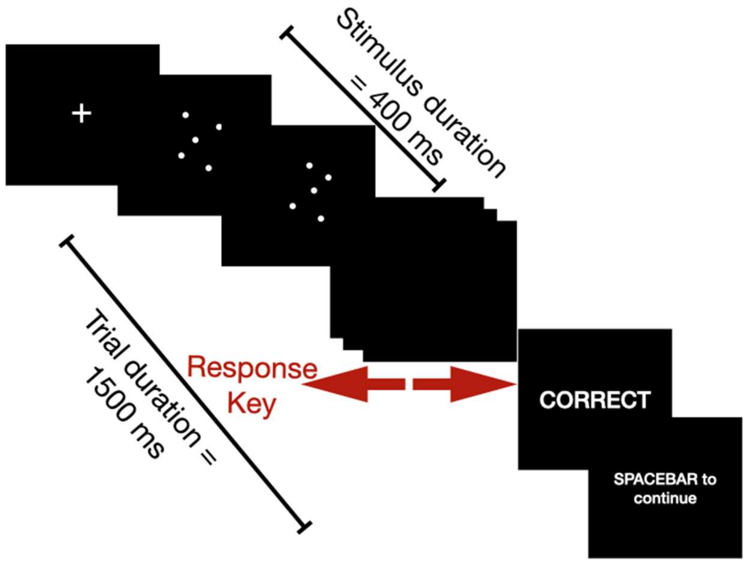
To isolate the sources of behavioural variability, we used the same RDM discrimination task within a test-retest reliability framework. This was accomplished by providing a week-long gap between two task sessions with the assumption that any learning effects are discounted via the gap while controlling for stimulus presentation and trial sequence. This provided us with a measure of the minimal source of variability, i.e., within-person variability.

In the linear modelling tradition, the observed outcome variability can be accounted for by the trial and subject variation. Any unexplained variation goes into the residuals or what is commonly known as measurement error. However, any shift in the observed variable across sessions assuming nothing has changed including the measurement error, subject, and stimulus variation can be explained in terms of the random fluctuations in the information processing system, i.e., internal noise.

This internal noise imposes limits on the reliability of the observed phenomenon. This has implications in the empirical claims made by individual differences, where the observed effect between two populations can be explained purely in terms of internal noise if it is less than the observed effect between repeated measurements. If the deviation is systematic, it owes its explanation in terms of the external noise - while internal noise will shift the observations randomly such that it limits the observed consistency between the two repeated sets of measurements.



(b) Overall task structure.



(a) General trial structure.

Figure 1. Experiment design.

2 Methods

Participants performed a random dot motion discrimination task inspired by the Ratcliff et al. (2018) paradigm, where the goal is to indicate the direction of apparently moving dots i.e., randomly positioned dots presented in succession at a frame rate of 60 Hz (Figure 1b). Each dot carries information about the direction of motion depending upon whether it is a signal or noise dot in a given frame. The probability that a given dot carries the signal is guided by the coherence variable, which changes every trial. The identity of a dot (signal/noise) expired every 3 frames to ensure that the decisions are based on the global motion of dots. Stimuli consisted of five dots moved at a speed of 4 pixels/frame in an invisible circular aperture of 100°.

We used the retest reliability paradigm with the same sequence and position of dots for the two main sessions separated by one week as shown in Figure 1a. Before the first main session, participants were given 90 practice trials to get familiar with the task. Thereafter we measured the accuracy threshold using the 4-up, 1-down staircase method with 30 reversals. The mean value of the last 6 reversals $\pm$ 5% provided us with three levels of subjective coherence.

In the main session, 19 participants (10M, with informed consent, normal or corrected to normal vision) were presented with 300 trials at each of the three coherence levels, randomly interspersed across 10 blocks of 90 trials each. They were compensated INR 500 for completing all four sessions.

3 Results

We measured response time (RT) and accuracy for each trial. Trials with missing responses in either session were excluded from both sessions, giving us a mean of 298 trials per coherence level.

A two-way repeated measures ANOVA was conducted with coherence and session as the main factors and mean RT as the dependent variable. We observed the main effect of Session (F (1,18) = 10.22, p = 0.005, eta^2 = 0.36) and Coherence (F (2,36) = 11.51, p = 0.001, eta^2 = 0.39) with no interaction.

Pearson correlation coefficient was used to measure the consistency across the two sessions. We observed a correlation coefficient r = 0.85 for mean RT and r = 0.8 for mean accuracy. As discussed before, less than perfect correlation indicates within-person random fluctuations.

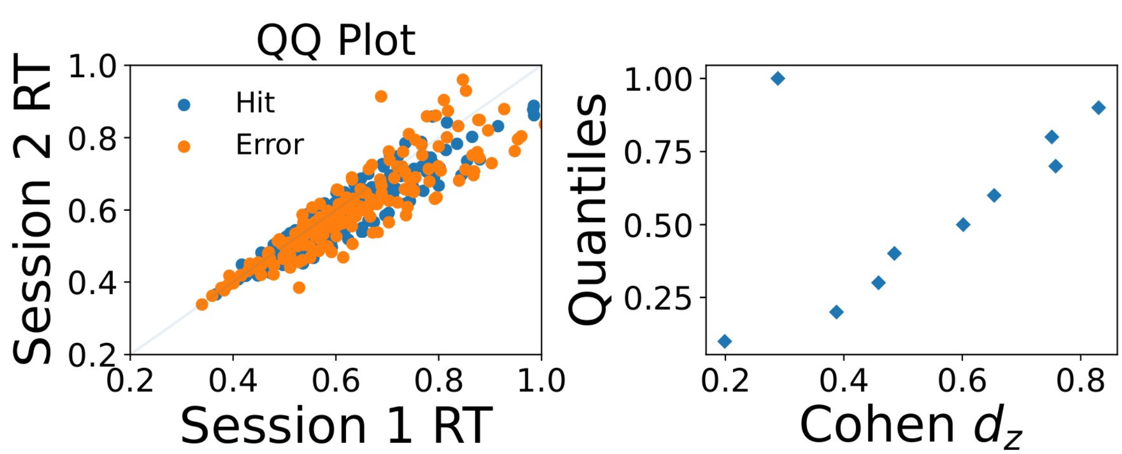


Figure 2. Divergence between sessions. a) Quantile-Quantile plot between sessions for observed Reaction time. b) Divergence as measured by Cohen's d_z at each quantile.

A QQ plot between the repeated measurements of RT quantiles shows the spread on account of the random fluctuations between the two sessions. In Figure 2, this analysis is collapsed for hit and error RT distributions. Each dot represents one quantile per coherence and subject. This variation cannot be accounted for by variation in a subject or task stimulus.

Quantifying this deviation, we measured Cohen's d between the two observations using repeated measurements. For every coherence level and subject, the aggregate measures for RT and accuracy were computed. The effect size between the repeated set is $d = 0.23$ for mean RT and $d = 0.2$ for accuracy.

Next, we also measured the consistency at the individual trial level. For RTs, collapsed across all coherence levels, Pearson correlation yielded $r = 0.16$. For accuracy, Cohen's kappa ($\kappa = 0.23$) is used for binary choices.

4 Discussion

Reliability forms the cornerstone of all empirical claims. The test-retest framework gives us the means to check how much a test/task measuring a construct of interest (e.g., perceptual decision) produces the same or different results at different time points. Interpreting empirical claims would then be affected by what this framework would show.

In this study, we aim to characterize the nature and contribution of different sources of variability by extending the double-pass manipulation to the retest reliability paradigm. The tradition of pushing the observed cognitive behavioural variability to the external noise assumes at the heart that there is some perfect encoding of the stimulus by a subject that is contaminated by noise. However, consistency measures across repeated sets of measurements give us a quantitative measure of the internal noise at the aggregate behaviour level, when subject and stimulus variation has been accounted for.

The presence of within-person random fluctuations provides us with a qualitative tool to assess the purely quantitative empirical claims. It provides boundaries to the empirical claims, for example, in individual differences studies where the internal noise indicates the overlap of two population distributions, which cannot be observed using one set of behavioural response elicitation.

In sum, using the consistency and deviation metrics over the longitudinal observations given everything is the same - we show that the notion of ground truth as assumed in the study of cognitive science needs to be reformulated with in terms of the within person behaviour fluctuations.

References

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, *12*(12), 4745-4765.

Gold, J. I., Shadlen, M. N. (2007). The neural basis of decision making, *Annual review of neuroscience*, *30*(1), 535-574.

Ratcliff, R., Voskuilen, C., & McKoon, G. (2018). Internal and external sources of variability in perceptual decision-making. *Psychological review*, *125*(1), 33.