

Please cite this abstract as: Clark, K., Pennington, C. R., Hedge, C., Lee, J. T., Petrie, A. C. P. (2019). Test-retest reliability for common tasks in vision science. *Journal of Vision*, 19(10), 86d. <https://doi.org/10.1167/19.10.86d>.

Test-retest reliability for common tasks in vision science
Vision Sciences Society Annual Meeting Abstract

Historically, research in cognitive psychology has sought to evaluate cognitive mechanisms according to the average response to a manipulation. Differences between individuals have been dismissed as “noise” with an aim toward characterising an overall effect and how it can inform human cognition. More recently, research has shifted toward appreciating the value of individual differences between participants and the insight gained by exploring the impacts of between-subject variation on human cognition. However, recent research has suggested that many robust, well-established cognitive tasks suffer from surprisingly low levels of test-retest reliability (Hedge, Powell, & Sumner, 2018). While the tasks may produce reliable effects at the group level (i.e., they are replicable), they may not produce a reliable measurement of a given individual. If individual performance on a task is not consistent from one time point to another, the task is therefore unfit for the assessment of individual differences. To evaluate the reliability of commonly used tasks in vision science, we tested a large sample of undergraduate students in two sessions (separated by 1–3 weeks). Our battery included tasks that spanned the range of visual processing from basic sensitivity (motion coherence) to transient spatial attention (useful field of view) to sustained attention (multiple-object tracking) to visual working memory (change detection). Reliabilities (intraclass correlations) ranged from 0.4 to 0.7, suggesting that most of these measures suffer from lower reliability than would be desired for research in individual differences. These results do not detract from the value of the tasks in an experimental setting; however, higher levels of test-retest reliability would be required for a meaningful assessment of individual differences. Implications for using tools from vision science to understand processing in both healthy and neuropsychological populations are discussed.