112. Individual Differences Neuroscience: From Within-To Between-Subjects Differences in Psychopathology

Greg Hajcak¹, Roman Kotov², Alexandria Meyer¹, Katherine Luking³, and Brady Nelson²

¹Florida State University, ²Stony Brook University

**Background:** Human neuroscience research has produced few significant advances in diagnosing, predicting, treating, and preventing mental illness. I’ll argue that one major reason for poor progress in clinical neuroscience is a failure to adequately consider the distinction between within- and between-subjects comparisons. Explaining differences between people requires measures with good psychometric properties. Robust, group-level differences from within-subject contrasts does not imply adequate internal consistency for understanding individual differences.

**Methods:** I will present a range of neuroimaging data (i.e., both fMRI and EEG) from multiple tasks (i.e., emotional face matching, simple reward). I’ll focus on internal consistency, and further issues related to difference scores in large sample of adolescents (N=177). Finally, I will examine sensitivity, specificity, and positive and negative predictive value of neural measures, when used alone and in conjunction with self-report measures.

**Results:** In our data, certain within-subjects effects (e.g., amygdala activation to faces versus shapes) were completely unreliable as contrast-based scores (r ~ .05). Other measures (e.g., increased striatal activation to reward versus loss) were both robust and internally reliable (r ~ .30). We demonstrate how neural measures can be used to improve classification when combined in series with self-report measures.

**Conclusions:** Neuroscience studies need to distinguish between robust within-subjects effects and reliable between-subjects variability. These psychometric issues may explain some poor progress in human neuroscience research. I’ll propose basic approaches for choosing and optimizing tasks and resultant fMRI/ERP measures, and how they might be combined in the real world to improve classification of psychopathology.

**Supported By:** NIMH; R01 MH097767

**Keywords:** Neuroimaging, Psychometrics, Classification

113. Lessons From Test Construction: The Establishment of Norms and Item Analysis for Cognitive Affective ERPs

Annmarie MacNamara¹, Daniel Foti², Michael Imburgio¹, David Clark¹, and Anna Weinberg¹

¹Texas A&M University, ²Purdue University, ³Michigan State University, ⁴McGill University

**Background:** Cognitive affective neuroscience tasks are not typically subjected to the rigors of test construction. For example, although the error-related negativity (ERN) has been shown to predict risk for psychopathology, it is unknown what values might constitute a “normal” or “abnormal” ERN (there are no established norms). In addition, task stimuli for assessing individual differences in neural measures of affective reactivity (e.g., positive and negative pictures) are rarely selected to maximize these differences (item analysis is not performed).

**Methods:** We present examples of norming and item analysis as two ways to improve neuroscientific measures for the examination of individual differences. First, we examined the distributions of event-related potential (ERP), the ERN, in an unselected sample of adults (N = 800). Second, we used item-response theory (IRT) modeling techniques to identify emotional pictures that best differentiated individuals (N = 80). Pictures that were the most informative about individual levels of latent trait theta (affective reactivity) were not necessarily the most likely to elicit the largest LPPs.

**Conclusions:** It may be worthwhile to establish norms for a variety of neuroscientific measures. In parallel, it may be fruitful to work to improve test design; item response theory may provide a useful means of selecting stimuli.

**Supported By:** A. MacNamara is supported by National Institute of Mental Health grant, K23MH105553

**Keywords:** Event Related Potentials, Individual Differences, Emotion, Error Monitoring, Assessment

114. Test-Retest Reliability of Task-Evoked Bold fMRI: Implications for Individual Differences Neuroscience

Annchen Knodt¹, Maxwell Elliott¹, and Ahmad Hariri¹

¹Duke University

**Background:** The utility of neuroimaging phenotypes for the study of individual differences depends on how reliably they can be measured over time. Task-evoked activity measured with BOLD fMRI is increasingly used for mapping variability in behavior and risk for mental illness. However, the test-retest reliability of these measures may not be sufficient for individual differences research.

**Methods:** Here, we calculate the test-retest reliability of regional activation for 7 fMRI tasks from the Human Connectome Project (HCP) where 45 participants were scanned twice, with a mean test-retest interval of approximately 140 days. Test-retest reliability of task-evoked activation within a priori regions of interest (ROIs) was quantified using the intra-class correlation coefficient (ICC).

**Results:** Robust (i.e., p < 0.05, FWE corrected) task-evoked signal was observed for all tasks within target ROIs at each time point. However, the test-retest reliability was generally poor regardless of the specific task-based contrast examined.