SYMPOSIUM
The Suitability of Neuroscientific Measures for Assessment of Individual Differences: Strategies for Improvement
12:30 p.m. - 2:30 p.m.
Chair: Annmarie MacNamara

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

Greg Hajcak[1], Roman Kotov[2], Alexandria Meyer[1], Katherine Luking[2], and Brady Nelson[2]

[1]Florida State University, [2]Stony 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.

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

Annmarie MacNamara[1], Daniel Foti[2], Michael Imburgio[1], David Clark[3], and Anna Weinberg[4]

[1]Texas A&M University, [2]Purdue University, [3]Michigan State University, [4]McGill 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) in terms of affective processing, as measured by ERP, the late positive potential (LPP).
Results: First, the 25th, 50th, and 75th percentile scores for the ΔERN (error - correct) were -2.37μV, -5.41μV, and -8.65μV, respectively. Second, item-response models fit the data well and were reliable (αs > .80; discrimination > .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 task design; item response theory may provide a useful means of selecting stimuli.

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

Annchen Knodt[1], Maxwell Elliott[1], and Ahmad Hariri[1]

[1]Duke 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 intraclass 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.