Theory Into Practice© (TIP) Assignment Series Reliability

Data to establish reliability coefficients for the *Theory Into Practice©* battery and individual instruments were gathered by administering each of six TIP assignments to 75 pre-service teacher candidates over the course of a semester-long Educational Psychology class at a university in the Southeast United States. Scoring of the assignments was completed by independent raters who were trained to use the assignment scoring rubrics. During rater training, inter-rater reliability was checked regularly to achieve and then maintain a coefficient of r > .90. Following the training, each assignment was scored by two raters. In cases where there was not satisfactory agreement, a third rater was used to settle the scoring. A split-half correlation was performed on the resulting data from the 75 subjects aggregated over all six TIP assessments. A full-battery (aggregated over all six instruments) reliability coefficient of r = 0.932 was calculated, then corrected for the split-half nature of the analysis (assuming equivalence of the groups) by applying a Spearman-Brown correction. The resulting corrected correlation of r = 0.965 was attained as the reliability coefficient for the entire battery. The table below presents disaggregated uncorrected and corrected reliability coefficients attained for each of the TIP assignments.

Table 1: Pearson's r Reliability Coefficients for Individual TIP Assignments (Uncorrected and Corrected Using Spearman-Brown Correction)

	TIP 1	TIP 2	TIP 3	TIP 4	TIP 5	TIP 6
Uncorrected	0.978	0.952	0.952	0.828	0.846	0.839
Corrected	0.989	0.975	0.975	0.906	0.917	0.912

Resulting corrected measures of instrument reliability fall within the "excellent" range (r >= 0.90) for all instruments (Cicchetti, 2001, p. 697). These results suggest that the instruments are statistically reliable both individually and in the aggregate.

Theory Into Practice© (TIP) Assignment Series Validity

The *Theory Into Practice©* battery and individual instruments include content commonly held as foundational to best educational classroom practices. All instruments are based upon current theory relative to human development, learning and cognition, and all performance expectations of respondents per the scoring rubrics for these instruments, reflect realistic and authentic applications of these theories into pedagogical practice. Because the instruments require respondents to apply theoretical content to classroom situations, the exercises are authentic to the "real" situations pre-service teachers may encounter as they begin their clinical experiences and, later, as they become

program completers and matriculate to full-time employment as teachers. For these reasons, both content and construct validity are claimed for the instruments.

Cicchetti, D.V., (2001). The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology, Vol.23* (5), pp 695-700.