

## Definition of Terms Used in the Additional Statistics Sheet of the Item Analysis File

(in their order of appearance on the sheet)

**Matched Respondents** are the respondents who answered both some pre-test and some post-test items and whose answers could be matched thanks to a code that protected their anonymity.

**Pre-test** is a test given to course participants at the beginning of the course to assess their level of knowledge of the topics covered by the course. The **Post-test** is a comparable test given at the end of the course.

**Mean Test Score (% correct)**: Each participant's test score is the percentage of correctly answered test items. The mean test score is the average of the test scores of all respondents or of all matched respondents.

**Median Test Score (% correct)** is the middle of all respondents' test scores; that is, half of the respondents have scores greater than the median, and half have scores less than the median.

**Standard Deviation** is a measure of how widely the respondents' scores differ from the mean test score. About 2/3 of all scores fall within the range of the mean test score minus one standard deviation and the mean test score plus one standard deviation.

**Standard Error of Mean** is an estimated standard deviation of the obtained mean test score. This statistic provides an estimate of how the mean test score might vary when the course/test is repeated with similar groups of participants.

**Skewness Ratio** is an indicator of the symmetry of the score distribution. A ratio of zero means the distribution is symmetric. A positive skew means that relatively more scores are below the mean score and a few scores are above the mean score. A negative skew means that relatively more scores are above the mean score and a few scores are below the mean score. Skewness could be a concern if the ratio is higher than +2 or lower than -2.

**Cronbach's Alpha** is an index of how consistently or reliably a test measures achievement. The index varies between 1 (a perfectly reliable test) and 0 (a perfectly unreliable test). Reliability in general is a matter of confidence in the dependability of a test. It usually increases when tests have more items, and when the items are of moderate difficulty and homogenous in content.

**Measurement Error** indicates how much an individual's score may vary with retesting. About 2/3 of the time an individual can be expected to score within one standard error of measurement of the individual's obtained score on the test when retested.

**Number of Respondents** is the number of participants in the course who answered at least a portion of the items in the test.

**Response Rate (%)** is the number of respondents divided by the total number of participants in the course reported by the training team.

**Number of Items** is the number of Level-2 evaluation questions in the pre-test or post-test.

**Total Number of Participants** is the number of people who attended the course as participants at any point in time, even if they did not complete the Level-2 evaluation test forms.

**Learning Gain** is the mean post-test score minus the mean pre-test score for matched respondents.

**Matched Gain** is the post-test score minus the pre-test score of each matched respondent.