Statistics Concept Inventory
The University of Oklahoma
NSF – Assessment of Student Achievement Grant # 0206977
PI: Teri Reed Rhoads
co-PI: Teri J. Murphy

The Statistics Concept Inventory (SCI) is designed to be similar in format to the Force Concept Inventory (FCI), which has been successful in assessing student understanding of Newton's law and transforming Physics teaching to improve understanding. The FCI is a multiple choice test of about 30 questions, with the incorrect answers identifying student misconceptions. The SCI is also a multiple choice testing instrument that is being developed to assess student understanding of fundamental statistics concepts and to identify commonly held misconceptions.

The SCI development began in fall 2002 with a 32-item instrument. The results from this pilot study were presented at the 2003 Frontiers in Education Conference (Stone, et al, 2003). The basic findings were that males significantly outscored females and mathematics majors significantly outscored social science majors. However, poor testing conditions in one class and the small magnitude of the differences make these findings less meaningful. The SCI was found to be slightly lacking on reliability, with a Cronbach's alpha of 0.6114.

Several methods have been used to improve the SCI. Objective analyses include the discriminatory index, alpha-if-deleted (how the overall alpha would change if each question were deleted), and analyzing answer distributions. Subjective methods include focus groups and the researchers' own opinions based on their experience with introductory statistics. This basic process is carried out after each semester. In fall 2004 the instrument will be reviewed by the External Advisory Board which consists of persons interested in the teaching of statistics, the assessment of the teaching of statistics, and experienced concepts inventory developers.

The reliability of the SCI has improved since the pilot study. Cronbach's alpha is around 0.65 on the pre-test and 0.72 on the post-test for most classes. There is slight concern that universities outside OU have been lower, but more data is needed to draw conclusions.

The content validity of the SCI is ensured by consulting a faculty survey which rated the importance of statistics topics. Priority is given to topics which scored highly, while some items have been discarded for lack of importance. The AP Statistics outline is also consulted for breadth of coverage. Concurrent validity is measured by correlating the SCI scores with the course percentage grade. Generally, the SCI is valid for engineering statistics courses but not for those in the mathematics department.

Scores on the SCI are remarkably consistent from class to class and semester to semester. The post-test scores for most classes are between 48% and 52%. Pre-test scores show slightly more variability, ranging from high 30%'s to mid 40%'s. The scores and gains (from pre to post) on the SCI are similar to those found on early testing of the FCI in classes which use the traditional lecture format. A full-information Maximum-Likelihood Factor Analysis or FIML Factor Analysis provided support for four areas of development for the SCI; probability, descriptive statistics, inferential statistics and graphical interpretations. Also, the Factor Analysis showed support for both a general or "broad" factor, and the specific domains as mentioned. A fifth area, advanced statistics topics, has been considered and further testing is required. More complete results can be found on the web page (http://coecs.ou.edu/sci). If you or your university is interested in utilizing our instrument in one or more of your statistics courses, please contact us at teri.rhoads@ou.edu for a full copy of the instrument.

Last updated July 23, 2004