International Statistical Literacy Project

Search ISLP for Go

HOME WHO WE ARE MISSION HISTORY PEOPLE EVENTS

Back to ISLP Home

ASSESSMENT OF STATISTICAL LITERACY

This page was created and maintained by : Sara Finney, Kenn Barron, and S. Jeanne Horst until July 2007.

INTERNATIONAL INITIATIVES FOR COMPARATIVE ASSESSMENT OF LITERACY IN COUNTRIES

PIAAC

Programme for the International Assessment of Adult Competencies PIAAC (OECD) is a plan by the OECD. It addresses critical statistical skills needed by adults as part of general everyday or workplace functioning. It is somewhat similar in general terms to the PISA assessment program of high school students which is now implemented in dozens of countries. One of the several domains of PIAAC will be numeracy, and one of the strands in it will be knowledge of statistica (data and chance). A talk by Iddo Gal at the IASE Satellite Conference in Guimaraes, August 19-21 2007 contains valuable information on this project and the potential contribution to it of statistics educators.

FREE GENERAL INTERNET RESOURCES

A.R.T.I.S.T Project

"This website provides a variety of assessment resources for teaching first courses in Statistics. Currently we provide articles and weblinks related to assessing student outcomes. In the near future, this site will contain assessment items and tasks, provide online testing, offer guidelines for using the assessment items and tasks, and allow for the collection and compilation of data for research and evaluation purposes" (Quoted from website). More detail is given below on some of the items contained in the website.

BOOK PROVIDING OVERVIEW OF ASSESSMENT IN STATISTICS EDUCATION

Gal, I., and Garfield, J. (Editors) The Assessment Challenge in Statistics Education

http://www.stat.auckland.ac.nz/~iase/publications/assessbk/ (Free electronic version)

ISBN 90-5199-333-1. Published by IOS Press and the International Statistical Institute, 1997, 296 pages. Cost: 10 Euro. For ordering information see http://isi.cbs.nl/sale-iase.htm. (Hardcover version)

Description of Book: "This book discusses conceptual and pragmatic issues in the assessment of statistical knowledge, reasoning skills, and dispositions of students in diverse contexts of instruction, both at the college and precollege levels" (p. xi). The authors contributing to this edited book do not simply present tests/instruments/measures for assessing statistical knowledge; instead, the 19 chapters focus on the importance of assessment, highlight pitfalls and challenges when assessing student understanding, and, when possible, present example items or tasks in order to help readers develop their own assessments. The chapters discuss different assessment methods (e.g., authentic assessment, multiple-choice exams, portfolios, projects, oral communication) used to evaluate student understanding of a wide variety of statistics-related content (e.g., data reduction, graphing, probability, randomness, combinatorics, interrelationships among statistics concepts, use of statistical software). Again, the chapters discuss different issues related to the assessment process and, in turn, provide guidelines, suggestions, and examples in order to create new assessment tools. Therefore, this book serves as a good introduction and overview of the assessment challenge in statistics education.

ASSESSMENT OF SKILLS

Assessment Resource Tools for Improving Statistical Thinking (A.R.T.I.S.T.)

https://app.gen.umn.edu/artist/ (homepage) https://app.gen.umn.edu/artist/user/login.asp (Assessment Builder) **Paper Presentation:** Garfield, J., DelMas, R., Chance, B. (2003) The web-based ARTIST: Assessment resource tools for improving statistical thinking. Paper presentation at the American Educational Research Association, Chicago, IL.

Population: Post-secondary students enrolled in their first statistics courses is the main audience. But, many of the items are applicable to other audiences.

Background: The ARTIST website is funded by NSF. The principal investigators are associated with the University of Minnesota and California Polytechnic State University. The website was developed as a resource for the assessment of statistical literacy, statistical thinking, and statistical reasoning.

Description: The ARTIST website contains an "Assessment Builder", which includes over 1,000 items. Use of the Assessment Builder is free; however, users are asked to register and create a password. Types of items that are offered include forced-choice, open-ended, and performance items. Content areas include the following: data types, data collection, data representation, measures of centrality and spread, comparing groups, measures of position, normal distribution, bivariate data, linear regression, categorical data, chi-square, probability, samples and sampling, significance testing, distributions, confidence intervals, and one-way ANOVA. After the user selects the type of item and content areas desired, the website presents a listing of items, a description of content areas, and processes involved (i.e., statistical reasoning, thinking, literacy, or computation). From the list, users select the items they wish to include in their assessments. The assessments easily download in .rtf format, and are saved for later access when visiting the website. In addition to the Assessment Builder, there are numerous resources available on the ARTIST website. Most of the resources include links to articles, research papers and national conference presentations. Research instruments, practical advice in a Q&A format, and information on upcoming events are also included. Faculty are invited to participate in pilot testing, contribute items or materials, and give feedback.

Scoring: Item responses are not provided for security reasons and to encourage instructors to formulate their own responses.

Psychometric Properties: Because each user's assessment will be comprised of different items, this will vary from assessment to assessment.

Location of Items: The Assessment Builder, which contains a bank of over 1,000 items, may be found at https://app.gen.umn.edu/artist/user/login.asp.

Balanced Assessment. Harvard Group Balanced Assessment in Mathematics Project. Assessing mathematical understanding and skills effectively

(AMUSE).

Publications are available for free download at http://balancedassessment.concord.org(All in English and some in Spanish and Hebrew): 1. *Assessing mathematical understanding and skills effectively (AMUSE).* 2. *Mathematics content and process scoring system: A multidimensional approach to the scoring of balanced assessment tasks;* and 3. *Sample scoring packet for a specific task (includes student work).*

Population: Kindergarten (Age: 5 years) students through advanced secondary school students (The items are written in English; however, some are available in Spanish and Hebrew.)

Background: The materials on the Balanced Assessment website were originally created by the Harvard Graduate School of Education, funded by a 1993 NSF grant.

Description: While the Balanced Assessment tools are specifically aimed at mathematics assessment, some of the items are relevant to statistical literacy assessment. The authors emphasize the distinction between objects (i.e., content) and action (i.e., process). Specifically related to statistical literacy, some of the objects include chance, relative frequency and probability, discrete and continuous data. Actions, or processes, related to statistical literacy include information gathering, necessary and/but not sufficient conditions, inferring and drawing conclusions, and communicating. Rather than referring to the items as "problems", they are termed "projects". The emphasis is on encouraging students to think through a variety of options, think critically, and develop "'habits of mind' about organizing and analyzing complexity" (Harvard Group, 1995b, p. 13). Group work, with both written and oral reporting of project results, is encouraged. For example, "Mixed-up Socks" is a task that was written for elementary level students, grades 3 to 5, which emphasizes chance. Students are told that different colored socks are mixed up in one drawer (6 of one color, and 4 of another). They are asked to state the fewest number of socks it would take, if pulled out one at a time, to be certain that they had a matching pair. Other examples include: Measure Me (elementary; interpreting and drawing charts), Broken Measures (elementary; quantity, measuring and interpreting), Graphs Resource (high school; identify a quantity and graph it), The Contest (high school; chance/data, inferring and drawing conclusions), Dubious Dice (high school; relative frequency distributions, histograms, chance, data, creating graphs, inferring and drawing conclusions, communicating), Genetic Codes (high school; chance/data, inferring and drawing conclusions, communication), Tie Breaker (high school; interpreting charts, inferring and drawing conclusions), Stock Market (high school; interpreting a graph, inferring and drawing conclusions), Survey Says (high school; interpreting percentages, creating a graph). These are just a few of the items that are available. They are listed here to enable the first-time user of the Balanced Assessment website to quickly sort through the strictly mathematical items to get a feel for what is available for statistical literacy assessment. Scoring: The authors have created scoring rubrics for each of the tasks/items. The tasks are weighted on both content (e.g., chance and data) and process (e.g., inferring/drawing conclusions and communicating) from 0 (not present at all) to 4 (a dominant presence). Scoring rubrics involve rating the student symbolically on their performance using icons. The ratings may also be considered ordinally, from 0 (little evidence of skill or understanding) to 3 (deep level of skill or understanding). The developers emphasized that students' skill levels should not be judged on the basis of one task. Two publications that explain the scoring method, as well as scoring rubrics for the individual tasks may be found at http://balancedassessment.concord.org.

Psychometric Properties: not provided

Location of Items: Items are listed both alphabetically and by number at http://balancedassessment.concord.org. Items are numbered with a prefix that indicates the educational level of the student for which they were written. Individual descriptions of the items include information on grade level, content area, and mathematical processes. Items are available as Word documents for download or printing at no cost. They may be duplicated for classroom use, but are otherwise copyrighted. Some of the solutions and scoring rubrics are freely available; others require a password, which is obtained by ordering materials through the Concord Consortium. Materials that are available include a CD-ROM, a Spanish packet, and packets that address various grade levels (most prices are in the $10 to $25 range). Scoring rubrics are password protected in order to give teachers control over whether or not they want the solution available to their students. However, teachers may freely share their password with other educators.

Garfield, J.B. Statistical Reasoning Assessment (SRA)

http://www.stat.auckland.ac.nz/%7Eiase/serj/SERJ2(1).pdf

Publications and Presentations: 1. Garfield, J. B. (1998). Challenges in assessing statistical reasoning. Paper presentation at the 1998 meeting of the American Educational Research Association, San Diego. 2. Garfield., J. B. (2003). Assessing statistical reasoning. Statistics Education Research Journal, Volume 2, Pages 22-38.

Population: Introductory statistics students. The instrument was developed with secondary school student samples.

Description: The SRA is comprised of 20 multiple-choice items that were designed to assess both correct and incorrect statistical reasoning.

Correct Statistical Reasoning: Areas of correct reasoning include: 1) correctly interprets probabilities, 2) understands how to select an appropriate average, 3) correctly computes probability (i.e., understands probabilities as ratios and uses combinatorial reasoning) 4) understands independence, 5) understands sampling variability, 6) distinguishes between correlation and causation, 7) correctly interprets two-way tables, and 8) understands importance of large samples.

Statistical Misconceptions: Areas of statistical misconceptions include: 1) misconceptions involving averages (i.e., averages are the most common number, fails to take outliers into consideration when computing the mean, compares groups based on their averages, and confuses mean with median), 2) outcome orientation misconception, 3) good samples have to represent a high percentage of the populations, 4) law of small numbers, 5) representativeness misconception, 6) correlation implies causation, 7) equiprobability bias, and 8) groups can only be compared if they are the same size.

Psychometric Properties: Reliability and validity evidence for the scores from the SRA have been studied via several methods. Experts reviewed item content and made suggestions for improvement during scale construction. Correlations between SRA scores and final scores, quiz totals, and project scores were low. Internal consistency of the scores was low, but most likely due to the fact that the items are measuring more than one skill. Test-retest correlations with one week between testing were .70 for correct reasoning responses and .75 for misconceptions. Cross-cultural studies compared SRA scores of Taiwanese and United States students. Although students in Taiwan had slightly higher total correct reasoning scores, both groups displayed similar patterns of correct reasoning and misconceptions. Location of Items: Items can be found in: Garfield., J. B. (2003). Assessing statistical reasoning. Statistics Education Research Journal, 2, 22-38. A copy of this article (which includes the SRA items), as well as many other items for building statistical reasoning assessments is at http://www.stat.auckland.ac.nz/%7Eiase/serj/SERJ2(1).pdf.

Hirsch, L.S., & O'Donnell, A.M. (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions.

In *Journal of Statistics Education*, *Volume 9*, *Issue 2*, Pages 1-22. Online journal article at: http://www.amstat.org/publications/jse/v9n2/hirsch.html.

Population: Post-secondary-level students

Background: The paper focuses on the concept of representativeness, a common misconception about probability (i.e., students estimate the probability of an event occurring, based upon how often it is represented in the population). For example, when asked on a test, students with this misconception would select a random sequence of heads and tails (e.g., THHTHT) as being more probable than a sequence of six tails (e.g., TTTTTT).

Description: The article provides 16 multiple-choice items that address students' misconceptions about representativeness. Fourteen of the 16 items have two parts: 1) the original multiple-choice item, and 2) a multiple-choice follow-up question, which asks students to justify their answer. The follow-up response options include a correct reasoning explanation as well as several explanations representing common misconceptions that students hold, such as “there ought to be roughly the same number of tails as heads”. An open-ended “other” response is also provided as the last response option of the follow-up questions. Students’ responses to the follow-up questions are used in scoring to verify whether or not their response to the first part of the question was correct or the result of a misconception. Scoring of the Two-Part Questions: Students receive both a correct response score and a misconception score. In order for students to receive a correct or misconception score, their answers to the two parts of the question must be congruent. For example, for a correct response to the first portion of the item to be counted as correct, students must select the response option in the second portion of the item that correctly justifies the answer that they gave to the first portion of the item. Likewise, in order for a misconception response in the first portion of the item to be counted as a misconception, students must select the justification in the follow-up portion that matches the answer they picked in the first portion of the item. Correct and misconception responses receive one point. If the students’ responses to the two parts of the questions are incongruent (e.g., selects the correct response to the first part of the question, but an incorrect justification), zero points are awarded. Possible correct and misconception scores for the two-part questions range from 0 to 14

Psychometric Properties: Item difficulty ranged from .44 to .81 for the first portion of the items and from .78 to .97 for the second portion of the items. Of the sample that was tested, the percent of students whose responses to the individual items were based upon misconceptions ranged from 5% (item 3) to 41% (item 12). A reliability estimate was .84. Scores from a group of students with more formal experience with statistics were higher than scores from a group of students with less formal experience with statistics.

Location of Items: The 16 items are listed in the appendix of the article located at: http://www.amstat.org/publications/jse/v9n2/hirsch.html.

Reston, Enriqueta (2005). The Statistical Literacy Assessment Scale (SLAS)

Paper Presentation: Reston, Enriqueta D. (2005). Assessing statistical literacy in graduate level statistics education. Paper presentation at the 55th Session of the International Statistical Institute, Sydney, Australia. A copy of the paper is available at http://www.stat.auckland.ac.nz/~iase/publications.php?show=13 and scrolling down the page to the Contributed papers portion. A copy of the SLAS is available at http://www.stat.auckland.ac.nz/~iase/cblumberg/reston.pdf

Population: Graduate students enrolled in a one-semester fundamental statistics course for teachers. The scale was developed with graduate students as the target population, but recently the scale was also administered to college statistics teachers.

Background: The author teaches graduate statistics courses for teachers at the College of Education, University of San Carlos, Cebu City, Philippines. The SLAS was developed as a classroom tool for assessing statistical literacy of graduate students who have diverse undergraduate backgrounds, but who have taken at least a 3-unit introductory statistics course in their undergraduate coursework.

Description: The SLAS is comprised of 15 items that are designed to provide a measure of adult statistical literacy along two dimensions: (1) understanding of basic statistical concepts and terminology used in the context of real-world situations and (2) understanding of claims and arguments based on data from various media outlets. For the first dimension, students are required to answer questions within a given real life problem that used statistical terms. For the second dimension, students are asked to interpret data from tables and graphs from various situations presented by newspapers, research reports, and popular advertisements. Although most of the items are based on specific examples occurring in Philippine contexts, the various skills encompassing statistical literacy which served as basis for the development of SLAS were drawn from the general literature on statistical literacy (Schield, 2000; Watson, 1997; Gal and Garfield, 1997; Rumsey, 2002). Specifically, each of the items presents a brief table, graph, or statistical conclusion, and then students are asked a close-ended question about the information presented (e.g., “Do you agree with the author’s claim?”). Items are answered using a “yes-no-cannot tell” format. If students cannot confidently state “yes” or “no”, they are asked to select “cannont tell” if they believe that more information about the assessment is necessary to understand the data being presented. In addition, students are required to justify/explain their answers using an open-ended response format.

Scoring: One point is given for each correct yes/no/cannot tell response. For the justification of the responses, a three-point rubric scale was devised for assessing the student’s level of statistical reasoning. A score of 2 was given if reasoning was correct and justified based on appropriate statistical concepts, principles, or procedures. A score of 1 was given if reasoning was partly correct but contained logical/conceptual flaws or errors. A score of 0 was given if no attempt to reason was made.

Psychometric Properties: The correlation between SLAS scores and final course grade in statistics was relatively high at 0.74. Internal consistency using Cronbach’s coefficient alpha was 0.65 while inter-rater reliability on the three-point scoring rubric was 0.86. Limited factor-analytic evidence of construct validity has been collected to date for a sample of 82 graduate students, warranting the need for increased sample size and continued work on the scale.

Sundre, D.L. Quantitative Reasoning Quotient (QRQ)

http://www.jmu.edu/assessment/ainstr.shtml (and then click on QRQ instrument)

Paper Presentation: Sundre, D.L. (2003). Assessment of quantitative reasoning to enhance educational quality. Paper presentation at the 2003 meeting of the American Educational Research Association, Chicago. Copies of this paper are at http://www.stat.auckland.ac.nz/~iase/cblumberg/sundreqrqpaper.doc (as a Word file) and http://www.stat.auckland.ac.nz/~iase/cblumberg/sundreqrqpaper.pdf (as a pdf file).

Population: Post-secondary students

Description: The Quantitative Reasoning Quotient (QRQ) is a 40 item measure that was designed to assess 11 quantitative reasoning skills and 15 quantitative misconceptions. A more detailed description is at http://www.stat.auckland.ac.nz/~iase/cblumberg/sundreobjectives.doc (as a Word file) and http://www.stat.auckland.ac.nz/~iase/cblumberg/sundreobjectives.pdf (as a pdf file). The scale is a revised version of the Statistical Reasoning Assessment (SRA; Garfield, 1998). The intention behind the QRQ was to create a quotient of correct reasoning over misconceptions.

Quantitative Reasoning: The quantitative reasoning skills include: 1) correctly interprets probabilities, 2) correctly interprets measures of central tendency, 3) understands how to select appropriate average, 4) correctly computes probability, 5) understands independence, 6) understands sampling variability, 7) distinguishes between correlation and causation, 8) correctly interprets two-way tables, 9) understands the importance of large samples, 10) understands sources of bias and error, and 11) recognizes features of good experimental design.

Quantitative Misconceptions: The quantitative misconceptions include: 1) misconceptions involving averages, 2) outcome orientation misconception, 3) good samples have to represent a high percentage of the population, 4) law of small numbers, 5) representativeness misconception, 6) correlation implies causation, 7) equiprobability bias, 8) groups can only be compared if they are of the same size, 9) failure to distinguish the difference between a sample and a population, 10) failure to consider and evaluate all of the data, 11) inability to create and evaluate fractions or percents, 12) only large effects can be considered meaningful, 13) failure to recognize potential sources of bias and error, 14) assumes that more decimal places indicate greater accuracy, and 15) inability to interpret probabilities.

Scoring: The original QRQ consisted of 43 items; however, it was later modified into a 40-item multiple-choice instrument. For each of the 40 items, a correct response is worth 2 points (80 possible points). Partial credit, 1 point, is awarded for response options that are partially correct, and no points are given for completely wrong answers. Each of the response options are written to directly address either a quantitative reasoning skill or a quantitative misconception. Scoring instructions are available for download from http://www.jmu.edu/assessment/wm_library/QRQscoring.pdf or as a Word document containing SPSS instructions at http://www.stat.auckland.ac.nz/~iase/cblumberg/sundreqrqspss.doc.

Psychometric Properties: The properties of the QRQ have been examined using scores from two large samples of post-secondary students. The original 43-item instrument was tested with sophomore students (*n*=803). Cronbach's coefficient alpha was .62. The modified 40-item instrument was tested using scores from first-year students (*n*=1,083). Cronbach's coefficient alpha was .55. Limited validity evidence has been collected to date, warranting the need for continued work on the scale.

Location of Items and Scoring Instructions: The QRQ and scoring instructions are currently available for free download at http://www.jmu.edu/assessment/ainstr.shtml (and then click on QRQ instrument or QRQ scoring) or at http://www.stat.auckland.ac.nz/~iase/cblumberg/sundreqrq.pdf (Note: Change the first item 17 to item 16.) . The original SRA items and many other items can be retrieved at the NSF-funded A.R.T.I.S.T. website at http://www.gen.umn.edu/artist/. For further information, contact the author at sundredl@jmu.edu.

ALTERNATIVE ASSESSMENT & AUTHENTIC ASSESSMENT TOOLS

Balanced Assessment. Harvard Group Balanced Assessment in Mathematics Project. Assessing mathematical understanding and skills effectively (AMUSE).

http://balancedassessment.concord.org (All in English and some in Spanish and Hebrew)

(See above under **ASSESSMENT OF SKILLS** for a detailed description)

Borough of Manhattan Community College. City University of New York Proficiency Examination

In their General Education Student Guide Brochure, the Borough of Manhattan Community College (BMCC) of the City University of New York (Cuny system) states as one of the education outcome goals "quantitative reasoning." Quoting, "students will acquire quantitative skills and the concepts and methods o mathematics to solve problems." This is part of the requirements to prepare students to take the CUNY Proficiency Examination (CPE) graduation requirement. The CPE tests require that students be able to: (a) summarize an author's argument; (b) Identify claims; © Analyze, integrate, and evaluate sources; (d) Discuss issues from their own perspective; (e) Provide support for their position; (f) Express ideas using proper written communication. Many of the aspects of the exam involve interpretation of graphs, statistics and quantitative arguments based on data. The problems are presented in a way that is likely to appear in the real life of the student and test really basic statistically and generally numerical literacy. See the examples, they provide excellent examples of statistical literacy assessment (some may call it, quantitative reasoning). Assessment like this can not be implemented unless the academic program supports it. Thus noteworthy is the attempt by the BMCC to spread quantitative literacy across the curriculum. BMCC's Match Across the Curriculum program has been very successful at bringing the quantitative reasoning to the most repote department on campus. This brings up the question: what is most effective, one course on quantitative reasoning, or quantitative reasoning across the curriculum?

Smith, G. (1998). Learning statistics by doing statistics.

In *Journal of Statistics Education, Volume 6, Issue 3, Pages 1-12*. Online journal article at: http://www.amstat.org/publications/jse/v6n3/smith.html.

Population: Post-secondary-level students

Background: The article describes the author's approach to both the teaching and assessment of statistics, which is based upon the concept of "authentic assessment" (e.g., Chance, 1997;

http://www.amstat.org/publications/jse/v5n3/chance.html; Garfield, 1993; http://www.amstat.org/publications/jse/v1n1/garfield.html). Authentic assessment involves the assessment of students' ability to apply what they have learned in class to everyday, realistic situations. It involves hands-on projects that apply the content from classroom lectures, followed by written and oral communication of the findings.

Description: The article provides 20 projects, which may be used as a supplement to homework and examinations. The author (Smith, 1998) describes the method that he used in restructuring his classes to incorporate realistic, every-day projects that students complete in small groups outside of class. The emphasis is based upon the concept that statistics involves careful thinking processes (i.e., statistical reasoning), as much as it involves mathematics (i.e., statistical methods). Additionally, it is emphasized that the communication of information, both orally and in writing, are an important component of statistical literacy. After using the projects in class, the author reported that the average (mean) midterm examination scores increased from 80.79 (s.d. = 16.00) without the projects to 92.13 (s.d. = 6.96) with the projects. Final examination scores increased from 80.27 (s.d. = 12.56) to 88.12 (s.d. = 8.28; Smith, 1998). (The classes had the same professor and textbook.) Additionally, only a few students had very low scores on the examinations when the projects were included in the course.

Scoring: Scoring rubrics are not provided.

Psychometric Properties: The psychometric properties have not been studied.

Location of Items: The 20 items/projects are listed in the appendix of the article located at: http://www.amstat.org/publications/jse/v6n3/smith.html.

ASSESSMENT OF Attitudes and Beliefs about Statistics

Schau, C., Dauphinee, T., Del Vecchio, A., Stevens, J. Survey of Attitudes Toward Statistics (SATS)

http://www.unm.edu/~cschau/downloadsats.pdf

Publications:

1. Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. Educational and Psychological Measurement, Volume 55, Pages 868 – 875.

2. Dauphinee, T. L., Schau, C., & Stevens, J. J. (1997). Survey of attitudes toward statistics: Factor structure and factorial invariance for women and men. Structural Equation Modeling, Volume 4, Pages 129 – 141.

3. Hilton, S. C., Schau, C., & Olsen, J. A. (2004). Survey of Attitudes Toward Statistics: Factor structure invariance by gender and by administration time. Structural Equation Modeling, Volume 11, Pages 92 – 109.

Population: Introductory statistics students

Description: Attitudes toward statistics is represented by four dimensions: 1) Affect—"positive and negative feeling concerning statistics" (6 items); 2) Cognitive Competence—"attitudes about intellectual knowledge and skills applied to statistics" (6 items); 3) Value—"attitudes about the usefulness, relevance, and worth of statistics in personal and professional life" (9 items); 4) Difficulty—"attitudes about the difficulty of statistics as a subject" (7 items). The scale was created to measure attitudes towards statistics at any point in time during the completion of a statistics course. This is possible due to the two versions of the scale (pre and post) that differ only in tense (e.g., "I will like statistics" versus "I like statistics"). The students respond to the 28-items using a 1 (strongly disagree) to 7 (strongly agree) response scale. Psychometric Properties: The results of three studies examining the properties of the SATS showed that the four-factor structure was supported (Schau et al, 1995) and was found to be invariant across gender (Dauphinee et al, 1997) and administration time (Hilton et al, 2004). In addition, reliability was found to be adequate across the subscales (a >.64; Schau et al, 1995), and hypothesized relationships with the ATS measure were supported (Schau et al, 1995).

Location of Items: Dauphinee, T. L., Schau, C., & Stevens, J. J. (1997). Survey of attitudes toward statistics: Factor structure and factorial invariance for women and men. Structural Equation Modeling, 4, 129 – 141.and at http://www.unm.edu/~cschau/downloadsats.pdf

Wise, S. L. Attitudes Toward Statistics (ATS)

Publication: Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. ***Educational and Psychological Measurement***, Volume 45, Pages 401 – 405.

Population: Introductory statistics students

Description: The 29-item attitudinal scale consists of two subscales: attitude toward the field of statistics (20 items) and attitude toward the course (9 items). Students are asked to respond to how they currently feel about a statement (i.e., "I feel that statistics will be useful to me in my profession") using a 1 (strongly disagree) to 5 (strongly agree) response scale.

Psychometric Properties: A two-factor solution was supported in addition to high (~.90) internal consistency for both subscale scores. The distinctiveness of the subscale scores was further supported by correlations with course grade. Specifically, the attitude toward the course subscale was a positive predictor of course grade, unlike the attitude toward the field subscale, which had a null relationship with course grade.

Location of Items: A copy of the **"Attitudes Toward Statistics"** items is at http://www.stat.auckland.ac.nz/~iase/cblumberg/wise1.pdf (in PDF format) or at http://www.stat.auckland.ac.nz/~iase/cblumberg/wise1.doc (as a Word document). An important memorandum concerning the scoring and use of these items is at http://www.stat.auckland.ac.nz/~iase/cblumberg/wise2.pdf (in PDF format) or at http://www.stat.auckland.ac.nz/~iase/cblumberg/wise2.doc (as a Word document). Contact the author at wisesl@jmu.edu for further information.

ASSESSMENT of self-efficacy related to Statistics

Finney, S. J. & Schraw, G. Current Statistics Self-efficacy (CSE)

Publication: Finney, S.J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. ***Contemporary Educational Psychology***, Volume 28, Pages 161 – 186.

Population: Introductory statistics students **Description:** Current statistics self-efficacy is defined as "confidence in one's *ability to solve* specific tasks related to statistics". Respondents are asked to rate their current confidence to complete 14 specific tasks related to statistics (e.g., "Interpret the probability value from a statistical procedure") using a 1 (no confidence at all) to 6 (complete confidence) response scale.

Psychometric Properties: Scores from the 14 items were unidimensional, demonstrated adequate reliability (α > .90) and related in hypothesized ways to theoretically-related constructs (e.g., anxiety, attitudes towards statistics).

Location of Items: Finney, S.J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses.

Contemporary Educational Psychology, Volume 28, Pages 161 – 186 or on the Internet at http://www.stat.auckland.ac.nz/~iase/cblumberg/finney2.pdf (in PDF format) or http://www.stat.auckland.ac.nz/~iase/cblumberg/finney2.doc (in Word format).

Finney, S. J. & Schraw, G. Self-efficacy to Learn Statisics (SELS)

Publication: Finney, S.J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. ***Contemporary Educational Psychology***, Volume 28, Pages 161 – 186.

Population: Introductory statistics students

Description: Self-efficacy to *learn* statistics is defined as "confidence in one's ability to learn the skills necessary to solve specific tasks related to statistics". Respondents are asked to rate their confidence in *learning* the skills necessary to complete 14 specific tasks related to statistics (i.e., "Interpret the probability value from a statistical procedure") using a 1 (no confidence at all) to 6 (complete confidence) response scale.

Psychometric Properties: Scores from the 14 items were unidimensional, demonstrated adequate reliability (a > .90) and related in hypothesized ways to theoretically-related constructs (e.g., anxiety, attitudes towards statistics).

Location of Items: Finney, S.J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, Volume 28, Pages 161 – 186 or on the Internet at finney1.pdf (in PDF format) or http://www.stat.auckland.ac.nz/~iase/cblumberg/finney1.doc (in Word format).

Back to ISLP Home

assess.txt · Last modified: 2007-08-28 20:24 by jsanchez