

# Building A Psychometric Portfolio: Evidence for Reliability, Validity, and Other Quality Indicators In Survey Data Collection

Association for Institutional Research  
50<sup>th</sup> Annual Forum  
Chicago, Illinois  
Wednesday, June 2, 2010



**Robert Gonyea**  
**Ali Korkmaz**  
**Allison BrckaLorenz**  
**Angie Miller**

# Overview

1. Introduction
2. NSSE Psychometric Portfolio
3. The Framework
  - A. Reliability
  - B. Validity
  - C. Other Quality Indicators
4. Discussion with you

# Introduction

Psychometric testing?

What is a psychometric portfolio?

Who is this for?

# NSSE Psychometric Framework

A framework for organizing and presenting studies about the quality of NSSE

Consists of three areas of analysis, each containing multiple approaches

1. Reliability
2. Validity
3. Other Quality Indicators

## Psychometric Portfolio

Validity

Reliability

Other Quality Indicators

### NSSE's Commitment to Data Quality

As part of NSSE's commitment to transparency as well as continuous improvement, we routinely assess the quality of our survey and resulting data, and we embrace our responsibility to share the results with the higher education community. This Psychometric Portfolio is a framework for presenting our studies of the reliability, validity, and other indicators of quality of NSSE's data, including analysis of data subsets defined by a variety of student and institutional characteristics. It serves higher education leaders, researchers, and professionals who have a vested interest in using NSSE data and trusting their results.

Within the portfolio, each study is described in a brief technical report that includes research questions, data, methods, results, selected references, and more. When available, complete reports and research papers are provided for studies that have been presented at conferences and published in journals and magazines.

Select the tabs above for a description of terms and links to studies organized according to validity, reliability, and other quality indicators.

Your feedback is most welcome. Please contact us at [nsse@indiana.edu](mailto:nsse@indiana.edu).

**The Psychometric Portfolio will be available on the  
NSSE website summer 2010: [www.nsse.iub.edu](http://www.nsse.iub.edu)**



## Psychometric Portfolio

[Validity](#) [Reliability](#) [Other Quality Indicators](#)

### Validity

Validity refers to "the degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of tests (Messick, 1989)." Validity is a property of the inference not the instrument. It involves empirical and theoretical support for the interpretation of the constructs. In this section we provide different forms of validity evidences for NSSE measures and items.

- Content validity
  - **Conceptual framework**
- Construct validity
  - **Deep learning scale**
- Concurrent validity
  - **BCSSE-NSSE Relationships**
- Predictive validity
  - **Connecting the dots**
  - **Retention & Cumulative credits**
- Response process
  - **Cognitive interviews**
  - **Focus groups**
- Known groups validity
  - **Group membership**
- Consequential validity
  - **NSSE Institute: Data use examples**

**The Psychometric Portfolio will be available on the  
NSSE website summer 2010: [www.nsse.iub.edu](http://www.nsse.iub.edu)**



ANNUAL RESULTS

PUBLICATIONS &  
PRESENTATIONSAGGREGATE  
TABLESPSYCHOMETRIC  
PORTFOLIOINSTITUTIONAL  
REPORTING

## Psychometric Portfolio

Validity

Reliability

Other Quality Indicators

### Reliability

Reliability refers to the consistency or stability of measurement. The reliability evidence presented here assesses the extent to which items within a scale are internally consistent or homogenous and the extent to which results are similar across periods of time or different forms of the NSSE survey. Use of a reliable instrument or scale implies that data and results are reproducible.

- Internal consistency [2009](#) , [2008](#) , [2007](#) , [2006](#)
  - Inter-item correlation tables [2009](#) , [2008](#)
- Temporal stability
  - Student level [2010](#) , [2009](#)
  - Institution level [2009](#) , [2008](#) , [2007](#)
- Equivalence
  - [How often is often](#)
  - [Co-curricular hours](#)
  - [Written amounts](#)

**The Psychometric Portfolio will be available on the  
NSSE website summer 2010: [www.nsse.iub.edu](http://www.nsse.iub.edu)**



ANNUAL RESULTS

PUBLICATIONS &  
PRESENTATIONSAGGREGATE  
TABLESPSYCHOMETRIC  
PORTFOLIOINSTITUTIONAL  
REPORTING

## Psychometric Portfolio

Validity

Reliability

Other Quality Indicators

### Other Quality Indicators

"Other Quality Indicators" includes procedures, standards, and other evaluations implemented by NSSE to reduce error and bias, and to increase the precision and rigor of the data. These studies assess NSSE's adherence to the best practices in survey design, and cover various stages of the survey, including sampling, survey administration, and reporting.

- Institution participation **2009**
- **Item bias**
- Measurement error
  - **NSSE policies compared with NCEES guidelines**
  - **Administrative processes**
- Data Quality
  - **Missing data**
  - **Item non-response**
- Mode analysis **2009**
- Nonresponse error
  - Non-respondents differ from respondents **2005**
  - **Do more engaged respond more?**
- **Sampling error**
- Social desirability **2010**

**The Psychometric Portfolio will be available on the  
NSSE website summer 2010: [www.nsse.iub.edu](http://www.nsse.iub.edu)**

# Report Template

Each brief report in the portfolio contains:

- Purpose of the analysis and research question(s)
- Data description
- Methods of analysis
- Results
- Summary
- References and additional resources

# Reliability

Reliability refers to the consistency of results

- Are results similar across different forms of the instrument or across time periods of data collection?
- Reliable instruments/scales imply that data/results are reproducible
- Strongly related to error—large amounts of error can lead to unreliable measurements
- Reliability measurements can be calculated with data from single or multiple survey administrations

# Reliability

## Heavy reliance on literature

- Framework construction
  - McMillan & Schumacher, 2001; Cook & Beckman, 2006
- Reliability standard of .70 or greater
  - Litwin, 2003; McMillan & Schumacher, 2001
- Reporting information for subpopulations
  - AERA, 2008
- Using proportions of consistent responses
  - Boote, 1981
- Assumption that reliability is independent of scale points
  - Matell & Jacoby, 1971

# Reliability

Generally four types of reliability

- **Internal Consistency**
- **Temporal Stability**
  - Also called test-retest reliability
- **Equivalence**
  - Also called inter-method or parallel forms reliability
- **Agreement**
  - Also called inter-rater reliability
  - Not relevant for NSSE

# Reliability

Example questions of NSSE reliability:

- How well do the items within the NSSE benchmarks intercorrelate?
- How stable are institutional benchmark scores over time?
- Does the NSSE survey produce similar results when administered to the same person at different times?
- Are students able to reliably estimate the number of papers they write?
- Do different versions of NSSE questions produce similar results, specifically, how often is often?

# Reliability

Results are presented

- At the student-level and/or institution-level
  - Student-level Temporal Stability
  - Institution-level Temporal Stability
- For various subgroups
  - Internal Consistency
- From original studies or previously written papers
  - Equivalence: How often is often
- From multiple years of testing
  - Reliability Framework

# Reliability

## Overview of results

- Benchmarks are highly reliable from year to year at the institution level
- NSSE benchmarks are generally reliable across various subpopulations
- Students reliably respond to vague quantifiers such as “sometimes” or “often”
- Enriching Educational Experiences is less reliable than other benchmarks

# Validity

Validity refers to “the degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of tests” (Messick, 1989).

- Validity is a property of the inference, not the instrument
- Involves empirical and theoretical support for the **interpretation** of the construct
  - Hypotheses driven (evidences collected)

# Validity

## Relied on literature

- No consensus on a framework but guided us to include different dimensions
  - American Educational Research Association, Psychological Association, & National Council on Measurement in Education. (1999), Messick (1989), Cook & Beckman(2006), Borden & Young (2008).
- Evolving concept
- Not a fixed characteristic (**depends on use, population and sample**)

# Validity

## General types

- Content validity (theory, expert reviews)
- Construct validity (factor analyses)
- Concurrent validity (relations to other variables)
- Known groups validity
- Predictive validity
- Consequential validity
- Response process (cognitive interviews, focus groups)

# Validity

Example questions of NSSE validity:

- Do BCSSE scales predict NSSE benchmarks?
- Is there a relationship between student engagement and selected measures of student success?
- Do students interpret the survey questions in the same way the authors intended?
- Do students' responses differ according to group membership in a predictable way?
- Do institutions appropriately use the survey data and results?

# Validity

## NSSE Validity Evidence

### Content validity

- Conceptual framework

### Construct validity

- Deep learning scales

### Concurrent validity

- BCSSE-NSSE relationships

### Known-group differences

### Predictive validity

- Connecting the dots
- Retention/Cumulative credits

### Response process

- Cognitive interviews
- Focus groups

### Consequential validity

- NSSE Institute: Data use examples

# Validity

## Sampling of results:

- Overall, cognitive interviews and focus groups found survey questions to be clearly worded and understandable.
- The NSSE benchmarks are able to detect differences between groups in a predictable way.
- Institutional uses of NSSE data coincide with the intended purposes of the NSSE instrument.
- Student engagement has a positive effect on first-year persistence and cumulative credits taken.

# Other Quality Indicators

“Other Quality Indicators” includes procedures, standards, and other evaluations implemented by NSSE to reduce error and bias, and to increase the precision and rigor of the data.

- Assesses NSSE’s adherence to the best practices in survey design, including sampling, survey administration, and reporting
- Related to both reliability and validity

# Other Quality Indicators

Created this additional section of the psychometric framework, because:

- We determined they were not mutually exclusive from either reliability and validity.
- We also examine administrative processes and design issues that impact survey quality.

# Other Quality Indicators

## Other Quality Indicator studies:

- Institution participation
- Item bias
- Measurement error
- Data quality (item non-response and missing data)
- Mode analysis
- Non-response error
- Sampling error
- Social desirability

# Other Quality Indicators

## Example questions:

- Do NSSE policies and practices adhere to NCES recommended standards and guidelines?
- Are institutions that participate in NSSE different from other baccalaureate-granting colleges and universities?
- Are responses to the NSSE questionnaire influenced by a tendency to respond in a socially desirable manner?

# Other Quality Indicators

## Sampling of results

- NSSE has stringent policies and procedures for data collection and reporting
- Institutions participating in NSSE are generally similar to other institutions and have an adequate number of respondents
- No strong evidence for mode bias
- Some evidence for non-response issues, such as high rate of drop-off for certain items

# Conclusion

1. This is responsible survey research.
2. This portfolio is about transparency – so we value feedback.
3. NSSE strives for continuous improvement.
4. We are very excited about the possibilities this presents (e.g., NSSE 2.0)



Robert Gonyea [rgonyea@indiana.edu](mailto:rgonyea@indiana.edu)

Ali Korkmaz [akorkmaz@indiana.edu](mailto:akorkmaz@indiana.edu)

Allison BrckaLorenz [abrckalo@indiana.edu](mailto:abrckalo@indiana.edu)

Angie Miller [anglmill@indiana.edu](mailto:anglmill@indiana.edu)

# Discussion

- Questions?
- Feedback and suggestions?

The Psychometric Portfolio will be available Summer 2010 on the NSSE web site: [www.nsse.iub.edu](http://www.nsse.iub.edu)



Robert Gonyea [rgonyea@indiana.edu](mailto:rgonyea@indiana.edu)  
Ali Korkmaz [akorkmaz@indiana.edu](mailto:akorkmaz@indiana.edu)  
Allison BrckaLorenz [abrckalo@indiana.edu](mailto:abrckalo@indiana.edu)  
Angie Miller [anglmill@indiana.edu](mailto:anglmill@indiana.edu)