**TASO Conference 2021 – survey design and validation**

**Introduction**

The following document provides a short summary of the key points from the Survey Design and Validation webinar delivered as part of the TASO Conference 2021. The document is broken down into five sections with introductory reading at the end of the document. The five sections are:

* What is a survey
* Constructs and dimensionality
* Validity
* Reliability
* Goals of measurement

**What is a survey?**

**Definitions**

* A survey is a tool used to collect primary data.
* Primary data is data that is collected directly to obtain the information needed to answer programme-specific research questions.
* Survey design includes all aspects of a survey, from the establishment of a need for data to the final output.
* Data collection is the process of gathering information and measuring the indicators.

**Use of surveys**

HE stakeholders using surveys:

* Widening participation teams
* Student success teams
* Evaluators
* Researchers.

Type of survey data collection:

* Face-to-face
* Telephone surveys
* Self-administered paper and pencil surveys
* Self-administered computer surveys (typically online).

Use of surveys in educational research:

* Testing students’ academic knowledge
* Assessing personal attributes (e.g., attitudes about one’s school, or university)
* Measuring aspirations (e.g., whether one plans to apply to university)
* Measuring personality characteristics (e.g., self-efficacy, self-esteem).

**Constructs and dimensionality**

**Measurement concepts**



*A construct* is an abstract idea, an underlying theme, or the subject matter that we wish to measure using survey questions (e.g., political party affiliation, sense of belonging, or parental engagement).

Simple constructs can be measured using one of two questions whilst more complex ones may require more questions to be fully operationalised.

*An indicator* is the direct answer to the construct, helping us find a way to measure a construct (e.g., hours helping kids doing homework as an indicator for parental engagement at home).

Often, there are various indicators for one construct, so it is important to find one that is as closely matched to the construct of interest as possible.

**Dimensionality**

Dimensionality is a concept that helps us to understand how to accurately measure the constructs of interest.

A unidimensional construct is expected to have a single underlying dimension and can be measured using a single measure or test (e.g., self-esteem, depression, IQ).

A unidimensional scale may include multiple items, but all these items attempt to measure the same underlying dimension.



A multidimensional construct consists of two or more underlying dimensions (e.g., personality).

Multidimensional scales employ different items or tests to measure each dimension of the construct separately. The scores of each dimension are combined to create an overall measure of the construct.



**Validity**

**Overview**

Validity is the relationship between a concept and an indicator. The closer these elements are, the higher the validity is.

Validity measures whether the survey items, or the indicators, are actually measuring the construct of interest.

Since social constructs can be difficult to translate into measurable data, validity is never assured but we can try to maximise it.

**Four types of validity**



* Construct validity indicates the capacity of an indicator or survey scale items to map the construct of interest. In other words, it tells us how closely related the indicator is to the underlying concept.
* Content validity tells us how representative an indicator or a test is, relative to all aspects of the construct.
* Face validity indicates the surface-level suitability of the content of a survey or a test.
* Criterion validity has two subtypes of validity that we are interested in:
	+ Predictive validity tells us about the capacity of the indicators or survey items to predict later performance on a related criterion.
	+ Concurrent validity indicates whether the new measure being developed relates to an existing measure.

**Reliability**

**Overview**

Reliability is the relationship between an indicator and the final data obtained. It assesses whether the indicator is measured consistently and precisely.

Often, the measures obtained do not match perfectly, and differences between the measurements and the reliability of your measure is more up for question.

**Three types of reliability**



* Internal consistency is the consistency of people’s responses across the items, on a multiple-item measure.
* Inter-rater reliability is the extent to which different researchers are consistent in their judgment.
* Test-retest reliability is the extent to which the scores obtained on a test or a survey are consistent over time.

**Goals of measurement**

When using surveys are a measurement tool, we aim for high reliability and high validity, the centre of the bullseye – the bottom right dartboard.

We may end up with sub-optimal measures:

* High validity and low reliability – top right dartboard.
* Low validity and high reliability – bottom left dartboard.

The worst case scenario:

* Low validity and low reliability – top left dartboard. 

**Introductory reading list**

**On survey design**

Maul, A. (2017). Rethinking traditional methods of survey validation*. Measurement: Interdisciplinary Research and Perspectives*, 15(2), 51-69.

Rossi, P. H., Wright, J. D., & Anderson, A. B. (Eds.). (2013*). Handbook of survey research*. Academic Press.

SAGE videos. (2018, October 2). *Designing a Survey* [Video]. YouTube. <https://www.youtube.com/watch?v=mdVWbuffdNY>

**On surveys in the HE sector**

Berends, M. (2006). Survey Methods in Educational Research. In American Educational Research Association, 2004, San Diego, CA, US; *Lawrence Erlbaum Associates Publishers*.

Cole, D., Kitchen, J. A., & Kezar, A. (2019). Examining a comprehensive college transition program: An account of iterative mixed methods longitudinal survey design*. Research in Higher Education*, 60(3), 392-413.

Klemenčič M., Chirikov I. (2015) How Do We Know How Students Experience Higher Education? On the Use of Student Surveys. In: Curaj A., Matei L., Pricopie R., Salmi J., Scott P. (eds) *The European Higher Education Area. Springer, Cham*.

Universities UK. (2016). *Student experience: measuring expectations and outcomes*. <https://core.ac.uk/download/pdf/74381689.pdf>

Walston, J., Redford, J., & Bhatt, M. (2017). *Workshop on Survey Methods in Education Research: Facilitator’s guide and resources.* U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. <https://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_2017214.pdf>

**On reliability and validity**

Al-Hemyari, Z. A., & Al-Sarmi, A. M. (2016). Validity and Reliability of Students and Academic Staff’s Surveys to Improve Higher Education. *Educational Alternatives, Journal of International Scientific Publications*, 14, 242-263.

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*, *22*(4), 195.

Dr. Cris Wildermuth. (2016, July 10). *Validity & reliability* [Video]. YouTube. <https://www.youtube.com/watch?v=F6LGa8jsdjo>

MeanThat. (2016, March 17). *3.11 Validity and Reliability Of Research* [Video]. YouTube. <https://www.youtube.com/watch?v=2fK1ClycBTM>

Prof Nikki Hozack. (2015, March 7). Improving Measurement Reliability [Video]. YouTube. <https://www.youtube.com/watch?v=_1tLkRmQbuU>

Prof Nikki Hozack. (2015, March 7). Reliability VS Validity [Video]. YouTube. <https://www.youtube.com/watch?v=tMERBdjs8Dw>

Prof. Nikki Hozack. (2015, March 7). Understanding Measurement Validity [Video]. YouTube. <https://www.youtube.com/watch?v=kkjjZtFV9ZE>

Quant Psych. (2019, September 25). The four types of validities [Video]. YouTube. <https://www.youtube.com/watch?v=a0wZTDiaUs8>

Zhao, J., & Gallant, D. J. (2012). Student evaluation of instruction in higher education: Exploring issues of validity and reliability*. Assessment & Evaluation in Higher Education*, 37(2), 227-235.