

Data Quality Dimensions

Dimension	Definition	Examples
Validity	<p>The extent to which an evaluation or instrument really measures what it is intended to measure. Validity applies to the evaluation design, methods, and data. There are two main types of validity: internal validity and external validity.</p> <ul style="list-style-type: none"> • Internal validity: The extent to which the study or instrument measures a construct accurately and is free of alternative explanations. • External validity: The extent to which an instrument or evaluation can be generalized to different contexts, such as other populations or settings. 	<p>Internal validity: A survey that is intended to measure pedagogical expertise measures an educator’s pedagogical skills and knowledge and does not measure other topics such as motivation or subject content knowledge.</p> <p>External validity: A survey that is intended to measure pedagogical expertise can be used in a variety of settings (e.g., rural or urban) and populations (e.g., middle school or high school teachers).</p>
Reliability	<p>Reliability is the extent to which a data source yields consistent results. There are three main types of reliability:</p> <ul style="list-style-type: none"> • Internal consistency: The extent to which a group of items on a data source consistently measure the same topic. • Inter-rater reliability: The extent to which multiple raters or observers are consistent in coding or scoring. • Test–retest reliability: The extent to which an individual would receive the same score if tested twice on the same assessment. 	<p>Internal consistency: Survey participants who respond positively to certain items about educator professionalism also respond positively to other related items on the survey.</p> <p>Inter-rater reliability: Two principals evaluating the same teacher indicate a similar rating independently.</p> <p>Test–retest reliability: A diagnostic screening test (with multiple forms) yields the same results if administered to the same individual multiple times.</p>
Timeliness	<p>Timely data are current, and the results of data analysis and interpretation are available when needed.</p>	<p>If the goal of a program is to improve students’ academic motivation upon middle school entry, data are collected when students enter middle school, not when they leave middle school, to</p>

Dimension	Definition	Examples
		determine the extent to which they are motivated.
Comprehensiveness	The data collected in an evaluation include sufficient details or contextual information and can therefore be meaningfully interpreted.	If the evaluation focuses on differences in math performance among boys and girls, the data collected include both students' math scores and their gender.
Trustworthiness	Trustworthiness is the extent to which data are free from manipulation and entry error. This is often addressed by training data collectors.	An evaluation team trains data collectors to ensure that there is no opportunity for participants to answer questions in a biased way.
Completeness	Data are collected from all individuals in the sample and are sufficient to answer the evaluation questions. Completeness also relates to the degree of missing data and the generalizability of the results to other contexts.	If the goal is to evaluate the use and implementation of a new K–12 math curriculum in a local district, then all math educators and their classes within that district are included.

Note. Adapted from *Data Quality Assurance Tool for Program-Level Indicators*, by W. Brown, R. Stouffer, and K. Hardee, 2007, MEASURE Evaluation (<https://www.measureevaluation.org/resources/publications/ms-07-19>).

This handout was prepared under Contract ED-IES-17-C-0005 by Regional Educational Laboratory Central, administered by Marzano Research. The content does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.