

TOOLS OF RESEARCH: RATING SCALE, ATTITUDE SCALE, QUESTIONNAIRE, APTITUDE TEST AND ACHIEVEMENT TEST, INVENTORY

Some measurement is very straightforward, using a single indicator to represent a variable. For example, you could measure a person's educational background by asking about the highest grade he or she had completed. Similarly, such variables as grade level, nationality, marital status, or number of children could be measured by a single indicator simply because these variables refer to phenomena that are very clear and for which a single indicator provides an acceptable measure. Other variables, however, are more complex and much more difficult to measure. In these cases, using a single indicator is not appropriate.

Selecting appropriate and useful measuring instruments is critical to the success of any research study. One must select or develop scales and instruments that can measure complex constructs such as intelligence, achievement, personality, motivation, attitudes, aptitudes, interests, and self-concept. Of all the types of measuring instruments available, cognitive, affective, and projective tests are the most commonly used in educational research.

TESTS

Tests are valuable measuring instruments for educational research. A test is a set of stimuli presented to an individual in order to elicit responses on the basis of which a numerical score can be assigned. This score, based on a representative sample of the individual's behavior, is an indicator of the extent to which the subject has the characteristic being measured. The utility of these scores as indicators of the construct of interest is in large part a function of the objectivity, validity, and reliability of the tests. Objectivity is the extent of agreement among scorers. Some tests, such as multiple-choice and true-false tests, are described as objective because the scoring is done by comparing students' answers with the scoring key, and scorers need make no decisions. Essay tests are less objective because scores are influenced by the judgment and opinions of the scorers. In general, validity is the extent to which a test measures what it claims to measure. Reliability is the extent to which the test measures accurately and consistently.

Cognitive Tests

A cognitive test measures intellectual processes such as: - thinking, memorizing, problem solving, analyzing, reasoning, and applying information. Most tests that school pupils take are cognitive achievement tests.

Achievement Tests

An achievement test measures an individual's current proficiency in given areas of knowledge or skill. Typically administered in school settings, achievement tests are designed to provide information about how well test takers have learned the material introduced in school. The tests are standardized, and an individual's performance is usually determined by comparing it to the norm, the performance of a national group of students in the individual's grade or age level who

took the same test. Thus, these tests can provide comparisons of a given student to similar students nationally.

Standardized Tests

Standardized tests are published tests that have resulted from careful and skillful preparation by experts and cover broad academic objectives common to the majority of school systems. Standardized achievement tests typically cover a number of different curriculum areas, such as reading, vocabulary, language, and mathematics. A standardized test that measures achievement in several curriculum areas is called a test battery, and the assessment of each area is done with a subtest. These are tests for which comparative norms have been derived, their validity and reliability established, and directions for administering and scoring prescribed. The directions are contained in the manuals provided by the test publishers. To establish the norms for these tests, their originators administer them to a relevant and representative sample. The norm group may be chosen to represent the nation as a whole or the state, city, district or local school.

Researcher-Made Tests

When using standardized tests of achievement is not deemed suitable for the specific objectives of a research study, research workers may construct their own tests. It is much better to construct your own test than to use an inappropriate standardized one just because it is available. The advantage of a researcher-made test is that it can be tailored to be content specific; that is, it will match more closely the content that was covered in the classroom or in the research study.

On the basis of the type of interpretation made, standardized and teacher-made tests may be further classified as norm-referenced or criterion-referenced.

Norm-Referenced and Criterion-Referenced Tests

Norm-referenced tests permit researchers to compare individuals' performance on the test to the performance of other individuals. An individual's performance is interpreted in terms of his or her relative position in a specified reference group known as the normative group. Typically, standardized tests are norm referenced, reporting performance in terms of percentiles, standard scores, and similar measures.

In contrast, criterion-referenced tests enable researchers to describe what a specific individual can do, without reference to the performance of others. Performance is reported in terms of the level of mastery of some well-defined content or skill domain. Typically, the level of mastery is indicated by the percentage of items answered correctly. A well-known standardized instrument, the National Assessment of Educational Progress (NAEP) is criterion referenced.

In norm-referenced tests, items are selected that will yield a wide range of scores. A researcher must be concerned with the range of difficulty of the items and the power of the items to discriminate among individuals. In criterion-referenced tests, items are selected solely on the basis of how well they measure a specific set of instructional objectives. They may be easy or difficult, depending on what is being measured. The major concern is to have a representative sample of items measuring the stated objectives so that individual performance can be described directly in terms of the specific knowledge and skills that these people are able to achieve.

Performance Assessments

Another way to classify achievement tests is whether they are verbal or performance tests. The most common achievement tests are paper-and-pencil tests measuring cognitive objectives. This familiar format, usually administered to groups, requires individuals to compose answers or choose responses on a printed sheet. In some cases, however, a researcher may want to measure performance— what an individual can *do* rather than what he or she *knows*. Performance assessment, usually administered individually, is a popular alternative to traditional paper-and-pencil tests among educators. A performance test is a technique in which a researcher directly observes and assesses an individual's performance of a certain task and/or judges the finished product of that performance. Performance tests are useful for measuring abilities and skills that cannot be measured by paper-and-pencil tests. However, they are time intensive and thus more expensive to administer and score.

APTITUDE TESTS

Aptitude tests differ from achievement tests in that aptitude tests attempt to measure general ability or potential for learning a body of knowledge and skills, whereas achievement tests attempt to measure the actual extent of acquired knowledge and skills in specific areas. Aptitude tests measure a subject's ability to perceive relationships, solve problems, and apply knowledge in a variety of contexts. Some critics question the distinction made between aptitude and achievement tests. They point out that an aptitude test measures achievement to some extent, and an achievement test has an aptitude element. Aptitude tests were formerly referred to as intelligence tests, but the latter term has declined in use because of controversy over the definition of intelligence and because people tend to associate intelligence with inherited ability. Aptitude tests should not be considered as measures of innate (or "pure") intelligence.

Researchers often use aptitude tests. Aptitude or intelligence is frequently a variable that needs to be controlled in educational experiments. To control this variable, the researcher may use the scores from a scholastic aptitude test. Tests of general aptitude are also referred to as scholastic aptitude tests and tests of general mental ability. Unlike an achievement test, which is used to assess what individuals have learned, an aptitude test is commonly used to predict how well an individual is likely to perform in a future situation. Aptitude tests can be administered to groups, or they can be individually administered.

Affective Tests

An affective test is an assessment designed to measure affective characteristics—mental characteristics related to emotion, such as attitude, interest, and value. Affective tests are often used in educational research and exist in many different formats. Most are non-projective; that is, they are self-report measures in which the test taker responds to a series of questions or statements about him- or her.

Instruments that examine values, attitudes, interests, and personalities tap the test takers' emotions and perceptions. Values are deeply held beliefs about ideas, persons, or objects. Attitudes indicate our favorable or unfavorable feelings; they reflect our tendencies to accept or reject groups, ideas or objects. Interests indicate the degree to which we seek out or desire to participate in particular

activities, objects, and ideas. Personality is made up of a number of characteristics that represent a person's typical behaviors; it describes what we do in our natural life circumstances.

ATTITUDE SCALES

An attitude scale is an instrument that measures what an individual believes, perceives, or feels about self, others, activities, institutions, or situations. Five basic types of scales are used to measure attitudes: Likert scales, semantic differential scales, rating scales, Thurstone scales, and Guttman scales. The first three are frequently used in educational research.

Likert Scales

A Likert scale requires an individual to respond to a series of statements by indicating whether he or she strongly agrees (SA), agrees (A), is undecided (U), disagrees (D), or strongly disagrees (SD). Each response is assigned a point value, and an individual's score is determined by adding the point values of all the statements. For example, the following point values are typically assigned to positive statements: SA = 5, A = 4, U = 3, D = 2, SD = 1. For negative statements, the point values would be reversed—that is, SA = 1, A = 2, U = 3, D = 4, and SD = 5.

Semantic Differential Scales

A semantic differential scale requires an individual to indicate his or her attitude about a topic (e.g., property taxes) by selecting a position on a continuum that ranges from one bipolar adjective (e.g., fair) to another (e.g., unfair). Each position on the continuum has an associated score value. This scale is typical of semantic differential scales, which usually have 5 to 7 intervals with a neutral attitude assigned a score value of 0. A person who checks the first interval (i.e., a score of 3) on each of these items has a very positive attitude toward property taxes. Totaling the score values for all items yields an overall score. Usually, summed scores (i.e., interval data) are used in statistical data analysis.

RATING SCALES

A rating scale may also be used to measure a respondent's attitudes toward self, others, activities, institutions, or situations. One form of rating scale provides descriptions of performance or preference and requires the individual to check the most appropriate description. A second type of rating scale asks the individual to rate performance or preference using a numerical scale similar to a Likert scale.

Circle the number that best describes the degree to which you state lesson objectives and give an overview before teaching a lesson.

5 = always

4 = almost always

3 = about half the time

2 = rarely

1 = never

1 2 3 4 5

Likert, semantic differential, and rating scales are similar, requiring the respondent to self-report along a continuum of choices. However, in certain situations—such as observing performance or judging teaching competence—Likert, semantic differential, and rating scales can be used by others (e.g., a researcher, a principal, a colleague) to collect information about study participants. For example, in some studies it may be best to have the principal, rather than the teacher, use a Likert, semantic differential, or rating scale to collect data about that teacher.

Thurstone and Guttman Scales

A Thurstone scale requires participants to select from a list of statements that represent different points of view on a topic. Each item has an associated point value between 1 and 11; point values for each item are determined by averaging the values of the items assigned by a number of judges. An individual's attitude score is the average point value of all the statements checked by that individual.

A Guttman scale also requires respondents to agree or disagree with a number of statements; it is then used to determine whether an attitude is uni-dimensional. An attitude is uni-dimensional if it produces a cumulative scale in which an individual who agrees with a given statement also agrees with all related preceding statements. For example, if you agree with Statement 3, you also agree with Statements 2 and 1.

INVENTORY

Interest Inventories

An interest inventory requires participants to indicate personal likes and dislikes, such as the kinds of activities they prefer. The respondent's pattern of interest is compared to the interest patterns of others. For example, for an occupational interest inventory, responses are compared to those typical of successful persons in various occupational fields. Interest inventories are widely used in this way to suggest the fields in which respondents may be most happy and successful.

Personality Inventories

A personality inventory includes questions or statements that describe behaviors characteristic of certain personality traits. Respondents indicate whether or to what degree each statement describes them.

Some inventories are presented as checklists; respondents simply check items they feel characterize them. An individual's score is based on the number of responses characteristic of the trait being measured.

General personality inventories frequently used in educational research studies include the Personality Adjective Checklist, California Psychological Inventory, Minnesota Multiphasic Personality Inventory, Mooney Problem Checklist, Myers-Briggs Type Indicator, and the Sixteen Personality Factors Questionnaire.

QUESTIONNAIRE

A questionnaire is a form prepared and distributed to secure responses to certain questions. It is a device for securing answers to questions by using a form which the respondent fills by himself. It is a systematic compilation of questions that are submitted to a sampling of population from which information is desired.

Questionnaires rely on written information supplied directly by people in response to questions. The information from questionnaires tends to fall into two broad categories – facts and opinions. It is worth stressing that, in practice, questionnaires are very likely to include questions about both facts and opinions.

Purpose

The purpose of the questionnaire is to gather information from widely scattered sources. It is mostly used in cases where one can not readily see personally all of the people from whom he desires responses. It is also used where there is no particular reason to see them personally.

Types

Questionnaire can be of various types on the basis of its preparation. They are like:

- Structured v/s Non Structured
- Closed v/s Open
- Fact v/s Opinion

Structured v/s Non-Structured Questionnaire

The structured questionnaire contains definite, concrete and directed questions; where as non-structured questionnaire is often used in interview and guide. It may consist of partially completed questions.

Closed v/s Open Questionnaire

The question that calls for short check responses are known as restricted or closed form type. For Example, they provide for marking a yes or no, a short response or checking an item from a list of responses. Here the respondent is not free to write of his own, he has to select from the selected from the supplied responses. On the other hand, in case of open ended questionnaire, the respondent is free to response in his own words. Many questionnaires also included both close and open type questions. The researcher selects the type of questionnaire according to his need of the study.

Fact v/s Opinion

In case of fact questionnaire, the respondent is expected to give information of facts without any reference to his opinion or attitude about them. But in case of opinion questionnaire the respondent gives the information about the facts with his own opinion and attitude.

Characteristics of a Good Questionnaire

Questionnaire should deal with important or significant topic to create interest among respondents.

- It should seek only that data which cannot be obtained from other sources.
- It should be as short as possible but should be comprehensive.
- It should be attractive.
- Directions should be clear and complete.
- It should be represented in good Psychological order proceeding from general to more specific responses.
- Double negatives in questions should be avoided.
- Putting two questions in one question also should be avoided.
- It should avoid annoying or embarrassing questions.
- It should be designed to collect information which can be used subsequently as data for analysis.
- It should consist of a written list of questions.
- The questionnaire should also be used appropriately.

Suggested Reading

Ary, D., Jacobs, L. C., Sorensen, C., and Razavieh, A. (2010), Introduction to Research in Education, Wadsworth Cengage Learning, Canada.

Best, J. W. and Kahn, J. V. (1995), Research in Education, Prentice Hall, New Delhi.

Cohen, L., Manison, L., and Morrison, K. (2018), Research Methods in Education, Routledge, London and New York.

Creswell, J. W. (2012), Educational Research Planning, Conducting, and Evaluating Quantitative and Qualitative Research, Pearson, New York.

Gay, L. R., Mills, G.E., and Airasian, P. W. (2012), Educational Research Competencies for Analysis and Application, Pearson, New York.

Kerlinger, Fred. N. (1978), Foundations of Behavioral Research, New York University.

Kothari, R.C. (2004), Research Methodology, New Delhi, New Age International (P) Limited, Publishers.

Langenbach, M., Vaughn, C., and Aagaard, L. (1993), An Introduction to Educational Research, Allyn and Bacon, Boston.