3

Measurement and scaling

1 Introduction

The first step in our research model is the theoretical preparation or formulation of the research topic — a step which is as complex as it is diverse and pluralistic. This step will establish the foundations for the remaining part of the study and is therefore very important, and deserves special consideration. However, before we commence our discussion of this step, we need first to consider another more general and equally fundamental element of the research process and one of the principles of social research, namely measurement. You would remember that in the the first chapter we stated that research is expected to be systematic, accurate and precise. Measurement is a central element of social research and also fundamental for the procedures which will be introduced in the next chapter, where the preparation step of the research is considered.

In this chapter, we shall first introduce the concept, nature and types of measurement and a number of issues associated with it. We shall look at measurement as an element of social research, and explore some of its qualities, such as validity and reliability. Following this we shall turn our attention to scaling.

2 Nature of measurement

Social research, irrespective of type and nature, entails a degree of measurement. In some cases measurement is exact, quantitative and complicated. In other cases it is qualitative, involving simple operations and aiming at classification or labelling of variables. Measurement is pluralistic and can be accomplished in a number of ways.

Measurement involves categorising and/or assigning values to the variables in question, and can be diverse in nature and level of operation. Generally, measurement is quantitative or qualitative. *Quantitative measurement* concentrates on numerical values and attributes. Qualitative measurement refers to labels, names and qualities.

Qualitative measurement describes attributes by using common concepts or symbols or introducing new ones; a common procedure involves description

of categories and classifications. The classification of Australian families for instance into Anglo-Saxon and ethnic families is a qualitative measurement. In the view of some writers, qualitative measurement does not qualify as 'measurement' since it does not demonstrate the main criteria of measurement, such as precision, reliability and validity; rather, it is thought to be a process of labelling, classification and description. Nevertheless, as we shall see soon, this view is not correct.

3 Variables

Definitions
Measurement relates to variables. A variable is a concept that can take two or more values; for example, sex (male, female), marital status (single, married, divorced, widowed, deserted), age and education are variables. Variables can be 'dependent' or 'independent'. An *independent variable* causes changes in another; a *dependent variable* is a variable that is affected or explained by another variable. For instance, in a research study of 'family status and scholastic achievement', the independent variable can be family status, and the dependent variable scholastic achievement.

Discrete and continuous variables
Variables can also be discrete or continuous. Discrete and continuous variables differ from each other in terms of *scale continuity;* the former are not continuous but use whole units only, while the latter are continuous and can be fractioned indefinitely. In *discrete variables* (also called *binomial variables*), measurement uses whole units or numbers, with no possible values between adjacent units. For instance, family size can be 3, 4, 5 and so on, not 2.5, 3.8, or 5.4. Family size is a discrete variable. However, weight is a *continuous variable*; it can use smaller increments of units, for example it can be 73.2, 78.1 or 85.6 kg. Discrete variables are counted, not measured; continuous variables are measured, not counted. Examples of discrete variables are ethnicity, race, sex, marital status, cause of death or blood type. Examples of continuous variables are height, distance, time, age, temperature or IQ scores.

More definitions
Some writers also differentiate between quantitative and qualitative variables. *Qualitative variables* use nominal scale measurement; racial origin, ethnic origin, religious affiliation or sex are qualitative variables. *Quantitative variables* use either ordinal or metric scales.

4 Levels of measurement

Measurement can be performed at four levels, which vary in many ways, but especially with regard to the degree to which they match the characteristics of

the real-number system. The four levels of measurement and four corresponding scales are: the nominal, the ordinal, the interval and the ratio level. Nominal-level measurement has the lowest and ratio-level measurement the highest match with the real-number system.

a Nominal-level measurement

This is the simplest, the lowest and the most primitive type of 'measurement'. At this level, measurement involves classification of events into categories that must be distinct, unidimensional, mutually exclusive and exhaustive; and the resulting scales are 'naming' scales. Such a measure indicates that there is a difference between the categories considered.

Such differences refer to nature but not to magnitude. Thus, dormitory no. 10 is not twice as large as dormitory no. 5. In a similar fashion, numbers assigned to categories have no mathematical meaning, are used only for identification and cannot be added, subtracted, multiplied, divided or otherwise manipulated mathematically. Classifying the respondents in categories such as male–female, black–white, young–old, single, married, cohabiting, separated, divorced, remarried or widowed, or Catholic, Protestant, Anglican, Orthodox is based on nominal measurement. Classifying respondents according to their place of birth, religious affiliation, political affiliation, car type and place of residence are additional examples. Further examples of nominal measurement are: nationality, type of shoes, skin colour, type of music and brands of drinks.

Characteristic of the nominal-level measurement is that:

- it involves nominal categories and is essentially a qualitative and a non-mathematical measurement; it actually names and classifies data into categories;
- it does not have a zero point;
- it cannot be ordered in a continuum of low–high;
- it produces nominal or categorical data;
- it assumes no equal units of measurement;
- it assumes the principle of equivalence: all units of a particular group are taken to be the same. All Greeks, Aborigines, jogging shoes, IBM computers, personal vehicles, etc. are the same.

It must be noted that only statistical measures designed for nominal data must be used.

b Ordinal-level measurement

Measurement at the ordinal level involves not only categorising elements into groups but also ordering of data and ranking of variables in a continuum ranging according to magnitude, that is, from the lowest to the highest point (transitive relationship). Here, numbers offer more information since they not only indicate differences between categories but they also rank them; however, they do not allow mathematical operations such as addition or subtraction. Characteristic of this level of measurement is that:

- it refers to ranks based on a clear order of magnitude of low and high signifying that some elements have more value than others;
- the numbers have actual mathematical meaning as well as having identification properties;
- it is essentially a quantitative measurement;
- it shows a relative order of magnitude.

With regard to the last point, order of magnitude allows categories to be ranked (who is first, second, last) but does not indicate the amount of difference between the groups (how much above or below a certain category neighbouring categories are). So the difference between the first and second may be different from that between the sixth and seventh category. The intervals between the categories are not necessarily equal.

Examples of such forms of continuum employed in ordinal measurement are: *status* — low, middle, high; *size* — smallest, small, big, biggest; *quality* — poor, good, very good, excellent; *class* — low, middle, high; *achievement at school* — poor, moderate, high; *income* — low, middle, high. Ranking occupations is another example.

Only nominal and ordinal statistical measures are permissible for ordinal data.

c Interval-level measurement

This level of measurement, as well as demonstrating the properties of ordinal-level measurement, provides information about the distance between the values, and contains equal intervals, ordering subjects into one of them. This method allows the researcher to judge differences between respondents and to obtain more detailed information about the research topic.

Characteristic of this level of measurement is that it includes equal units, and that it is essentially a quantitative measurement. In addition to allowing differentiation and classification and also incorporating orderings, it specifies the numerical distance between the categories. In other words, interval measurement allows the researcher to determine whether two values are the same or different (as in nominal measurement), whether the one is greater or smaller than the other (as in ordinal measurement) and the degree of difference between them. Nevertheless, it does not have a true zero point, and if a zero is used it is set arbitrarily, is done so for convenience and does not mean absence of the variable.

For example, if the IQ of two students is 105 and 125 respectively, in nominal terms this means that they have a different IQ; in ordinal terms that the first student has a lower IQ than the second; and in interval terms, that the IQ of the second student is 20 points higher than that of the first student, but not, say, one-fifth smarter than the other student.

In mathematical terms, at this level numbers assigned to categories are used to count and rank, but can also be added to and subtracted from each other. This indicates that interval-level measurement is superior to the other two. However, given that there is no true zero, they cannot be multiplied or divided. Statistical measures for nominal, rank and interval data can be used.

Examples of this type of measurement are degrees of temperature, calendar time (day, week, month), attitude scales and IQ scores.

d Ratio-level measurement

Measurement at this level includes all the attributes the other three forms offer, plus the option of an absolute true zero (0) as its lowest value, which in essence indicates absence of the variable in question. This allows the researcher to make statements about proportions and ratios, that is, to relate one value to another. For instance, a comparison of speed of response of two students to a stimulus, say, 10 seconds and 20 seconds, allows the researcher to conclude that the first is twice as fast as the second.

In the social sciences this level of measurement is employed mainly when measuring demographic variables; however, it is considered inappropriate for measuring attitudes and opinions. This is because indicating a 0 option in an attitude scale means no attitude, or no opinion; given the nature of the research question, this is incorrect since even stating that the respondent has no opinion is in itself an opinion.

In terms of mathematics, numbers arrived at through ratio ordering indicate counting as well as ranking, and can also be added, subtracted, multiplied or divided. As far as statistical tests are concerned, measures appropriate for all levels can be used.

Examples of this type of measurement are those given for the interval level above, with the addition of a 0 point in the continuum. Other examples could, for instance, come from the following areas: number of family members, weight, length, distance, number of books that subjects own, reaction time and number of products produced per hour.

e Measuring variables

Variables are not by nature measured at one specific level only. Whether a variable will be measured one way or another depends very much on how it is conceptualised and on the type of indicators used during measurement. The same variable can be measured in a variety of ways. Age, for instance, can be measured nominally, if it is defined in broad and discrete categories, such as infancy, adolescence, adulthood, middle age and old age; or as young and old. It can be measured also at the ordinal level, when respondents are ranked according to age from the oldest to the youngest. Age can also be measured at the interval level, given that units are equal and we can determine how many units of difference there is between age levels. Interval-level measurement tells us here not only whose age is higher (as in ordinal-level measurement) but also how much higher it is. Age can, finally, be measured at the ratio level, since it has an absolute (non-arbitrary) zero. One cannot be younger than 0; and a 20-year-old person is twice as old as a 10-year-old person.

Table 3.1 Levels of measurement: a summary

Criteria	*Nominal*	*Ordinal*	*Interval*	*Ratio*
Properties of measurement	Naming	Naming and ranking	Naming, ranking and equal intervals	Naming, ranking, equal intervals and zero point
Nature of measurement	Categorical	Ranking	Scoring	Scoring
Mathematical functions	None	None	Addition and subtraction	All four functions
Relevant statistical tests	Lambda test χ^2 test; CI%	Spearman's ρ Mann–Whitney *U* test Sign test	Pearson's *r* *t*-test ANOVA	Pearson's *r* *t*-test ANOVA
Nature of underlying construct	Discrete	Discrete or continuous	Continuous	Continuous
Examples	Marital status, gender, race, birth place, residence, ethnicity	Income status, achievement, social class, size	Temperature, calendar time, IQ scores, attitude scales	Length, Weight, Distance, No. of children, Age
Typical answers to questions	Male, female single, married, divorced, widowed	Very high, high, moderate, low, very low	Scores Likert scales Degrees	Years Kilograms Kilometres

Despite this degree of freedom researchers enjoy when measuring variables, there is a rule of thumb which states that *variables are measured at the highest level possible.* Overall, discrete variables are measured at the nominal or ordinal level, and continuous variables at the interval or ratio level.

Arbitrary and true zeros

The use of true zero as the distinguishing characteristic of ratio scales has caused some confusion. This is due to the fact that students find it difficult to distinguish true zeros from arbitrary zeros. Zeros are not always *true.* True zeros are meaningful, arbitrary zeros are not. For instance, when we measure temperature, a zero degree reading does not mean no temperature at all! And measuring attitudes, a zero does not mean no attitude at all (having no opinion on an issue is an opinion!). These zeros are not *true zeros,* they are arbitrary zeros. However, when measuring income, number of cars or number of children, a zero indicates no presence of these criteria: it means no children, no income, no cars, etc. These are true zeros; and only measurement based on using these true zeros is conducted at the ratio level.

f Summarising

All levels of measurement are effective and useful in their own context and the purpose for which they have been developed. However, nominal-level measures are the least precise, followed by ordinal-level measures, and then interval-level measures, with ratio-level measures offering the highest degree of precision. Measuring at the interval and ratio level has many advantages, but not all variables can be measured at these levels (Wang and Mahoney, 1991).

Measurement is a very important and also relatively complicated process, but it is associated with many problems and errors. It is limited by the nature of our social world and the variables in question, the perceptions of researchers and their personal bias. Its value depends on the accuracy of the instruments used and the model of operationalisation employed.

5 Validity

a Validity in quantitative research

As stated in the first chapter, attainment of validity is one of the basic principles of social research. Validity means the ability to produce findings that are in agreement with theoretical or conceptual values; in other words to produce accurate results and to measure what is supposed to be measured. If an instrument employed to measure the extent of cheating in examinations revealed that 32 per cent of the students regularly cheat, the measure used has validity if the proportion of students who cheat is actually 32 per cent. A valid measure produces true results that reflect the true situation and conditions of the environment it is supposed to study.

There are two ways of checking the validity of an instrument: empirical validation and theoretical validation. In the former, the validity of a measure is checked against empirical evidence. In the latter, the validity of an instrument is ascertained through theoretical or conceptual constructs. In both cases, validity is claimed if the findings produced through the measure in question are supported by empirical evidence or by theoretical principles.

Empirical validation

Empirical validation tests *pragmatic* or *criterion validity*. If an instrument has, for instance, produced results indicating that students involved in student union activities do better in their exams, and if this is supported by available data, the instrument in question has pragmatic validity. Again, validity here is assumed if the findings are supported by already existing empirical evidence. In this case the validity is *concurrent validity.*

Quite often, the validity of a measure is checked by the degree to which predictions made by the results of this measure are supported by findings that appear later. Validity is then claimed if new data support the predictions of the measure in question. For example, if a study found that an eventual introduction of advanced statistics into the social sciences degree would