Validity Assessment in Qualitative Research

*Dr. Rajesh Sharma, *Associate Professor, Jaipuria Institute of Management

**ABSTRACT**

The two most important and fundamental characteristics of any measurement procedure are reliability and validity. Validity is the extent to which a test measures what it claims to measure. A test needs to be valid in order for the results to be accurately applied and interpreted. Validity isn’t determined by a single statistic, but by a body of research that demonstrates the relationship between the test and the behavior it is intended to measure. The “validity” issue is not about singular truths, and it certainly is not limited to quantitative measurement; rather, by validity we mean that a research study, its parts, the conclusions drawn, and the applications based on it can be of high or low quality, or somewhere in between. The purpose of this article is to introduce the core concepts of reliability and validity of measurement, and understand the related research issues.
1.0 INTRODUCTION

In social sciences research and practice, the use of measurement is ubiquitous. Measurement instruments are frequently used for various kinds of assessments e.g. intelligence tests, aptitude tests, interest inventories, behavioral procedures etc. Using such tests involves some kind of measurement procedure and statistical theories for characterizing the results. Since the conclusions of empirical studies are based on values measured on research objects, it is therefore crucial to assess the quality of the measurements.

The most important property of measurement is validity. In general, validity is concerned with whether a measuring instrument measures what it intends to measure in the context in which it is to be applied. However, the reliability of measurements is also important because the researchers can rely on the accuracy of the measuring instrument. Establishing measurement reliability & validity is therefore of utmost importance in both applied and theoretical research. These two basic types of assessments i.e. validity and reliability, need to be conducted before the researcher can claim that the measurement process is sufficiently valid.

There is an inextricable link between validity and reliability (Black & Champion, 1976; Kerlinger, 1964; Hammersley, 1987). For example an unreliable measure cannot be valid (Crocker & Algina, 1986). Reliability is a necessary but insufficient condition for validity i.e. a valid instrument must be reliable, but a reliable instrument may not necessarily be valid.

Insofar as the 'validity' definitions are concerned, two important points emerge. Firstly, whether the means of measurement are accurate. Secondly, whether they are actually measuring what they are intended to measure. However, the notions of accuracy are associated with 'reliability' also. Most authors seem to attribute to 'reliability' more commonly than to 'validity' is the degree of replicability. Thus 'validity' refers to accuracy while 'reliability' refers to replicability.

Violations of instrument validity severely impact the function and functioning of a testing instrument. In some ways, validity inadequacies impart even more serious consequences on an instrument than reliability because validity is a comprehensive construct that cannot be definitively measured in any one given statistic, and also this instrumental testing property is often even less understood than reliability (Crocker & Algina, 1986; Gregory, 1992).
In this paper we shall first discuss the concept of validity in the research paradigm. First, we define “validity”. Secondly, we discuss the need to understand validity in qualitative research. Then by using a typical framework we explain the important types of validity. Finally, we summarise, discuss and conclude.

2.0 VALIDITY

The key concept frequently used to judge an instrument’s usefulness is its validity. Determining the validity of a test involves the use of data and other information both internal and external to the test instrument itself.

2.1 Definitions of 'Validity'

The exact nature of ‘validity’ is a highly debated topic in social science research since there exists no single or common definition of the term. A much cited definition of ‘validity’ is by Hammersley. An instrument is valid or true if it represents accurately those features of the phenomena, that it is intended to describe, explain or theorise (Hammersley, 1987). Validity measures the extent to which the set of indicators accurately represents a construct (Hair et al., 2006).

Validity has been defined as “the extent to which a test measures what it claims to measure” (Gregory, 1992). A measure is valid if it measures what it is supposed to measure, and without including other factors. The focus here is not necessarily on scores or items, but rather inferences made from the instrument which need to be “appropriate, meaningful, and useful” (Gregory, 1992).

In order to understand range of meanings attached to 'validity', it is essential to review a selection of the range of definitions given by leading authors.

Table 1.1 : Validity definitions by different authors

<table>
<thead>
<tr>
<th>DEFINITION</th>
<th>AUTHOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>'An agreement between two efforts to measure the same thing with different methods'</td>
<td>Campbell and Fisk (as cited in Hammersley, 1987)</td>
</tr>
<tr>
<td>'The measure that an instrument measures what it is called'</td>
<td>Black and Champion (1976)</td>
</tr>
</tbody>
</table>
From table 1.1., as far as 'validity' definitions are concerned, two common strands emerge: Firstly, whether the means of measurement are accurate. Secondly, whether they are actually measuring what they are intended to measure.

### 3.0 THE NEED TO UNDERSTAND VALIDITY IN QUALITATIVE RESEARCH

Researchers have a choice of two broad streams of research design that can be used to uncover the intricacies of a given phenomenon; quantitative or qualitative. Data gathered by either quantitative or qualitative methodologies can be used to verify or generate new theory (Glaser & Strauss, 1967). The decision of which research stream a particular researcher uses lies within the essence of the research question. If the researcher is seeking to verify an existing set of defined variables of an established theory (Hom & Kinicki, 2001) or to test defined variables of an untested theory (Mitchell, Holtom, Lee, Sablynski & Erez, 2001) then quantitative research would be appropriate. If the aim of the research is exploratory in nature, and seeks to unearth an understanding about an area that little is known about (Hendry, 2003) then qualitative methodology would be appropriate (Strauss & Corbin, 1998; Symon & Cassell, 1998).

The traditional criteria for 'validity' find its roots in a positivist tradition. Within the positivist terminology, 'validity' was the result of other empirical conceptions e.g. evidence, objectivity, truth, actuality, deduction, reason, fact and mathematical data to name just a few. It is within this tradition and terminology that quantitative research is traditionally defined.
Qualitative research, however, has concerned itself with the meanings and personal experiences of individuals, groups and sub-cultures. Some qualitative researchers have even argued that the term validity is not applicable to qualitative research. Few such researchers have espoused their own theories of 'validity' and have often generated or adopted what they consider to be more appropriate terms, such as 'trustworthiness', 'worthy', 'relevant', 'plausible', 'confirmable', 'credible' or 'representative' (Denzin & Lincoln, 1998; Guba & Lincoln, 1989; Hammersley, 1987; Mishler, 1990; Wolcott, 1990).

To summarise, qualitative research is based on subjective, interpretive and contextual data; whereas quantitative research attempts to control and/or exclude those elements (Auerbach & Silverstein, 2003, Glaser & Strauss, 1967, Maxwell, 1992, Strauss & Corbin, 1998). Thus, the positivist viewpoint of validity and the canons of rigor that are applied to quantitative research are not entirely applicable to qualitative research (Maxwell, 1992, Strauss & Corbin, 1998). To solve the dilemma of the measurement of validity, qualitative researchers have developed measurement concepts in line with the qualitative paradigm (Maxwell, 1992).

## 4.0 FRAMEWORK TO UNDERSTAND THE CONCEPT OF VALIDITY IN QUALITATIVE RESEARCH

There are many approaches to the generation and analysis of qualitative data, including phenomenology, grounded theory, qualitative content analysis and narrative analysis (Priest et al., 2002). In general, qualitative researchers consider that whichever approach is used, there is a need to determine how far the research findings are believable, accurate and useful (Creswell, 1998).

Joseph A. Maxwell (1992) developed five typologies to judge the validity of qualitative research: descriptive validity, interpretive validity, theoretical validity, generalizability, and evaluative validity. Although other researchers have also developed various categories; Maxwell’s five categories offer the best explanation thorough conceptualization and are discussed in the following part of this section.

### 4.1 Descriptive Validity
Descriptive 'validity' is concerned with the initial stage of research, usually involving data gathering. The central issue is factual accuracy in the informational statements that describe what was observed and experienced. The choice of language and selection of relevant data are the greatest threat to validity. If different observers or methods produce descriptively different data or accounts of the same events or situations, then the descriptive validity is questionable. (Maxwell, 1992).

Descriptive validity forms the base on which all the other forms of validity are built upon. Without an accurate account of the formative data all else is irrelevant (Glaser & Strauss, 1967). Walsh (2003), and Glaser and Strauss’s (1967) ‘credibility’ captures the same concept. The data must accurately reflect what the participant has said or done. The reporting of the data must also reflect the same accuracy, which means that the transcription is an accurate account of what was said or the transcription of the videotapes portrays the unfolding of events in an accurate manner.

### 4.2 Interpretative Validity

Within the qualitative paradigm, interpretation is typically viewed as an unavoidable element of data collection. Interpretive validity' captures how well the researcher reports the participants’ meaning of events, objects and/or behaviours. Here, the interpretations are not based on the researcher’s perspective but that of the participant. Interpretive validity is inherently a matter of inference from the words and actions of participants in the situations studied (Maxwell, 1992). Walsh (2003) terms interpretive validity as ‘conformability’ while ‘justifiability’ is the term used by Auerbach & Silverstein (2003).

### 4.3 Theoretical Validity

Theoretical validity is a more abstract analysis than the descriptive and interpretive validities. Theoretical validity goes beyond the concrete and descriptive and concerns itself with the constructions that researchers apply to, or develop, during the research (Maxwell, 1992). Theoretical validity seeks to evaluate the validity of the researcher’s concepts and the theorized relationships among the concepts in context with the phenomena. Auerbach and Silverstein (2003) termed it as ‘coherence’. The patterns, concepts, categories, properties, and dimensions must fit together to create the constructs so as to tell the story of the phenomena.
4.4. Generalisability

Most findings from experiments, survey designs, and quasi-experimental studies are intended to be generalized from the respondents sampled to some wider population. Such generalization is warranted only where subjects have been sampled randomly from the entire population to which the findings are applied (Reichardt and Cook, 1979).

The ability to generalise findings to wider groups and circumstances is one of the most common tests of ‘validity’ for quantitative research but is considered to be of little, or even no importance by many qualitative researchers. Qualitative research limits itself to 'internal' generalisations. Quantitative research, although, attempts to deal with both 'internal and 'external' generalisations, referring these as 'internal validity' and 'external validity' respectively (Maxwell, 1992).

External validity is however often of no importance to qualitative research since qualitative findings are generalisable to the development of theories and not wider populations. Unlike quantitative research whose validity depends upon the established canon of standardised tests and procedures, qualitative research embodies a vast and evolving body of techniques that can be modified or developed as the research demands.

4.5 Evaluative Validity

This form of validity refers to the application of an evaluative framework. ‘Evaluative validity’ moves away from the data itself and tries to assess the evaluations drawn by the researchers. The claims may be drawn from the researchers’ own understanding of the situation and not on the data. Evaluative frameworks are similar in both qualitative and quantitative research and many researchers make no claim to apply any evaluation to their research whatsoever (Maxwell, 1992).

Qualitative researchers work from the viewpoint that research findings are the result of an interpretive effort and are not objective but subjective (Denzin & Lincoln, 2000; Glaser & Strauss, 1967; Strauss & Corbin, 1998; Walsh, 2003). Qualitative researchers understand that they are part of the research process and thus affect the results. By applying these five categories of validity, qualitative researchers and readers can assess the validity of the findings.
In the following section, we present a short summary of various aspects of validity in qualitative research discussed so far in this paper.

5.0 SUMMARY

Validity is the extent to which the instrument measures what it purports to measure. Maxwell (1992), identified five types of validity in qualitative research: descriptive validity (i.e., factual accuracy of the account as documented by the researcher), interpretive validity (i.e., the extent to which an interpretation of the account represents an understanding of the perspective of the underlying group and the meanings attached to the members’ words and actions), theoretical validity (i.e., the degree to which a theoretical explanation developed from research findings is consistent with the data), evaluative validity (i.e., the extent to which an evaluation framework can be applied to the objects of study, as opposed to a descriptive, interpretive, or explanatory one), and generalizability (i.e., the extent to which a researcher can generalize the account of a particular situation, context, or population to other individuals, times, settings, or context).

Validity refers to the extent to which the inferences made from a test is justified and accurate. Further, reliability does not imply validity. That is, a reliable measure is measuring something consistently, but not necessarily what it is supposed to be measuring. In terms of accuracy and precision, reliability is precision, while validity is accuracy.

Validity necessitates demonstration that the propositions generated, refined, or tested match the causal conditions. There are two issues involved in matching scientific explanations of the world with actual conditions in it. First, do scientific researchers actually observe or measure what they think they are observing or measuring? Second, to what extent are the abstract constructs and postulates generated, refined, or tested by scientific researchers applicable across groups? The first issue addresses the issue of internal validity while second issue addresses external validity. Internal validity refers to the extent to which scientific observations and measurements are authentic representations of some reality while external validity addresses the degree to which such representations may be compared legitimately across groups.

6.0 DISCUSSION
The concept of validity is described by a wide range of terms in qualitative studies. This concept is not a single, fixed or universal concept, but “rather a contingent construct, inescapably grounded in the processes and intentions of particular research methodologies and projects” (Winter, 2000). Although some qualitative researchers have argued that the term validity is not applicable to qualitative research, but at the same time, they have realised the need for some kind of qualifying check or measure for their research. For example, Creswell & Miller (2000) suggest that the validity is affected by the researcher’s perception of validity in the study and his/her choice of paradigm assumption. As a result, many researchers have developed their own concepts of validity and have often generated or adopted what they consider to be more appropriate terms, such as, quality, rigor and trustworthiness (Davies & Dodd, 2002; Stenbacka, 2001).

The issue of validity in qualitative research has not been disregarded by Stenbacka (2001) as she has for the issue of reliability in qualitative research. Instead, she argues that the concept of validity should be redefined for qualitative researches. Stenbacka (2001) describes the notion of reliability as one of the quality concepts in qualitative research which "to be solved in order to claim a study as part of proper research".

‘Generalizability’ refers to the ability to apply the theory resulting from the study universally (Auerbach & Silverstein, 2003, Maxwell, 1992), which Walsh (2003) puts under the heading of ‘transferability’. For qualitative research generalizability is problematic. Qualitative research is concerned with the concepts and idiosyncratic characteristics of a select group; therefore, the findings or theory may only applicable to a similar group (Auerbach & Silverman, 2003, Maxwell, 1992; Strauss & Corbin, 1998).

Patton (2001) advocates the use of triangulation by stating that triangulation strengthens a study by combining methods. This can mean using several kinds of methods or data, including both quantitative and qualitative approaches. However, the idea of combining methods has been challenged with the argument that while mixing paradigms can be possible but mixing methods within one paradigm, such as qualitative research, is problematic (Barbour, 1998). Barbour (1998) does not disregard the notion of triangulation in qualitative paradigm and she states the need to define triangulation from a qualitative research’s perspective in each paradigm.
Both qualitative and quantitative paradigms strive to ensure that their findings are generated from an appropriate sample size and are valid. Quantitative results provide the researcher with hard facts and figures to validate and generate theory while a qualitative analysis uncovers a subjective viewpoint. Qualitative and quantitative methods compliment each other. Qualitative research can uncover new theories and variables. Quantitative research can test these new theories and variables. Quantitative research can highlight causal variable in theoretical models while qualitative research can provide the understanding of the detailed operationalization of the variable.

It is sometimes claimed, however, that traditional tests of rigour are not relevant in the qualitative paradigm - qualitative research should be judged against a different set of standards, using alternative terminology, and evaluated using an alternative set of strategies to quantitative research. In the light of this claim, qualitative researchers have devised an array of procedures for demonstrating validity.

Using either established or more novel approaches to assess the validity of research is one way of producing useful and trustworthy research findings. In determining the validity of research, reducing error is of prime concern. However, while adhering as closely as possible to a set of procedures to pursue truth and limit error is important, an approach that seeks to ensure rigour in research is equally important.

Perfect validity is unachievable; rather the goal should be achieving sufficient validity for a particular purpose of the researcher. No study can address all issues in measurement, but every study should consider at least some aspects of validity. While efforts can be made to minimise such risks, particularly systematic errors, they are acknowledged as a limitation in all types of research. While researchers should use as many approaches as possible to ensure validity, there remains the possibility that flaws may occur at the design, measurement or analysis stage, resulting in a less than perfect study.


