Chapter XLIII
A Cross Sample Analysis:
To Examine the Predictive Validity of an Instrument

Leping Liu
University of Nevada, Reno, USA

ABSTRACT

There are different methods to examine the predictive validity of an instrument. In this chapter, the author presents a method of validation—cross sample analysis, using a study as an example. This study demonstrates the procedures to determine whether a technology attitude instrument can predict student technology learning achievement consistently across four featured samples, with the data from two universities over a nine-year period. A base-model of prediction is first developed and then tested. The predictive validity of the instrument is confirmed by the model testing results that no significant differences exist between the means of the predicted and observed learning achievement scores in each featured sample group. Background knowledge and other relevant methods of validation are also reviewed in this chapter.

INTRODUCTION

In the field of using information technology in education, assessment has been a weak area (Liu & Handerson, 2003; Liu & Maddux, 2008). Literature shows that lack of validation of instruments used in the field is one issue related to this weakness (Maddux & Cummings, 1999). Educators and researchers often use self-developed instruments to assess student performance, instructional design, learning outcomes, or the effectiveness of technology integration (Christensen & Knezek, 2001). Unfortunately, in many studies conducted with such self developed instruments, there are no reports about the validity of the instrument (Liu & Maddux, 2008). If the validity of the instruments is not confirmed, that is, if they are not measuring what they are supposed to measure, the results of the studies cannot be considered meaningful; and such studies would not provide any valuable

insights to the field, to the literature, or to the practice of other educators.

There are multiple forms of validity, depending on the purpose of the instrument, research questions, and specific type of inference the study intends to make. Procedures and statistics methods of validation are varied, and thorough validation of any instrument is a complex task that requires careful design and may take years to complete (Bryant, 2004; Moody, 2001; Strube, 2004). This chapter will focus on *predictive validity* and an applied method of validation—*cross sample analysis.* The cross-sample analysis is performed in a study that examines whether a technology attitude instrument measuring four attitude variables can predict student technology learning achievement consistently across four featured sample groups, with the data from two universities over a nine-year period.

BACKGROUND

Overview of Predictive Validity

Validity of an instrument is usually defined as the extent to which the instrument measures what it is supposed to measure. In social science studies, three major types of test validity are often considered important: (a) *content validity*, (b) *criterion validity*, and (c) *construct validity. Content validity* concerns the extent to which an instrument measures all relevant aspects or domains it is supposed to measure. Multivariate statistical procedures, such as exploratory factor analysis or confirmatory factor analysis, are often used to assess content validity (Bryant & Yarnold, 1995). *Criterion validity* deals with the degree to which an instrument can accurately predict a well-accepted indicator of a given concept, or a criterion. Statistics procedures such as the Pearson correlation coefficient, logistic regression, multiple regression, and structural equation modeling are usually used to assess criterion validity (Anastasi, 1988; Keith, 2006; Wright, 1995). *Construct validity* concerns the conceptual accuracy of measurement, to determine "whether a given measure actually assesses the underlying conceptual variable, or construct, that the measure is intended to represent" (Bryant, 2004, p.111). Construct validation remains among those most difficult challenges in the field of social sciences (Webb, Campbell, Schwartz, Sechrest, & Grove, 1981).

Predictive validity falls into the category of *criterion validity.* Again, it is defined as the degree to which an *instrument* can predict *what it is intended to predict.* Two concepts need to be clarified: (a) the test scores from the "instrument" are the measures of the predictor variables, which can be, for example, attitude variables, or any pre-obtained knowledge and skills; and (b) "what it is intended to predict" is the criterion variable (or response variable) to be predicted and the scores are obtained from another measure; for example, the criterion variable can be learning outcomes measured by standardized examinations, or preference to online learning obtained from survey questionnaire.

Researchers are usually able to examine three types of criterion validity—*predictive validity, concurrent validity*, and *retrospective validity* (Bryant, 2004); depending on the time when the test scores and the criterion measure are obtained. First, if a researcher obtains the test score *before* measuring the criterion, to examine how accurately the former predicts the latter, this type of criterion validity is known as *predictive validity.* Second, if the researcher acquires the test score and the criterion measure *at the same time* to assess the degree to which the two correlate, the criterion validity being examined is *concurrent validity.* Finally, to assess another type of criterion validation, the researcher may also obtain the test score *after* the criterion measure is performed; this is known as *retrospective validity.* This chapter will focus on the method to examine *predictive validity.*