Survey Design and Data Analysis : Methodological Considerations to Reduce Error and Increase Instrument Utility in Quantitative and Mixed Methods Research

Lasonja Kennedy, Presenter

Background

Researchers regularly develop measures to assess concepts of interest. It is not the answer (i.e. response) to a question that is of interest, but rather the extent to which the answer can be shown to have a predictive relationship to facts or subjective states that are of interest [1]. Proper survey research, therefore, requires instruments are valid and reliable for the population examined.

A myriad of options exist for development of surveys, whether designed as prototypes or constructed from previously validated scales. The researcher must decide upon the type of instrument (form, electronic), theoretical basis for the tool, and a reasonable number of survey items. Additionally, the order of questions, stem for each item, and format for response options has to be determined. Although items from the aforementioned lists may be easily addressed, additional considerations must be given to conceptualization and validation of the tool. Most importantly, foundational applications are needed to maximize the utility of survey items and information collected.

The current presentation will review important research methods to reduce error when designing novel instruments and techniques useful for proper analysis of data collected from the tool. This includes basic information for those new to the field of survey design and survey research methods as well as relative information for participants skilled within the area.

Aims

Aims of the presentation are:

I. Review basic elements of hypothesis testing;
II. Present essential survey design techniques that may impact variability of scores and standard error (SE);
III. Demonstrate analytic methods to identify latent factors; and,
IV. List tools most useful in design and analysis for various expected outcomes.

Analytic Methods to Identify Latent Factors

To examine the latent structure of an instrument, an exploratory factor analysis utilizing principal component analysis (PCA) is often conducted. The factor analytic process is used to determine the extent to which shared variance exists among a set of variables [4]. The mathematical procedure allocates variance across a smaller number of underlying, hypothetical, and unobservable variables called factors [5]. Items that cluster together or contain shared variances are usually grouped within the same construct. Factor analysis permits researchers to eliminate extraneous test items or variables and group items with similar themes.

An exploratory factor analysis (EFA) requires the subjective analysis of the researcher. The researcher must determine which factors will be retained and which ones removed in the exploratory process. The goal of EFA is to extract the maximum amount of variance from a set of variables and to produce a linear set of components useful for research[6]. There are general guidelines to determine the number of factors to retain when conducting a factor analysis. Five steps for the analytic process include: select desired variables for the analysis; prepare the correlation matrix for factor or component analysis; extract variables; perform a rotation to increase interpretability; interpret results.

Elemental Components of Hypothesis Testing

Hypothesis testing allows researchers to make inferences about a population based upon data from a sample [2]. Basic procedures for hypothesis-testing includes (1) statement of a hypothesis about the population; (2) predict characteristics for sample; (3) select sample; (4) compare sample data with hypothesis' prediction[2]. Additional details are provided below exclusive of directional hypotheses and treatment effects.

Hypothesis Statement:

1. Null Hypothesis (H_0): states that the independent variable (IV) has no effect on dependent variable (DV); notes no relationship or impact
2. Alternative Hypothesis (H_1): also known as scientific hypothesis; predicts the IV will have an effect on the DV.

Decision Criteria:

1. Establishment of probability values commonly recognized as level of significance (i.e. alpha level).
2. Most often used- the alpha level for the hypothesis test: $\alpha = .05$ (5%); $\alpha = .01$ (1%); and $\alpha = .001$ (0.1%)
3. Values defined by alpha levels fall within a *critical region* that define outcomes inconsistent with the H_0
4. Sample means located in critical region indicates a need to reject the null hypothesis

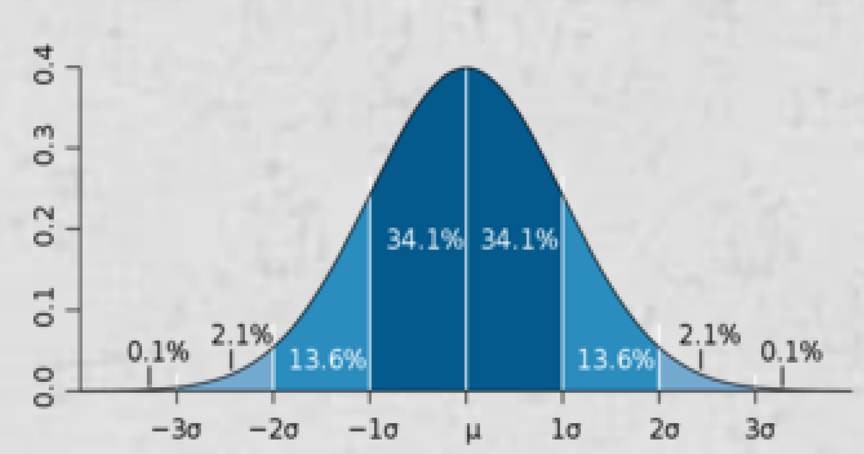

Figure 1: Distribution curve displaying critical regions

Data Collection

1. Conducted after statement of hypothesis and establishment of decision criteria
2. Random sample selected (in most cases)

Determine Final Outcome

1. Reject the null hypothesis
2. Fail to reject the null

Survey Design Techniques Impacting Variability of Scores and Standard Error (SE)

A plethora of techniques exist for development of research tools. In addition to the research question, additional uses for and purposes of survey data are necessary to consider during the developmental process. It is important to also consider which (and how) various techniques may impact variability of scores and the standard error (SE) as each provides an indication of how sample data reflect the population examined.

Factors to consider include [3]:

I. Proper scale development: item content reflective of the underlying latent variable
II. Construction of useful items with appropriate reading levels
III. Equally weighted scale items
IV. Response formats and application within the survey

Design and Analytic Tools

Analysis of survey data depend, in part, on the instrument's design. Outside of descriptive statistics, tests to analyze survey data are often classified as parametric or nonparametric. Parametric tests are commonly used inferential techniques and require that certain assumptions are met; nonparametric tests are used when parametric tests are inappropriate for use. A decision map for choosing a parametric test is provided below [2]. Similar maps exist for choosing nonparametric tests and methods to evaluate relationships between variables.

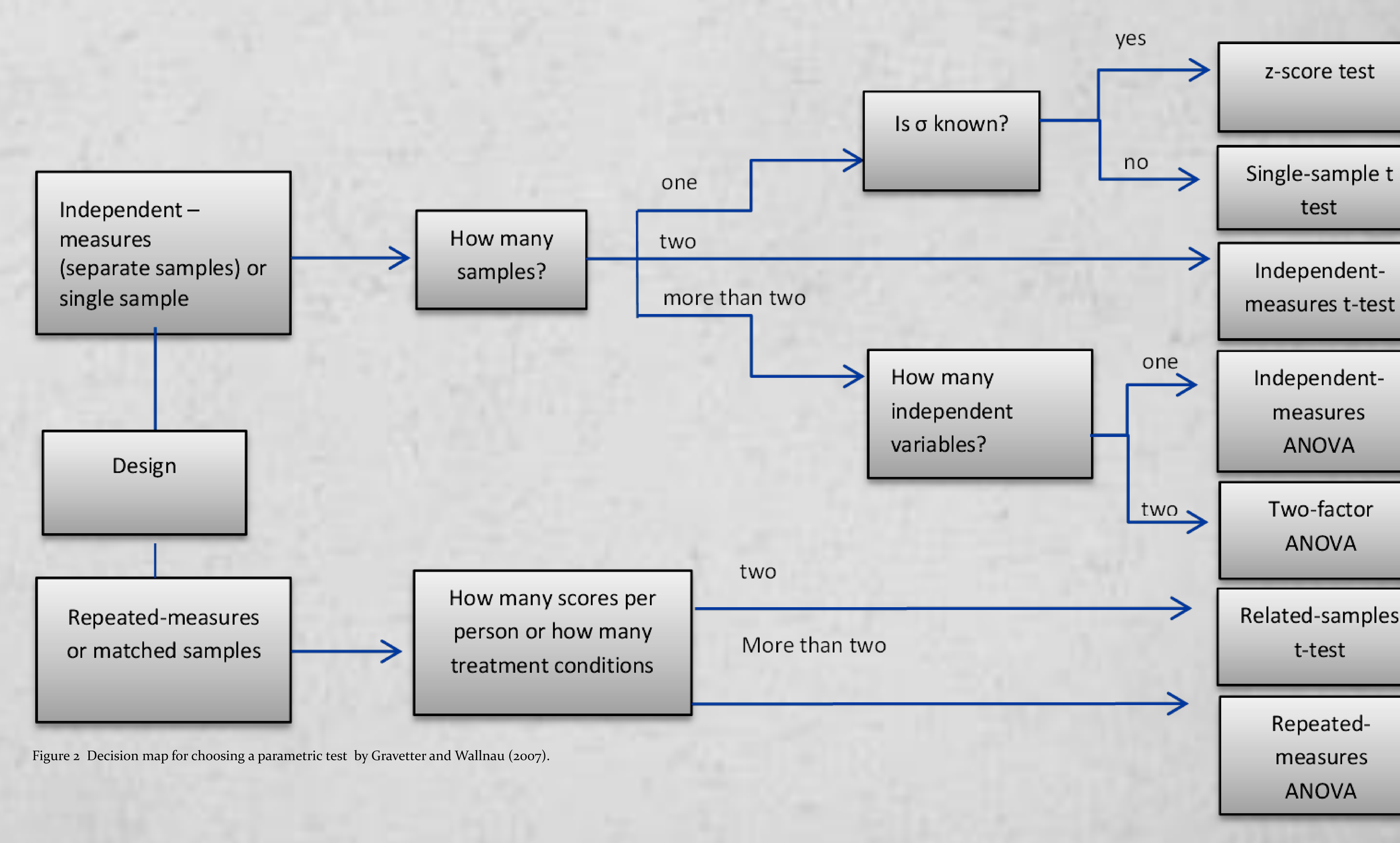

Figure 2 Decision map for choosing a parametric test by Gravetter and Wallnau (2007).

References:
1.: Fowler, F. (1984). *Survey research methods*. Thousand Oaks, CA: Sage.; 2.: Gravetter, F.J. and Wallnau, L. B. (2007). *Statistics for the behavioral sciences*. Belmont, CA: Thomson Wadsworth.; 3.: DeVellis, R. F. (2003). Scale development. Thousand Oaks, CA: Sage.; 4.: Williams, F. (1992). *Reasoning with Statistics: How to Read Quantitative Research* (4th ed.). Fort Worth, TX: Harcourt Brace Jovanovich.; 5.: Mertler, C., & Vannatta, R. (2005). *Advanced multivariate statistical methods* (3rd ed.). Glendale, CA: Pyrczak; 6.: Tabachnick, B.,& Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.