Quantative Research Primer

For

EDUC 500

Dr. Franc Feng
University of British Columbia
June 2014
http://blogs.ubc.ca/educ500/

University British Columbia
EDUC 500 (2014)
Methodologies in Quantative Research

Franc Feng

Quantitative Methods (Obtrusive)

- Experimental
  - Blind
  - Double Blind
- Quasi-Experimental
- Ex Post Facto
- Survey
- Historical Trend Aanalysis

Quantitative Techniques

- Frequency Distribution
  - Central Tendency
    - Mean
    - Median
    - Mode
- Probability
- Variability
  - Standard Deviation
    - Variance
- z-scores
- t statistic
- Analysis of Variance
- Repeated Measures
- Correlation
- Regression
- Chi-Square

Ends of Quantitative Research

- Prediction
- Description
- Evaluation
- Interpretation
- Normative Measurement

Methodological Issues in Quantitative Research

- Validity
- Reliability
- Generalizability
- Instrumentation
- Sampling
- Power
- Ideology
- Equity
- Deception

Quantitative Research— Working Definitions

Initial considerations: limits and responsibilities

Rather than beginning with the parameters we need to concern ourselves with quantitative research, in this primer, I want to try a different approach- that can be glossed over in the attempt to convey the “objective facts” about quantitative research. In so doing, I am trying to write an introductory section that anticipates complexities that lie might ahead in your development as researchers, as you struggle with related questions of research, method and analysis, insofar as quantitative analysis is concerned. This introductory section also serves another purpose, in trying to relate the practice of quantitative research with ethical questions concerning its operations (e.g. in terms of translating the complex world into numbers, the need to pay attention to these transformations, etc.). A third rationale for this section – is to try to situate core elements of quantitative research approach within a whole, and a context- wherein specific terms like constructs and operationalize are introduced as core to the whole research endeavor, *before* they are explained in isolation of context.

Research Problem

Whether with qualitative or quantitative research, here is where it all begins. Yet, does the problem drive the method? Or does method drive the research problem? The latter option appears to be problematic, since it implies that regardless of the question, precedence goes to methodology. By extension, this logic suggests that whatever question one might have in mind, it will be driven by method. Following such a path, we could also argue method almost dictates the problem. So it is the other way around then? Does the research problem drive the method? While one might observe that certain questions are more amenable towards certain methodology, the question/answer is likely more nuanced. The answer is more like: it depends. While one cannot preclude the kinds of research questions by insisting on methodology beforehand, the reverse does not necessarily hold. True, in the traditional books, that is what one would find- clearly the problem drives the whole research. In the real life context however, it is more like the question is emergent with the method, and it is only is through multiple iterations of engaging with literature, problem and method that one clarifies a question, which by that time, could be amenable to one or several methods (as in a mixed-design research).

Method

When is student in interested in the possibility of conducting quantitative research, whether by itself or as part of a mixed-research model, as s/he crafts her/his research design, s/he needs to be aware of both the possibilities and limitations of quantitative research methodology. Given the power of quantitative research, its ambit, its place, and it ability to influence public policy, it is prudent in her/his development as a quantitative researcher, to be also reminded of humility, always bearing limits in mind, about quantitative research methods. The student needs to remind her/himself, for instance, that whether researching literacy, dropout rates, motivation, anxiety or self-efficacy- that which s/he defines as dependent and independent variables in our research design, might be more appropriately termed "proxy-variables". This is because s/he needs to be aware that we can never *really* measure literacy, drop-out rates, motivation, anxiety or self-efficacy, only variables we have operationalized through constructs that stand in proxy for these- informed by the literature, our narratives, experience, backgrounds, etc. Importantly, with regards to instrumentation, s/he also cannot lose sight of her/his responsibility when creating constructs that operationalize complex notions like literacy, in assigning numbers to human attributes, around which quantitative analysis is design and conducted.

Analysis

Quantitative analysis thus codes the complex world through research models, wherein findings are imputed to the concrete everyday "real world" through independent and dependent variables that are operationalized from abstract constructs. The researcher need to be aware of the double move that takes the general situation in which we find ourselves immersed, creates a specific model that in turn, we generalize back- to the real world. Given the turns, the researcher need to be aware of possible slippages that could occur in transformation. As well, the researcher needs to be reminded that we can never be certain that we have found an effect or not, only that we are confident within margins of error- or, differently put, willing to take the chance that we could be wrong, given our set probability levels. Even if we have done our best in observing protocols attendant to our chosen method, the researcher needs to be reminded that with quantitative research, it is possible to commit the error of discovering an effect when there is none, or missing an effect that is present. For we are after all, interpreting through variables around a model of the world, with our findings only as good as our models, based upon arbitrary level of confidence we have allowed ourselves. In sum, even as we realize the power/potential of quantitative research to assess change (including natural changes) on a large scale, produce instruments with which we can assess or better ourselves, and improve our literacy, we need to be at the same time, informed by, and humble about, limitations of quantitative research while remembering our responsibility as researcher, reminded that there might be factors we cannot possible control, have not considered, or changes that we might not be aware of, etc., and our large–scale findings carry direct implications for public policies.

Review of Core Concepts in Quantitative Research

Having situated key concepts of quantitative research within the research context, in this section, I list these, in accordance to conventionally categories of quantitative research in three supplementary sections: 1) Basics, 3) Descriptive Statistics and 3) Inferential Statistics. [1]

The Basics

Phenomenon – the object under study, which is amenable to investigation. Although we often associate phenomenon with the study of individuals, or their characteristics, we are reminded that these objects could also be documents, archival records or natural events.

Construct – as indicated above, our research is underpinned and informed by constructs that we conceptualize through combining our reading of the literature with our narratives, experiential learning and our research interests.

Variable – this is the observable characteristic of the object or event that we have classified and described through operationalization. The variable is expressed symbolically with numerals and values. In quantitative research, we categorize variables as independent variables and dependent variables, wherein we attempt to learn the nature of relationships between them.

[1] In these preliminary pages, I have sketched the outlines of a quantitative primer, one that is a work-in-progress, with more to follow. Although some of the wordings are similar to texts I have consulted, since I have written this primer mostly as a conscious stream of thought, there are likely errors. Comments, critiques and suggestions are welcome.

Independent variables – this is often known as the variable that is manipulated by the researcher. Often, as with gender or class, this involves the classification and coding of variables are numbers through operationalization.

Dependent variables – this is the variable that we measure, it is the variable that we sometimes predict, a variable that we observe, that we seek to account for. Often, but not always, this variable is measured as interval data (e.g. achievement scores, Likert scores).

Operationalization – we assign "meaning" to the constructs, and hence the variables, when we classify and convert phenomena like human attributes into numbers that we can measure and manipulate.

Measurement – this is a numeric expression of a given observation. Measurement is often expressed in terms of levels of sophistication, in terms of scales. There are four types of measurement scales: namely the nominal, ordinal, interval and the ratio that are often understood in terms of their relative "strengths".

Nominal scale – otherwise known as categorical data, here we are assigning numbers or symbols to designate our variables. Of the four possible, the nominal scale is considered to be the "weakest" level of measurement, since it conveys the least amount of information of variables. These are typically used to classify independent variables; examples include gender, job or course, wherein each are coded into numbers.

Ordinal scale – one level "up" from nominal scales are ordinal scales. Otherwise known as rank data. Here is have data grouped and hierarchically ranked. Ordinal data are whole numbers. Like nominal data, it can be coded. An additional advantage concerns ranking, we can rank order items on an interval scale; examples include rank ordering educational attainment.

Interval scale –moving up further we have a continuous scale, wherein we have decimal points. There is however one limitation. Even though we can compare multiple scores, etc. – we cannot say one measure is a multiple of another. This is because there is a limitation –we have no “true-zero” point that is equivalent to the absence of the variable being measured. Examples include exams scores, age (including fractions) and distances.

Ratio scale - Unlike the interval scale, here is it possible to compare numbers and argue one is a multiple of another, examples include instances, where it is possible to compare dependant variables like achievement and argue that one score is twice as much as the other. This is because an ratio scale has an “true-zero” point in which the attribute being measured is absent.

Validity – this is a measure of the how well the research design captures the variable under inquiry. Thinking back to how variables are operationalized, validity then concerns the degree to which the variables in question purport to measure a given phenomenon.

Reliability- this is measure of the repeatability and stability, with the question of reliability, we are concerned with whether we can expect results to be consistent under similar conditions.

Internal Validity – this is a term with origins in work of Campbell and Stanley (1963) that concerns the attribution of observations to the hypothesis of concern, rather than to rival hypothesis. Associated with the concept is the notion of threats to internal validity

External Validity – this concept concerns the extent to which we can generalize the findings of our research, from our specific study to other participants and contexts.

Descriptive Statistics

Descriptive statistics – this is the branch of statistics used for summarize population and sample data. Often associated with frequency counts, histograms and graphing techniques, descriptive statistics emphasize measures of central tendency like mean, mode and median, and variability/ dispersion like range and standard deviation.

Central Tendency – concerns statistical measures that describe how scores cluster in particular distributions. When we speak of central tendency, we are often concerned with finding a score that defines the centre of a given distribution, often expressed as the need to find the most representative score. The central tendency, given the way that it describes the middle point of the distribution, is often seen as the single most important point in the distribution. There are three indices of central tendency: the mean, mode and median,

Mean – also known as the arithmetic mean or arithmetic average, the mean is arrived at through adding up all the scores that make up the given distribution and dividing that total by the number of observations/scores. Means can be both population and samples means, and these are differentially identified: generally Greek letters are used to represent population characteristics and conventional letters for the sample. The mean is the most sensitive measure of central tendency, sensitive to all scores in the distribution, the mean is sometimes known as the "centre of gravity" of a distribution, wherein the mean can be construed as a balance point.

Median – by definition, the median is the 50^{th} percentile of the distribution, the point at which the distribution divides into halves, where there will be an equal number of cases/participants/scores above and below the median. Technically a point rather than a score per se, the median is the most cited measure next to the mean.

Mode – by definition, the mode is the most frequently occurring observation. The mode is thus the most common observation among a group of cases; it is the score with the greatest frequency in a frequency distribution, appearing tallest in that graph. It is also possible to have no mode, as with rectangular distributions, when every score occurs the same number of times. It is also possible to have multiple modes, as with bi-modal, tri-modal or multi-modal distributions.

Distribution – this is made up of a collection of measurements that are usually viewed in terms of frequency wherein observations are assigned to points on the distribution. With statistics we are interested in measures of central tendency and variability in describing a given distribution.

Range – the range allows us a crude idea of the spread of any given distribution since it is merely the difference between the highest and lowest scores in a distribution. Based on only two scores, in a given distribution, the range is insensitive to nearly all but a few scores. For instance, given that the range measures only the "ends" of the distribution, it does not change the range if we were to add any number of scores in between.

Standard Deviation – by definition, the standard deviation is the square root of the squared deviation scores about the mean of the distribution, or most simply stated, the square root of the variance. Having an idea of the mean and standard deviation allows us an idea of the shape of the distribution and to make inferences, whether for instance a test is easy, or whether a population is more/less homogenous, given the scores distributed around the mean.

Variance - since by definition, the variance is computed as the mean of square deviations about the mean of a distribution, similar to the mean, the variance is sensitive to all scores in a distribution. Since deviations are computed around the mean, variance being a measure of the square deviation of scores about the mean of a distribution, also hover around a balance point.

Skewness – in terms of distribution, the normal curve constitutes a special case, where the mean, median and mode are equal due to the symmetrical distribution. When distribution deviates from the normal curve, the shape becomes skewed, with the mean being "pulled" towards one "tail", either positively or negatively. While the mean is not as useful in skewed distribution, since the median is always between the mean and the mode, it is still preferred for inferential statistics.

Kurtosis – here is another instance of deviations from the normal curve, only in this case rather than becoming skewed in being elongated and flattened along the horizontal axis, the distribution peaks or flattens, in the first instance, resulting in changes in standard deviations.

Correlation – this is a measure of the relationship between two sets of variables that ranges from –1 to 0 to +1. Correlations are represented graphically, in the form of scatterplots and as correlation coefficients and correlation matrix showing inter-correlations between sets of variables of interest. It is important to emphasize that correlations only apply to linear relationships and they do not impute cause.

Variance- also known as the coefficient of determination, the variance is obtained from squaring the correlation coefficient. The variance represents the proportion of variance that the two measures have in common. Whereas the correlation coefficient cannot be thought of in terms of a portion of percentage of a relationship, the variance can be.

Inferential Statistics

Inferential statistics – this is the branch of statistics concerning the study of samples rather than whole populations. The idea however, inherently links sample and population – given that in inferential statistics we obtain representative samples from the population, with which we in turn, make inferences back to the population of interest.

Population – this is the set of all individuals we are interested in our quantitative research. While it would be ideal to study entire populations, for reasons of economy, scale, management, feasibility, etc. – we often end up with samples that are drawn from populations.

Sample – when we cannot study intact populations, we often work with samples. In quantitative research, the idea is to work with representative samples, representative samples enable us to make inferences back to the parent population.

Sampling- this is the technique through which researchers draw participants from a population. When we draw a sample from the population, it is not necessarily representative of the population. In order for the sample to be representative, it has to be random. Randomness is a requirement that is central to quantitative research methodology.

Random sampling – this is a statistical method of obtaining a sample from a population, wherein every participant has an equal chance to be selected. All things being equal, since a random sample is supposed to be an unbiased sample, it should theoretically be representative of the population from which it is drawn.

Treatment /Treatment effects– this is the measure of our intervention and the effect that we will attempting to establish through positing relationships between our research variables, given alpha levels of statistical significance, typically at the 0.05 or 0.01 levels.

Level of statistical significance- otherwise known as the alpha level in quantitative research– typically 0.05 or 0.01. This is the level that we set to designate the presence of treatment effects that is greater than chance. While typically in educational research, we set the alpha level at 0.05 (where we are willing to be wrong one time out of twenty), if more stringent assurance is needed, when the consequences of being wrong carry life and death implications, as with pharmaceutical tests, we set the alpha level at 0.01, or even 0.001, where we are only willing to be wrong one out of a hundred times, and a thousand times respectively.

Hypothesis – Typically in quantitative research, when establishing relationships between variables or to posit the tenability of treatment effects, we work with a set of complimentary hypothesis, the *null hypothesis* and the *alternative hypothesis*. Theses hypotheses are in turn, constructed by and informed by our search of the literature and our experience concerning the relationships between the variables under investigation, wherein the presence of such relationship, suggest the tenability of the alternative hypothesis given the presence of treatment effects that are outside the probability of set chance levels, we call alpha levels, typically 0.01 or 0.05.

Null hypothesis – This is tentative hypothesis, expressed as H_o, the researcher forwards in terms of *no* treatment effects that she attempts to call into question through establishing the tenability of an alternative hypothesis, H_1, at probability levels, typically 0.01 or 0.05.

Alternative hypothesis – This is tentative hypothesis, expressed as H_1, that the researcher establishes as tenable, when findings indicate the tenability of treatment effects at probability levels, typically 0.01 or 0.05.