

Chapter Six

Fundamentals of Quantitative Research in the Field of Teaching English as a Foreign Language

ANTHONY WATT

Introduction: Why Do Quantitative Research in English as a Foreign Language?

Adopting a quantitative approaches in researching teaching English as a Foreign Language (EFL) will achieve much the same purpose as utilising this methodology in the broader fields of education and psychology. Thinking beyond the stereotype of quantitative designs being only applicable to experiments or data collections of a scientific nature is critical in creating opportunities for new knowledge and information in the domain. The renowned psychological and educational researcher Lee Cronbach provided a valuable perspective as to how quantitative educational research framed in the true scientific method may not always be feasible.

The habits of the psychologist and biologist do not fit research on classroom instruction. Rats receiving a drug or placebo are properly considered to be independent subjects; what one rat does has no effect on the score of the next (unless the experimenter somehow introduces correlated errors). Students in a class, however, do not provide independent evidence. What the class experiences goes beyond the treatment specified by the experimenter (Cronbach, 1976, p. 10).

Although research in the education domain has recently deviated from a focus on the quantitative approach, with a greater emphasis on qualitative methods, the use of numbers can be a very useful tool, either as part of a larger project that employs multiple methods or as a basis for a complete piece of work (Berry, 2005). Brown (2011) suggested that research in the area of EFL involves a systematic and principled inquiry of an area of interest. Furthermore, he defined quantitative research as involving studies that focus on “counting things and the patterns that emerge from those counts” (p. 192).

If we consider EFL within the tenet of critical pedagogy, an amalgam description could be considered as the development of communicative abilities in English and the ability to apply them to developing a critical awareness of the world and the capacity to act for positive change and equitable outcomes for all (Crookes & Lehner, 1998). Quantitative research practices can play an important role in generating evidence to substantiate the viability of this goal, or evaluate exemplars of how EFL teaching and learning practices are aligned with the critical pedagogy philosophy.

Quantitative Research Design in EFL

Trochim (2006) defined research design as “the glue that holds the research project together. A design is used to structure the research, to show how all of the major parts of the research project—the samples or groups, measures, treatments or programs, and methods of assignment” (p. NA). In support of this endeavour a range of descriptors exist that can be utilised to categorise quantitative research designs suitable for the EFL domain. It is useful to consider these designs along a continuum from the simple through to the more complex. It then becomes the responsibilities of the researcher or team when developing the project to determine the level of complexity required to achieve the research goals. De Vaus (2001) outlined two straightforward questions that can serve as the stimulus to establishing the design framework: (a) What is going on (descriptive research)?; and (b) Why is it going on (explanatory research)? Furthermore, design can also be considered against the following classifications: (a) Descriptive which involves describing the information being collected using measures such

as surveys or assessments records; (b) Associational, which further than description by attempting to determine how the characteristics assessed are related so can better understand phenomena; and (c) Intervention, a more complex approach whereby through intervening and evaluating resultant effects, change in the characteristics of interest can be established (Fraenkel & Wallen, 2011).

High quality literature related to research designs in the social sciences, education and EFL currently exists (e.g., Brown, 2011; Fraenkel, Wallen, & Hyun, 2011; Trochim, 2006). In considering relevant work, the following set of designs that can be applied in the context of EFL will be described individually in greater detail. These designs are descriptive, correlational, comparative, causal, quasi-experimental, and experimental.

Descriptive Designs

Descriptive designs are a necessary element of all quantitative research and can serve as an independent study, or more commonly as the platform on which to construct subsequent analyses. Descriptive research confirms or rejects patterns in the information collected, promotes theoretical discussion of observed data, directs the subsequent design of cause and effect studies, and provides an important base framework for interpreting and generalizing results derived from more complex designs (e.g., inferential, experimental) (McEwan, 2008). The data generated within these designs details scores, behaviours, human characteristics, occurrences, equipment counts, outcomes etc. and is derived from descriptive statistics such as frequency, percentages, mean, mode, median, range, and standard deviation (Brown, 2005). A simple descriptive design could involve determining the country of origin frequencies of the individuals who are completing an off shore course in ESL. Results could generate information that could serve to allow course instructor to set up tutorial groups of individuals of a similar background or indeed work to diversify the groups by assigning members of different nationalities to different groups. Although the simplest research design, descriptive findings can often be the most readily accessible to students and researchers with less of a background in quantitative methods.

Correlational Designs

The correlational design is used to examine the relationship or degree of association between two or more characteristics, attributes, numbers, or quantities that can be measured or counted. Correlational designs can only provide information in regards to the existence of an association and do not allow the researchers to infer a cause. Researchers who incorporate a correlational research design do not manipulate any elements of interest but collect data on existing factors or characteristics and typically examine any relationships within a single defined group (Johnson, 2000). This design allows for any two scores collected to be compared, however, it is important to use logic to support evaluating the association. The basic reporting statistic is a correlation value, in which the higher the value the stronger the link (e.g., height and weight), whereas in the negative direction, the value indicates the scores operate in opposite directions (e.g., days of sunshine and precipitation). Correlational designs also allow for the prediction of how a score might influence another score in the future. Many studies in education and EFL could be developed around the correlation design. For example, evaluating the relationship between the students entry level English skills and time required to complete an ESL program of studies establish. Furthermore, entry level English skills could also be compared with scores on the first course exam, and provide information that could assist in predicting course performance based on entry level English skills. As with descriptive designs, correlational designs are considered a simple quantitative approach.

Causal-Comparative Designs

Causal-Comparative designs involve research in which the researcher attempts to answer questions in regard to simple differences between two or more groups (Lauer, 2004). The contrast effects have already occurred and the researcher attempts to determine whether one characteristic may have influenced another characteristic. The main difference between this design and the early designs discussed is that the researcher may now put forward opinions or reasonable inferences regarding the cause of any possible find-

ings (Gay et al., 2006). Investigators will typically decide to use a causal-comparative design when the groups (and associated variable) involved cannot be manipulated for ethical and practical reasons. The forming of groups is based on natural occurrence and where the characteristic of interest is present or absent (Gall, Borg, & Gall, 2003) and the observation relates to natural characteristics and not manipulated variables. For example, it may only be possible to compare scores regarding different tutorial groups on the basis of assessments that constitute part of a set course. Students may not have the time to prepare for a specific measure developed for the research, or the researchers may not have the time or funding to manage administration of a separate measure. In this case the findings may tentatively allow us to infer that differences in scores between tutorial groups may be due to teaching or teacher characteristics. Simple causal-comparative designs provide an excellent source of evidence upon which to base the development of more complex manipulation or intervention studies.

Quasi-Experimental Designs

Quasi-experimental design often used in educational research because investigators cannot always randomly assign participants to groups in particular settings or manipulate an intervention because of ethics and system constraints. It is not typically appropriate to withhold or administer an educational technique to one group and purposely do the opposite to a similar group. Therefore, a quasi-experimental design could be described as a best attempt at an experiment when it is impossible, or not reasonable, to meet all the criteria of a true experiment. Yet, there is still an attempt to isolate the treatment so that inferences can best be attributed to the treatment or intervention. Typically, as its primary goal, research incorporating quasi-experimental methodologies attempts to answer questions such as: “Does a treatment or intervention have an impact?” and “What is the relationship between program practices and outcomes?” (Dimsdale & Kutner, 2004). In quasi-experimental studies, researchers attempt to control for differences between non-randomly assigned groups in a number of ways. Although groups are not randomly assigned the design may necessitate the matching of groups (e.g., year one ESL students with year one ESL students). Additionally, the

investigators can incorporate statistical controls such as pretesting of the matched groups before the treatment. It is likely in research within the ESL domain, that the quasi-experimental design will serve as an appropriate approach in which new ideas or techniques could be evaluated.

Experimental Designs

The critical element in experimental design is the random assignment of the participants to one of several treatment or intervention groups. A basic experimental design involves randomly assigning the members of the group of interest to either an experimental cohort, that receives a treatment, or the control cohort, who receive no treatment. Anderman (2006) succinctly describes the principal goal of an experimental design that reflects an education context.

differences in treatment between the experimental and the control group are tightly controlled, and if subsequent to the experiment there are measurable differences between the two groups that were not present before the experiment, then researchers often conclude that the experimental manipulation “caused” the differences to occur (p 5)

In ESL research, an experimental design could involve the random assignment of a second year Masters of TESOL class into three groups. All students are using a text books and receive the ascribed lectures and tutorials and have completed their first course progress language test. Treatment group one will access a new language tutorial website for three weeks following the course progress language test, treatment group two will join a language conversation club for three works, and the control group will complete the course in standard delivery. After four weeks the students will complete the second language test of the course. Test score differences between the three groups will be considered as an indication of the efficacy of the treatments. It should be recognized that in experimental research in education settings that many other factors can influence results, such as members of the control group working harder in the second phase, or preparing more effectively for the second test.

Planning a Quantitative Research Project in EFL

Quantitative research is typically framed by the important decisions made prior to the commencement of a project. Critical areas for consideration include the formulation of aims that connect to theory, detailing of specific quantifiable research questions, and the selection and implementation of a logical and viable design.

General Aims of the Research in EFL

Developing aims in quantitative research necessitates substantive reflection and evaluation of both theory and previous research that match the general ideas being proposed. Further to this, in the ESL domain aims can also be derived from the observation and experience of practical problems in the field (Trochim, 2006). Once an idea is formulated, the researcher should establish that literature and previous research exist that will support planning. At many levels of quantitative research it is acceptable to replicate a previous study using different sample demographics or to pursue an alternative interpretation of a theory or construct. Aims, therefore, should constitute the project team's ideas in an area of ESL research that are detailed in a manner that others from the ESL field can easily comprehend. Similar to the aims, in quantitative research a hypothesis can also be proposed. The hypothesis is still a general statement without connection to measurement or findings, but represents a specific statement of the predictive thinking that underpins the aims of the research (Fraenkal et al. 2011).

Quantitative Research Questions in EFL

Research questions will serve as the main link between the theoretical or practical aims of investigation and the capacity to collect and analyse data that reflects the aims (Fraenkel et al., 2011). The research questions need to be fully developed prior to the final selection of a methodology or sample (Anderman, 2006). Fraenkal et al. (2011) proposed that research questions should be: (a) feasible, capable of investigation using available resources; (b) clear and specifically defined in operational terms; and (c) developed from definitions recognizable within the field of research.

Specific details of research questions should integrate descriptions of the sample, the variables, and the analysis. Sample information should pertain to demographics elements (e.g., boys or girls, country of origin etc.) or grouping characteristics (e.g., treatment or control, grade 8 or grade 9). Variables also need to be clearly articulated within the research question. In quantitative design the main types of variables (i.e., elements of the research that can vary) are (a) Independent-the cause (demographic characteristic; manipulated, treatment or experimental variable); (b) Dependent -the effect (measured or outcome variable); (c) Quantitative (variables measured as a matter of degree using real numbers, i.e. age, number kids); (d) Categorical (no variation... either in a category or not, i.e. gender, hair color); (e) Extraneous or uncontrolled independent variables (Fraenkal et al. 2011). Analysis needs to be generally framed within the research questions in terms of the types of findings you are expecting. For example, one group will be significantly higher than another group on the dependent variable as an outcome of the treatment or independent variable. Terminology will change because of the statistics that are used to answer the research questions.

In ESL research an example of a research question within a study that incorporates a treatment and a scored measure of language skills could be: “Significant differences in written English language scores will exist between a treatment group of first year university ESL students that receive access to an online story telling program, than first year university ESL students who participate in the standard program. It is predicted that the treatment group will score higher than the standard group”. Fraenkal et al. (2011) suggest that directional research questions (e.g., one group higher or lower than another) can be riskier to present and should only be used when researchers are confident of the expected findings. Otherwise, non-directional (differences between the two groups will exist) research questions should be presented.

Matching a Research Design Methodology With Aims

Irrespective of the type of issue or idea the study will attempt to clarify, strong connections are required between the research aims and the design and methodology of the research. De Vaus (2001) has noted that design and methods are not the same and must be considered independently, but within

the project, operate in unison. Decisions pertaining to characteristics of the sample, availability of resources, time to implement the project, and skills of the research team will influence the approaches that can be implemented within the project. Research in ESL is suitable to a variety of methods that operate within the different quantitative designs outlined earlier.

Knowledge of the general demographic characteristics of the sample prior to the study is important in relation to the capacity to undertake group comparisons based on factors such as age, gender, country of origin etc to ensure it is representativeness. For example, if the size of the groups to be contrasted is not relatively similar in relation to the independent variable a number of several important statistical comparisons cease to be viable. Early design phases of the project will substantiate the availability of resources. Experimental designs involving interventions or treatments can be costly, or require specific technical skills so the research team must take these types of factors into consideration before finalizing the project plan. Time constraints will affect whether the project can utilize a cross-sectional or a longitudinal methodology within a correlational or a causal-comparative design. Cross-sectional methods involve the collection of data from multiple groups within a single time period. Longitudinal methods require access to the same cohort of participants over an extended interval involving multiple phases of data collection (Lauer, 2004). Finally, quantitative designs necessitate specific skills from the within the research team which should be apparent or acquired over the course of the project. More complex causal or experimental designs typically utilize more advance statics and it is imperative that a member of the team has the knowledge to ensure the data is collected correctly, and their skills are sufficiently developed to effectively use these techniques.

Data Collection In EFL Research

Data collection in EFL research will require undertaking a series of important steps to ensure that the design is fully instigated and aims are achieved. In many areas of educational research, the setting in which the data is to be collected has a major influence on viable procedural strategies (e.g., primary school, university, language college). Researchers must also work through

tasks associated with sampling procedures and participant recruitment. Selection of instruments and interventions will require close review to ensure applicability to the samples recruited. Finally, the processes and procedures used to formulate a sample and administer the measures or treatment will require thorough detailing to support the fluidity of the project as a piece of useful research.

Evaluating the Setting

Quantitative data collection necessitates that the setting must be capable of facilitating access to the targeted groups. This can be either on-site or in the case of survey research and on-line environment can be created. The research team will need to negotiate with those who manage the settings (or create the on-line space) that participant access, working spaces, and equipment are available. Interaction regarding organisational and participant (or parent) consents must be undertaken as a prerequisite to the data collection commencing.

Recruiting and Selecting the Sample

The sample is the set of individuals that the researchers involve in the specific data collection, whereas, the population is the set of individuals to whom findings could be generalized (Dornyei, 2007). Recruitment and selection of participants should be equitable within the limitations of the procedural elements of the study. Researchers may not exclude participants on the basis of gender, race, national origin, religion, creed, education, or socioeconomic status.

Dornyei (2007) suggested that the procedures associated with formulating a sample can be considered in two main groups that are probability sampling and non-probability sampling. Probability sampling can be complex and often challenging to implement in research in ESL. The main probability sampling strategies include random, stratified random, systematic and cluster. The approach commonly adopted in The ESL field is non-probability sampling, which necessitates the acceptance of a level of statistical error in facilitating the selection of a representative sample that can be recruited

by the ordinary researcher. Procedures used regularly are quota sampling, dimensional sampling, snowball sampling and convenience sampling. Each of these procedures are often adopted because of the restraints that a research must operate within in educational settings. Participants are often targeted as a combination of the features that the individuals represent the target group, and are available to the research team. Interestingly, in line with this notion of availability and convenience, Dornyei proposed that determining the size of the sample should be considered in terms of not how large a sample should be but rather in terms of what is the minimum acceptable sample size that maintains the viability of the project.

Instruments and Interventions

Quantitative data collected in ESL research is more commonly derived from tests relating to language skills or course performance, or survey questionnaires that source information related to participant attributes or program characteristics. Less common is the implementation of a treatment or intervention to determine if it can serve as an agent of change.

Surveys and questionnaires. Developing and administering surveys and questionnaires is a very important data collection methodology within ESL research. The popularity of this approach is primarily due to the ease, adaptability, and generalizability with which it can be undertaken and results conveyed (Schutt, 2012). Griffiee (2012) provides a clear overview of how the development and use of questionnaires in language research should be managed. Firstly, in terms of administration select from the options of paper and pencil, on-line, or via telephone. Secondly, format must be considered to ensure that the components of demographics, closed-ended items, and/or open-ended items are included. Open-ended items are typically multiple choice or Likert type response formats. In ESL research, many projects will use a group-administered survey completed by individual respondents assembled together. Response rate is less of an issue in the group setting because most group members will participate, however, a concern is the possibility that respondents will feel coerced to participate and may not answer questions honestly (Schutt, 2012).

A survey questionnaire should be designed as an integrated whole, with each question and section serving some clear purpose and complementing the others. A standard introductory statement should be included that highlights appreciation, describes the steps of the survey, and reinforces that the survey is not the same as a test. Schutt presents an excellent summary of the basic strategy for developing both survey items and structure.

Questions must be worded carefully to avoid confusing respondents, encouraging a less-than-honest response, or triggering biases. Inclusion of “Don’t know” choices and neutral responses may help, but the presence of such options also affects the distribution of answers. Open-ended questions can be used to determine the meaning that respondents attach to their answers. Answers to any survey questions may be affected by the questions that precede them in a questionnaire or interview schedule. (p. 183)

Overall, when the research team decides to utilise a survey instrument for data collection they must also determine if existing measures with substantiated reliability and validity exist that meet the aims of the project, or there is the need to develop a new measure. If the latter is the case, researchers will also need to follow expected psychometric protocols associated with instrument development such as reliability, content validity, and construct validity. Data collection through the use of surveys is generally effective when the measure represents a coherent and accurate representation of an aim of the project.

Interventions. The use of interventions in experimental studies in ESL can strengthen the quality of the research, particularly in relation to specific teaching activities, programs, curriculum changes, or textbooks that purport to increase educational achievement and success. By identifying and empirically evaluating what is viable and effective in the management of teaching and learning in ESL through trialling specific treatments or program changes, evidence is generated to support the broader community of ESL practice. Gill and Hahs-Vaughn (2010) further proposed that by:

Determining the efficacy of an intervention on a local scale is not only practically useful—it can help one’s day to day teaching—but it is also important for the broader research community in that it may reveal conditions under which the treatment X works or does not work as well (p. 25).

Experimental treatments are typically more successful when the intervention is grounded in strong substantive theory, when implementation of treatment remains fully committed to that theory, the research setting is well managed, and when intervention is consistent over its delivery. In many educational field experiments these conditions are not met. Because teaching and learning settings are often large, complex, social organizations with multiple programs, disputatious politics, and conflicting stakeholder goals, the fidelity of the treatment programs can be highly variable (Shadish, Cook, & Campbell, 2002). At the simplest level experimental interventions in ESL that are well designed and managed can make a contribution by simply probing whether an 'intervention-as-implemented' makes a marginal improvement beyond other associated influences or variables. The generation of useful data as an outcome of interventions is only possible when a substantive effort is made by researchers to ensuring the quality of that treatment both prior to and during implementation.

Processes and Procedures

Adopting both ethical and viable processes and procedures in collecting data, is a critical component in addressing research questions or supporting hypotheses. Data gathering as previously discussed in this section incorporates consideration about what variables to investigate, the unit of measurements, participants of the study (population and sample), participant protections, procedures used for selecting participants, the methods and procedures used for data collection, and any reliability or validity of collection methods (Cresswell, 2004). Procedures do not necessarily occur independently of each other but should operate in a uniform and orchestrated manner. For example, recruiting participants and decisions associated with the measures or treatments of the research can and, possibly should happen in unison. Although, many of the procedures of research require different attributes (e.g., communication skills, statistical knowledge, academic acumen), good quantitative researchers are capable of establishing procedural frameworks that demonstrate a logical and integrated sequence of formulating, implementing, and finalising the essential tasks of their investigation.

Quantitative research in ESL should also be bound by these expectations of quality practice

Data Analysis in EFL Research

The purpose of this section is to provide a general but brief overview of the main types of quantitative statistics that could be adopted in ESL research. Interestingly, the types of statistics used in educational research have remained relatively consistent and tend to represent the simpler end of the complexity continuum (Karadag, 2010). An essential preparatory phase in data analysis is to ensure that following collection the data is coded, entered, screened and cleaned (Dörnyei, 2007). Data analysis will be considered within the categories of descriptive statistics (including correlation), inferential statistics, construct statistics, and model testing statistics.

Descriptive Statistics

All quantitative research requires the reporting of basic descriptive statistics. Descriptive statistics constitute the basis of all quantitative analysis and its contribution must not be overlooked because advanced statistics are also used in the investigation (Karada., 2010). The role of descriptive statistics is to summarize sets of numerical data in order to present a concise simple overview of the trends and patterns of the data (Dörnyei, 2007). Data when considered as variables can be classified as numerical, non-numerical, discrete, categorical, nominal, ordinal, interval scale, and continuous. The main univariate statistics reported should be considered within the three categories of distribution, central tendency, and dispersion (Trochim, 2006). The details of information provided for these statistics are based upon descriptions by Ary et al. (2006) and McDonough & McDonough, (1997).

Distribution. The distribution is a summary of the frequency or occurrence of individual values or ranges of values for a variable. It is considered as a representation of the layout of the data by highlighting where scores occurred and how the data is spread. For example, the distribution of gender composition at different year levels of an ESL course, or numbers of individuals who score at different grading values for a particular test. The two

main presentation formats are tables that detail the percentage or number of occurrences for a variable and formulated categories (e.g., nationality, age group) and a frequency distribution chart such as histogram or line graph that presents a figure format of the data pattern.

Central tendency. The central tendency of a data set is an estimate of the “center” of the distribution of values. The three major types of estimates of central tendency are mean, median, and mode. The mean is the average of all scores from the sample; the median is the middle score of the range of the set of scores; and the mode is the score that occurs most frequently within the set of scores. An example of how central tendency data is shown is a set of 20 grammar test results in which 10 questions of 1 mark each were asked, and descriptive results indicated the mean was 6.6, the median was 7, and the mode was 6.

Dispersion. Spread of scores or dispersion of scores is typically reported as range, standard deviation and relative position. The range is basically the detail associated with the lowest and the highest score of the set of scores, and is calculated by subtracting the low score from the high score and adding one. Standard deviation measures how much on average individual scores of a given group vary or deviate from the mean score for this same group, or the average of the differences of scores from the mean. The main statistics that demonstrate relative position or where a score lies in relation to the other scores in the sample are z score, stanines, and percentile rank. A z score is defined as an indication of positive or negative difference above or below the mean. Stanine scores are a system of detailing Z scores in a format from 1-9 rather than negative to positive. A percentile rank indicates the percentage of scores from the set of scores that are above or below a given scores. If we consider the test example provided for central tendency, the range was 6, the standard deviation was 1.37, the z score spread was -2.2 to +2.2. In terms of percentile, if we refer to a score of 6, then 35 per cent of scores were at or below this score and 65 percent were above.

Descriptive statistics constitute a necessary and important element of the analysis procedure within the quantitative approach. The resultant data not only provides the basis for additional complex analyses, but just as importantly generates statistical information for the researcher to make

judgements on the trends and patterns observed. For example, in a study examining final test scores of students completing EFL programs delivered in their country of origin compared with those who complete the program in a country where the first language is English, descriptive statistics allow the researcher to make initial commentary on how the trends pertaining to subgroups might generally differ. This could be illustrated through contrasts such as women may be performing better than men locally, older participants scoring lower when away from their country of origin, individuals with multiple language backgrounds score highest overall etc. In general, descriptive data serves as the succinct commencement point in considering how the data could be further explored or broadly communicated. The reporting of descriptive results also regularly utilises presentation formats such as tables, charts, and graphs as an illustrative framework for detailing trends and general contrasts.

Inferential Statistics

Inferential statistics are associated with formulating predictions or inferences about a population from observations and contrasts of sample data. These analyses utilise tools that allow researchers to estimate a level of confidence in inferences “that phenomena observed in samples will also appear in the populations from which the samples were drawn” (Ary et al., 2006, p. 210). Inferential results extend on the trends generated within descriptive findings by determining if differences between groups or test occasions are statistically significant. Statistical significance represents the likelihood that a result or relationship observed in a sample is caused by something other than random chance and therefore generalizable to the whole population (Dörnyei, 2007). Statistical hypothesis testing is employed to determine if a result is statistically significant or not and provides a “p-value” representing the probability that random chance could explain the result. A figure of a 5% or lower p-value is considered to be statistically significant.

The main set of basic inferential statistics includes t-tests, Analysis of Variance (ANOVA), Chi-square test, and correlation. The t-test procedure is used to compare two groups or a pair of repeated scores from a single group (e.g., males and females, semester one exam and semester exam). ANOVA is

a similar contrast that is used when there are more than two groups or repetitions of a score to compare (e.g., low, middle, high income groups; language skill scores before, during, and after course completion). Chi-square analysis involves the examination of differences between proportions or frequencies of nominal or categorical data (e.g., yes or no scores contrasted with country of origin). The correlation procedure considers the relationship between variables, scores, or categories rather than the differences between the values. Results of correlational analysis provide an indication of the strength and direction of the association between scores or variables (e.g., strong positive relationship between years of previous English language experience and EFL course performance).

Advanced inferential statistical procedures include Multi-level analysis of variance (MANOVA), Analysis of co-variance (ANCOVA), and regression. MANOVA is a group comparison technique that examines multiple variables and multiple groups. The procedure could be implemented in an investigation of English reading and writing skill scores for EFL course students and contrasted against gender or year level in the program. Application of ANCOVA analysis is undertaken when the researcher is aware that differences in scores on a variable of interest exist prior to the implementation of a treatment. For example, students from a range of countries of origin are completing an on-line English language skills enhancement program. The researcher is interested in the efficacy of the program in improvement of written English skills. It was observed that the pre-treatment scores varied substantially between cohorts according to their first language background. The ANCOVA procedure controls for these differences to ensure that post test comparisons are a consideration of the effect of the training program and not simply the influence of pre-treatment written English skills. Regression analyses are an extension of correlation that involve the examination of the relationship between multiple variables. The technique determines the capacity of one variable or scores to predict the scores on subsequent variables or scores. For example, a regression analysis could be used to evaluate if the selection exam for an EFL studies program is more effective in predicting scores on the final year written exam or students' scores on the National English language skills test.

Construct and Modelling Analysis

The process of data collection in EFL research will often involve the development and use of measures that are designed to assess a construct or set of attributes to represent that construct. Further to this, researchers will on occasions formulate studies that attempt to determine the mediating influence or linked association of one score or variable on others measured at a similar time or within a predetermined sequence. These outcomes can often be achieved through the use of statistical procedures associated with factor analysis and structural equation modelling.

The two main forms of factor analyses are Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Both forms are used to determine the factors that items within a questionnaire or survey could represent. In general, EFA is used when the researcher has developed a measure and serves to provide an indication of which items group together to form factors. Factors are considered to be linked to the construct on which the development of the measure was based. For example, in EFL research if a measure was developed to determine how students perceive the importance of their overall set of English skills (i.e., reading, writing, communicating) it can be analysed to ascertain if these domains around which the items were based exist as factors or are grouped in another pattern. If all the items related to reading, writing, and communicating skills group together, respectively, the measure could be considered to demonstrate a level of logical coherence. In subsequent studies in using the measure with different EFL populations, researchers could utilise CFA to evaluate that the factors proposed as an outcome of the EFA can be identified as an a priori model of the construct and factors of the measure. The use of both EFA and CFA are critical in the establishing the validity and viability of both newly developed and existing measures.

Structural Equation Modelling (SEM) is a statistical technique that allows investigators to consider a hypothesised model of relationships between variables of interest to their research question. The researcher constructs an a priori model, or path diagram, of possible associations and then uses the SEM to test the adequacy of the fit of these variables. For example, a project could be undertaken in which a path model is constructed to consider

the predictive capacity of English language skills score, high school performance, and frequency of international travel on undergraduate performance in an ESL course. This score could then be linked with further study score performance at the post-graduate level, and then considered against relationships with post-graduate outcomes such as published output, scholarship application success, and ESL teaching opportunities (see Figure 1). This path model can also be evaluated on the basis on mediation effects of demographic variables such as age or gender. The broader aim of SEM and path analysis is to work towards the incorporation of available data to support or refute hypotheses associated with areas of interest to the field. For more in depth overviews of SEM and path analysis please see Hair, Black, Babin, and Anderson (2010) and Tabachnick and Fidell (2001).

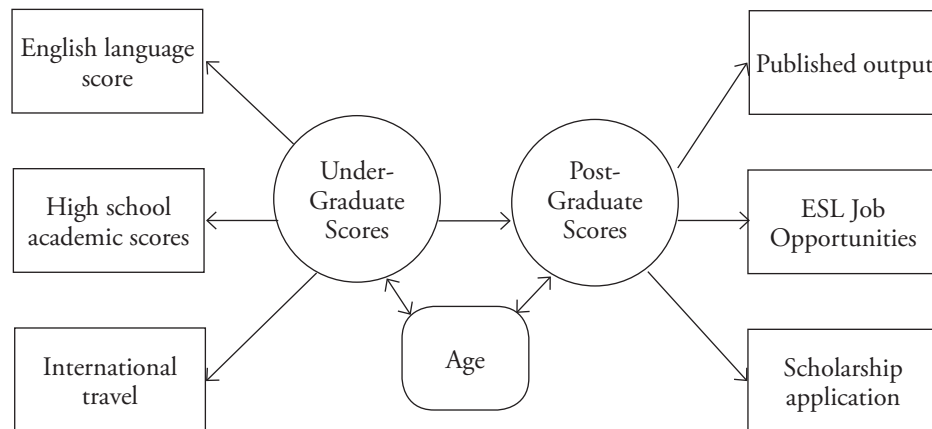


Figure 6.1 Hypothesised model of Post-graduate success in EFL studies (Tabachnick & Fidell, 2001)

Data Interpretation and Reporting

A critical phase in the quantitative research process is the interpretation and reporting of findings. It is accepted that a specific approach for detailing numerical results is required, generally involving the adoption of a pre-defined manuscript presentation design and a set of statistical styles as outlined in procedural texts for academic publishing such as the American Psychological Association (APA) publication manual (American Psychological Asso-

ciation, 2009) or the reporting chapter of Dornyei (2007). The typical elements of a quantitative research report adopting APA style include: Title page, Abstract, Introduction, Method, Results, Discussion, and References. As a broad guide to framing interpretation and reporting phase of a research project this section of the chapter will cover the following themes of results as numbers, contextualizing the numbers, and establishing meaning from the findings.

Numbers as Results

Quantitative statistics as described in previous sections of this chapter are first reported in a numerical context. The researcher should describe the patterns of descriptive data and outcomes of inferential and modeling analyses in a manner that accurately reflect the data, and strictly follow the underlying assumptions of the statistical procedures utilized. The results section need only present the information that will be absorbed into later discussion. If findings are detailed in a table then there is no need to repeat that information in the text, unless that information extends or refines the table results (Dornyei, 2007). A blend of precision and brevity should be an aim in the formulation of the results section. The detail contained within the numbers is important, and readers with basic experience in quantitative analysis should be able to work from that information. There is little need to extend numerical descriptors pertaining to the analyses into a text based overview of what the numbers already reveal. For example, if the analysis is related to a comparison of two groups on several variables associated with English skills, then in a results section focus on a table of resultant data and briefly summarize any required inferential detail related to significant group differences. I recommended “The Owl at Purdue Writing Lab” as an excellent site to guide both structuring of results and how this is integrated into the manuscript (Purdue Owl, 2014).

Contextualizing the Numbers

Descriptions of quantitative information should be focused on locating the resultant data relative to the research questions, and as a representation of

the population of interest. McDonough and McDonough (1997) have proposed in relation to EFL research that “the researcher has to go back to the design and purpose of the research to decide what the result actually means” (p. 115). For example, a set of exam results in which the average score is very high could indicate that either the test is too easy or that the teaching was of exceptional quality and the students mastered the topic content. Depending on the possible reason for considering test scores as data, such as to determine the qualities of a new teaching technique, or two compare course performances of males and females then a high score as a result would have contrasting meanings. As an outcome of a new teaching technique this would be a very pleasing result, whereas in identifying possible differences between males and females it would add limited new information.

In considering and interpreting inferential results, attention should be directed toward the meaning of both significant and non-significant results. Using the example of pre-requisite skills that might contribute toward end of course results in an EFL program of studies, significant differences between groups classified on the basis of how many additional languages they speak could be influential in establishing selection criteria for courses (i.e. speaks more than two languages). In contrast, no differences between course outcome results when programs are delivered on-line or on-site could serve as evidence to administrators that off-campus teaching is equally as effective as face to face delivery.

Overall, it is essential in quantitative reporting that the numbers do not hide the meaning from all but those who have higher order statistical knowledge (Lazaraton, 2005). When the results are considered by the reader, the complexity of the presentation should not operate as a deterrent to those individuals with an interest in the research questions and context but are restricted in interpretation by a limited quantitative history. The framing and presentation of the numbers and the statistics in that manner may lead to the reader to not fully engage in the discussion sections of the manuscript.

Establishing Meaning

The final phase in the quantitative data collection and reporting process is to ensure that details regarding numerical and statistical patterns are inter-

preted in relation to existing research, and receive logical consideration on the part of the researcher. Statistical patterns generated as an outcome of any investigation cannot exist in isolation, and must be framed against the results reported in similar studies, connected to relevant theory, and coherently linked with the original research questions. If a researcher is interested in determining if linguistics progression rates of students located in Asia demonstrate any discriminating characteristics then these findings need to be considered in relation to similar studies in Asia but also investigations based in Europe and first language English speaking countries. Highlighting the contrasts and similarities in the patterns observed with earlier research, and considering the associations the other researchers have proposed in relation to theory will facilitate a framework to support the implications drawn in relation any new research. The consideration of the research questions is best managed when the logical inference perceived as an outcome of the descriptive and inferential analyses is clearly linked with current and established ideas accepted within the EFL research domain.

Overall, findings can most effectively generate a rationality of meaning when the researcher and the readers share, and acknowledge (not necessarily accept), matching or comparable assumptions (Chu, N.D.). It is critical that the manuscript or report serves as the primary source to explain your assumptions and how these connect to your data and findings. When the quantitative research project is framed upon a solid theoretical foundation it facilitates greater opportunity to generalize your findings. Chu proposed a simple but accurate point to be conscious of in reporting your research in that “as you extrapolate or generalize, you must examine your assumptions in order not to exceed the boundaries of your framework or the quality of your data”. (see <http://digitalcommons.unl.edu/libphilprac/advice.html>).

Summary

Highlighting the perceived advantages of the quantitative analyses methodology such as precise conceptualization and measurement, sophisticated statistical analyses, and defined reporting procedures (Voils, Sandelowski, Barroso, & Hasselblad, 2008) may serve as a stimulus for new researchers to consider this approach. However, these advantages will only be of benefit if

the researcher is organized in the pre-data collection phases of the investigation. Deciding upon the type of design to be implemented and connecting this to the broader planning associated with the aims of the research is critical. Data collection procedures in relation to setting, sample, and instrumentation must be closely matched to the research design. For example, awareness of the requirements associated with intervention period and participant numbers (i.e., quasi-experimental), will necessitate different strategies than a large scale survey project in which correlational and descriptive data will be incorporated. Subsequent data analyses will be of greatest value when the original design and planning is strong. New researchers should seek as much information as possible regarding the wide variety of techniques that can be used to analyse the data after collection. Although current analysis software provides ample opportunity for the researcher to perform a variety of descriptive, inferential, construct, and modelling analyses, it is recommended that the researcher must be confident that they can accurately interpret the resultant output. Finally, irrespective of design, procedure, and analysis, results of any study in the EFL will only be of value when interpretation is coherent and located within the broader body of literature that the research community supports.

Dr. Anthony Watt
College of Education
Victoria University, Australia

References

- American Psychological Association. (2009). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.
- Ary, D., Jacobs, L. C., Razavieh, A., & Sorensen, C. (2006). *Introduction to research in education* (7th ed.). Belmont, CA: Thomson Wadsworth.
- Anderman, E. (2003) *Research methods: An overview*. The Gale Group <http://www.education.com/reference/article/research-methods-an-overview/>
- Berry, J. (2005). *Quantitative Methods in Education*. Research Centre for Teaching Mathematics, University of Plymouth. <http://www.edu.plymouth.ac.uk/resined/quantitative/quanthme.htm>
- Brown, J. D. (2011). Quantitative research in second language studies. In E. Hinkel (Ed.), *Handbook of research on second language teaching and learning* (Vol. 2) (pp. 190-206). New York: Routledge.
- Chu, F. T. (N.D.) *Factors you can control when writing a paper*. Malpass Library Western Illinois University. <http://digitalcommons.unl.edu/libphilprac/advice.html>
- Creswell, J. W. (2004). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*, 2nd ed. Columbus, OH: Prentice Hall.

- Cronbach, L. (1976). *Research on classrooms and schools: Formulation of questions, design and analysis*. Stanford Univ., Calif. Stanford Evaluation Consortium.
- Crookes, G., & Lehner, A. (1998). Aspects of process in an ESL critical pedagogy teacher education course. *TESOL Quarterly*, 32(2), 319-328.
- De Vaus, D. A. (2001). *Research design in social research*. London: SAGE.
- Dimsdale, T., & Kutner, M. (2004). *Becoming an educated consumer of research: A quick look at the basics of research methodologies and design*. Meeting of the Minds Practitioner-Researcher Symposium. American Institutes for Research, Sacramento, CA.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.
- Fraenkel, J., Wallen, N. & Hyun (2011). *How to design and evaluate research in education*. New York: McGraw-Hill
- Gall, M. D., Borg, W. R., Gall, J. P. (2003). *Educational research: An introduction*. (7th Edition). White Plains, New York: Longman.
- Gay, L., Mills, G., & Airasian, P. (2006). *Educational research: Competencies for analysis and applications*. New Jersey: Pearson Education, Inc.
- Gill, M. & Hahs-Vaughn, D. (2010). Does it work? A guide to investigating the efficacy of interventions in educational research. *Current Issues in Education*, 13(4), 1-32.
- Griffiee, Dale T. (2012). *An introduction to second language research methods: Design and data*. Berkeley, CA: TESL-EJ Publications.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Johnson, B. (2000). It's (beyond) time to drop the terms causal-comparative and correlational. Research in Education ITFORUM PAPER #43 <http://itech1.coe.uga.edu/itforum/home.html>
- Karadağ, E. (2010). An analysis of research methods and statistical techniques used by doctoral dissertation at the education sciences in Turkey. *Current Issues in Education*, 13(4), 1-21.
- Lauer, P. A. (2004). *Policy maker's primer on education research. How to understand, evaluate, and use it*. Aurora, CO: Midcontinent Research for Education and Learning (McREL) and Denver, CO: Education Commission of the States (ECS). Available from: www.ecs.org/researchprimer.
- Lazaraton, A. (2005). Quantitative research methods. In E. Hinkel (Ed.), *Handbook of research in second language learning* (pp. 209-224). Mahwah, NJ: Lawrence Erlbaum.
- McDonough, J. & S. McDonough. (1997). Principles and problems—what makes good research? *Research Methods for English Language Teachers*. Arnold: London.
- McEwan, Patrick J. (2008). Quantitative research methods in education finance and policy. In Helen F. Ladd and Edward B. Fiske (Eds.) *Handbook of research in education finance and policy*, (pp. 87-104). New York: Routledge.
- Purdue OWL (2014). *The writing lab & the OWL at Purdue*. Purdue University. <https://owl.purdue.edu/owl/> retrieved Jan 2014.
- Schutt, Russell K. 2012. *Investigating the social world: The process and practice of research*, 7th ed. Thousand Oaks, CA: SAGE Publications.
- Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.
- Trochim, W.M. (2006) *The research methods knowledge base*, 2nd Edition. <http://www.socialresearchmethods.net/kb/> Retrieved May, 2013.
- Voils, C.I., Sandelowski, M., Barroso, J., & Hasselblad, V. (2008). Making sense of qualitative and quantitative findings in mixed research synthesis studies. *Field Methods*, 20 (1), 3-25.