Evaluation Research: A Comparative Analysis of Qualitative and Quantitative Research Methods

Zanyar Nathir Ghafar[1*]

[1]Interior Design Department, Bright Technical and Vocational Institute, Sulaymaneayah, Kurdistan Region, Iraq

Abstract: Researchers often use a mixture of qualitative and quantitative approaches drawn from various fields. It is possible to record or quantify quantitative data. Qualitative data are descriptive, linguistically related, and rely on interpretation. Quantitative data provides information such as how many, how much, or how often something occurs. On the other hand, qualitative data could assist us in comprehending the processes, causes, and occurrences of activities. Each approach has many advantages as well as some disadvantages. This research aims to evaluate the various language testing and assessment approaches and compare and contrast their benefits and drawbacks. These advantages included the opportunity to learn more about the development, administration, and interpretation of assessments and tests and the capability to research test-takers behavior, perceptions, attitudes, and understanding. The study's findings showed that the amount of time required gathering data using quantitative research techniques is much less than the amount of time required collecting data using qualitative research methods. Quantitative research methods, on the other hand, have several drawbacks, such as the difficulty of explicating what a researcher means when using a specific word and the fact that these methods can only provide an overarching comprehension of the topic being studied.
Keywords: Benefits, drawbacks, testing, evaluation, qualitative, quantitative.

Review Paper

*Corresponding Author:
Zanyar Nathir Ghafar
Interior Design Department, Bright Technical and Vocational Institute, Sulaymaneayah, Kurdistan Region, Iraq

How to Cite this Paper: Ghafar, Z. N., (2023). Evaluation Research: A Comparative Analysis of Qualitative and Quantitative Research Methods. *Middle East Res J Linguist Lit, 3*(2): 25-32.

1. INTRODUCTION

Understanding the techniques used in producing a scientific publication, such as a manuscript, dissertation, or work to complete a university course, requires an instrument known as a scientific methodology. This tool is of crucial significance for doing so. Studying the procedures or apparatus essential to producing an academic piece is part of the scientific methodology. According to Flanagan (2013), the scientific method is the most effective way to develop new hypotheses, investigate current theories, and put those ideas to the test via experimentation. A scientific investigation is carried out this way by conducting rigorous and careful inquisitions to find and examine the facts in a given reality. Research in the scientific community may be quantitative, depending on the approaches used. Qualitative research aims to convey a topic in statistical terms but rather to understand it better. When doing qualitative research, the researcher often acts as both the subject and the object of his or her study.

The purpose of the qualitative method is to give detailed and illustrative data in order to get an understanding of the different characteristics that are associated with the topic. The scope of qualitative research is expansive, and it elucidates the workings of social relationships by focusing on intangible aspects of human experience that cannot be measured. According to Maxwell (2012), the qualitative study investigates the world of meanings, motivations, aims, beliefs, values, and attitudes, which translates to a more profound realm of interactions, processes, and occurrences that cannot be reduced to fundamental components.

As an additional possibility, quantitative research might quantify the data. The findings provide a thorough and adequately complete perspective of the whole society (Martin & Bridgman, 2012). This is because the samples are often extensive and are believed to represent the general population. The investigation and generalization of the uncovered findings are greatly aided by mathematics and statistics, both of which play an essential role. Quantitative research, which is oriented

Peer Review Process: The Journal "Middle East Research Journal of Linguistics and Literature" abides by a double-blind peer review process such that the journal does not disclose the identity of the reviewer(s) to the author(s) and does not disclose the identity of the author(s) to the reviewer(s).

25

on objectivity, is particularly well suited for situations in which it is possible to deduce conclusions from population percentages and gather quantifiable measurements of variables. The gathering of data in quantitative studies often involves the use of predetermined protocols and predefined apparatus. The information is gathered thoughtfully and fairly.

Finally, statistical techniques are employed to examine the numerical data, and software packages like SPSS, R, and Stata are often used in this process. Qualitative and quantitative Research processes and methodologies are often used in various academic disciplines, including sociology, psychology, history, etc. When academics from different schools of positivism and interpretivism strive for dominance, there are continuous "paradigm wars" about research methods. Positivism is a school of thought in the academic world that believes that the social world comprises an unchanging material reality that can be measured. Interpretive scholars, on the other hand, claim that reality is a social construction constructed by people that can be modified and understood subjectively, thus contradicting the positivist view of reality (Corbetta, 2003; Marcon & Gopal, 2005; Kroeze, 2012). These scholars point out that reality may be perceived differently depending on the person doing the understanding (Kroeze, 2012).

In order to undertake language testing and assessment research, this study will examine the advantages and disadvantages of employing quantitative and qualitative procedures and approaches. The historical examination of research methodologies and procedures (both of them) is the focus of the investigation's first section. Following a brief overview of the language, tests and evaluations that are accessible, as well as the drawbacks of employing qualitative and quantitative methodologies, are discussed and the research methodologies that are the most used in language difficult and estimation are reviewed. Concerns of an ethical nature are also addressed as a last point.

2. Assessing One's Capability in a Language through Testing and Evaluation

A test may be described as anything that indicates a person's degree of competency or incompetence, ability or incapacity, and where they fall on a scale that includes categories such as fail, pass, average, acceptable, outstanding, and remarkable. This is the broadest possible definition of the term. Whether or not a student should be permitted to advance to the next level is one of the most significant choices that can be made with an academic exam. It enables us to analyze a student's progress, whether or not the student needs assistance, and compare the performance of other pupils. In addition, the examination is an essential tool for the execution of public policy, such as ensuring that only students who have the highest test scores are permitted to advance to the next level of education by distributing nationally standardized tests of the same level of difficulty all over the country (Douglas, 2014). This can be accomplished by administering national tests of the same difficulty level everywhere in the country. In this respect, the examination functions as a crucial tool for the execution of governmental policy. One such instrument is the necessary examination for admittance into higher education programs, yet, this test, loaded with a significant level of risk, must be taken. Evaluators of a candidate's performance on a language test are concerned with the extent to which the candidate's findings accurately reflect the candidate's level of expertise in a particular domain, such as reading, writing an argumentative piece, vocabulary, or conversing with other people (Weir, 2005). This is because the evaluators want to know how well the candidate can read, write an argumentative piece, vocabulary, or converse with other people in the target language. Standardized tests, in any industry, including education, are intricately woven into the fabric of society, according to Fulcher (2010). The examination of language skills is standard. However, one must pass a language exam to progress in their job, become a citizen of the United States of America, move to another nation, or seek refuge in another country. Because of this, education is a necessary component that contributes to the formation of how individuals live in society.

According to McNamara and Roever (2006), standardized testing of languages has the unintended consequence of legislating what topics should be taught in schools. This finding comes from the researchers' examination of the phenomenon. According to the information presented up to this point, language assessments play a significant role in the lives of many individuals (McNamara, 2000). Despite this, many individuals have put significant importance on the reliability and correctness of language testing. For instance, Fulcher and Davidson (2007) said that test validity is emphasized in some capacity throughout every book and publication about linguistic evaluation. This is because the fundamental idea behind testing and evaluating is known as the validity of the test. As a direct result of this, it is clear that many different ideas are involved in evaluating and assessing a person's language skills; as a result, the study of education must include it as an essential issue.

3. Methodologies and Strategies for Qualitative Research

Given that it is the primary focus of this section, defining qualitative research is of the utmost importance. However, doing so may take much work. Ritchie *et al.*, (2013) noted that it needs a distinct set of procedures or practices solely (Denzin & Lincoln, 2011) and a distinct theory or paradigm. This word refers to a wide range of approaches and strategies that may be used in several fields of study. As a direct result, the writers have offered unique qualitative research descriptions. For instance, Strauss and Corbin (1990, p. 11) declared "When we say 'qualitative research,' we mean any kind of study that

doesn't use statistical methods or other ways to count things." It can include studies of people's lives, experiences, opinions, and feelings, as well as studies of how groups work, social changes, cultural events, and exchanges between countries.

Additionally, it may refer to studies on international interactions. This shows how different realities may be integrated via qualitative research rather than statistical analysis. Second, qualitative research is "keen on examining personal significance or socially generated creation of issues, events, or practices by collecting non-standardized data and analyzing words and pictures rather than quantity and statistics," as stated by Flick (2014, page 542). The influence of this concept on people's perceptions of the world was significant. As a result of this, there are a few different components that are associated with qualitative research. In addition, qualitative research "is multi-method in concentration, including a naturalistic, interpretive approach to its subject matter," according to Denzin and Lincoln (1994, p.). According to Van Maanen (1979, page 520), the definition of qualitative research is "an all-encompassing term wrapping an array of methods of interpretation which seek to describe, decode, translate, and otherwise come to the context with the meaning, not the degree of occurrence, of certain more or less naturally occurring phenomena in the social world." Van Maanen provided this definition. This word makes it abundantly evident that several points of view are taken into consideration in qualitative research. Because it considers both optimistic and pessimistic points of view, qualitative research is a broad notion that may include many different difficulties.

4. Benefits of the Qualitative Study

Firstly, using qualitative research techniques and procedures may be advantageous in several different ways. According to Denzin (1989), the initial outcomes of a qualitative research technique are an in-depth and thorough account of the thoughts, participants' opinions, and experiences of the part and an analysis of the factors that lie beh behavior. For instance, Bachman (1998) demonstrated in his study that qualitative research outcomes demonstrate the precise and in-depth link between information processing and performance regarding language examinations. He found that the findings of the research reveal this relationship. According to Chalhoub-Deville and Deville (2008), qualitative methodologies are used so that concerns concerning the development, implementation, and interpretation of language assessment may be better understood.

Secondly, one school of thought contends that the interpretive approach to qualitative research is the one that most accurately portrays the human experience in a given set of conditions. According to Denzin and Lincoln (2002), for example, qualitative study is a discipline that incorporates a wider variety of epistemological viewpoints, research methodologies,

and interpretative tools to comprehend human experiences. This makes qualitative study a multidisciplinary area. Researchers in the field of language assessment have started to use qualitative research methodology in order to concentrate on issues that require an analysis of content-related variables, such as the influences of test-takers characteristics on performance on tests, the techniques used for finishing tasks related to evaluation, and so on (McNamara, 2001). This shift in focus has allowed researchers to concentrate on issues that require an analysis of content-related variables. According to McNamara (2001), any evaluation of a language from a philosophical position has to consider the history, culture, and values of the society in which it was utilized.

Thirdly, the research uses an interpretivism strategy, an ideographic technique that investigates particular situations or occurrences (Kelin & Myers, 1999). This approach can comprehend various individuals' perspectives, meanings, and happenings. As a consequence of this, the importance of individual occurrences is the source of information that is used by this tactic (Richardson, 2012). According to Lazaraton and Taylor (2007), qualitative research methodologies in language testing investigate not only the behavior of the interviewer and interlocutor's behavior but also the participants' conduct and the impact of culture on the participants' behavior.

Finally, the framework of a qualitative research design that uses an interactive method is adaptable (Maxwell, 2012) since it may be developed and rebuilt higher than in other types of designs. According to Flick (2011), qualitative research methods have the potential for in-depth and accurate evaluations of a subject because they allow participants to determine what elements of the study are most important to them. As a direct consequence of this, difficult situations are pretty straightforward.

For instance, due to the dynamics of such settings, researchers have found that a variety of various related to their primary study area may be able to alter the behavior of students' behavior enrolled in courses. This was discovered as a result of the fact that such settings are so dynamic. When seen from this angle, a qualitative research approach is essential for recording the processes mentioned above. Because of the complexity of the methods involved in language evaluation, Mohan (2012) suggested using a qualitative research approach. As a result, qualitative research may be helpful in better comprehending the many aspects of language evaluation. The disadvantages of using qualitative research methods and strategies will be discussed in the following section.

5. Disadvantages of Qualitative Study

Despite the benefits discussed above, a few drawbacks must be ignored. First, according to

Silverman (2010), many qualitative research methodologies neglect contextual knowledge in favor of putting more emphasis on meanings and experiences.

For example, the phenomenological approach (Wilson, 2014; Tuohy *et al*., 2013) seeks to discover, examine, and better understand what the participants have gone through. In a manner analogous to this, Cumming (2001) placed participant experiences above any pressing concerns about the situation. He featured people who taught writing and the inhabitants of their respective nations from different countries," This meant that he was not attempting to represent educators working in particular nations or institutions. The examination's situational effect is not considered in this study since the emphasis is on the talent of the writing instructors while overlooking, for example, the effects of these nations' cultures on writing evaluation. This is because the research concentrates on the ability of the writing instructors.

Secondly, it is unlikely that the authorities will take the outcomes of a qualitative approach seriously. According to the findings of Sallee and Flood (2012), national and state officials in the United States educational system measured the performance of both teachers and students (Ravitch, 2010). Furthermore, quantitative approaches are typically accorded greater weight in various social science fields (Berg, 2009), as stated previously. Additionally, qualitative research is the only kind in which the social and cultural construction of the variables are being studied (Stuhards *et al*., 1994).

Thirdly, when it comes to research techniques, having a smaller sample size raises the question of whether or not the findings can be generalized to the whole study population (Harry & Lipsky, 2014; Thompson, 2011). Lam (2015) conducted a qualitative investigation on the language testing assessment in Hong Kong. She pointed out that due to the small sample size of the research, the findings could not be broadly extrapolated to various settings.

Richards *et al*., (1994) state that if this occurs, the interpretation and analysis of the data may become more challenging or convoluted. According to Berg and Lune (2012, page 4), "Qualitative investigation is a lengthy hard road, with mysterious data on one side and high standards for interpretation on the other." According to Darlington and Scott (2003), it is more challenging to transform a developing topic into a shape that is amenable to investigation, and the difficulties with data interpretation and analysis in qualitative research may persist for the whole study. The researchers Rumsey *et al*., (2016) focused more on the theme analysis than they did on the research question, language assessment, or language requirement. Nevertheless, in subsequent data analysis, the information was employed again to answer a fresh topic that had yet to be handled in the first study. This particular query had yet to be tackled in the initial research. When analyzing the data and creating the research subject while using the same data already obtained, conducting a qualitative study is more complex and continuous.

When everything is said and done, assessing the instances is time-consuming, and there are few ways to generalize the findings to a broader population (Flick, 2011). For example, if a lawmaker has to cast a vote on a particular issue, they must immediately start qualitative research to be finished in thirty days (Sallee & Flood, 2012). Like this, policymakers may seek quantitative rather than qualitative research in language testing and evaluative research to expedite the production of brand-new rules.

6. Strategies and Methods for Conducting Quantitative Studies

In Bryman's definition (2012, P. 35), quantitative research is described as the relevance of quantitative research should be considered, as this implication suggests. The answers to how many, how much, and to what degree serve as the basis for this study approach (Rasinger, 2013). In order to put it another way, while using this strategy, a considerable amount of attention is devoted to various aspects of the social environment or the components that make up the social environment. Payne *et al.*, (2004, page 180) state that the world of society can be broken down into empirical components, which researchers may obtain via stimuli and systematic measurement. The study of society then moves on to include these components. These tactics often use deductive reasoning and focus on pursuing regularities in human lives as their primary objective. Additionally, Bryman (2012) defined positivism as the research that leads to behaviors that are either general or legal. According to Richardson (2012), the positivist approach to knowledge development involves evaluating hypotheses with evidence. However, when the investigators assert that knowledge is gained via the interpretation of events and that interpretive thinking itself takes place (Kelin & Myers, 1999). (Richardson, 2012) is the antithesis of positivism in qualitative research. This is the case when the researchers state that knowledge is gained by comprehending experiences' significance. Even if Weber (2004) asserts that the contention between positivism and interpretivism is without foundation, it is again clear from this discussion that academics continue to be divided into these two camps. According to Bryman (1988), the phrase "quantitative methods" is a catch-all term that, like the term "qualitative methods," may be used for a range of distinct research strategies. This is because both terms refer to types of study that may be qualitative or quantitative. The following list discusses the advantages and disadvantages of doing quantitative research.

7. Some Benefits Associated with Using a Quantitative Method

It is feasible to extend the quantitative results to either the whole population or to a particular subset of the population thanks to the larger sample size and the random selection of participants in the study (Carr, 1994). According to Connolly (2007), using statistical software like SPSS makes it feasible to do data analysis more timely. More timely the fact that using such software speeds up the sampling process. Researchers by the name of Powers *et al.*, (2015) looked into TOEIC tests, and they were overjoyed to find that the study sample (2300) accurately mirrored the majority of the individuals who took the TOEIC test. The reliability of the findings was therefore improved as a consequence of this. According to research by D. Powers and A. Powers (2015), most individuals who sat for the TOEIC exam have a minimum of a bachelor's degree. According to Kauber (1986), the positivist method of assessing components must be the foundation for quantitative research. This is because the positivist approach favors objectivity. For example, Carroll and Bailey's (2016) research on the assessment of language competency found that there are a variety of criteria; include EFL and non-EFL students and assessments in communication, writing, reading, and listening. Include native speakers of the target language and additional languages. In addition, the findings of the research showed that there are several different criteria. When researching the level of fluency possessed by individuals who spoke a second language, Préfontaine, Kormos, and Johnson (2016) considered various aspects.

8. Disadvantages of the Quantitative method

There are drawbacks to doing quantitative research besides the benefits that were just discussed. Traditional explanations of social phenomena are not permitted under the positivist research paradigm. In addition to this, it ignores essential truths and answers that are more deep. Bouwer, Béguin, Sanders, and van den Bergh (2015) employed various variants, a range of genres, and random mistakes to study how genre influences writing scores. The investigation did not go into great length on the impact's causes or what they signified in relation to that particular event. Utilizing quantitative research methodologies has additional drawbacks. in positive psychology is that it does not consider how individuals evaluate their behavior and the behavior of others (Blaikie, 2007). Studies on quantitative testing of languages have demonstrated that these tests only look at and evaluate a candidate's linguistic talents, competency, score, and other criteria. Katzenberger and Meilijson (2014) researched the evaluation of the Hebrew language for preschoolers. The study's goal was to discover whether or not the examination can differentiate between growing children and those with phrase-impaired speech. The research did not investigate the factors contributing to the ease with which confident children learn a second language compared to other children. In other words, the research did not consider the truth or the children's impressions about their ability to learn Hebrew. It does not matter whether the person in the picture was at their best or incredibly disorganized; it merely considers those attributes at that time (Schofield, 2007). This is another disadvantage of quantitative research that tends to capture happenings in a single instant. For example, to demonstrate the statistical validity of the tests, the research conducted by Fidalgo *et al.*, (2014) included an additional 400 individuals. Because of this, it has only been able to provide a broad overview of the variables rather than fully processing all of the data.

Hammersley (2007) pointed out that criticism of unsolved issues in "positivist" research is also apparent, precisely the challenges in quantifying the benefit of schooling. It is possible not to appropriately estimate the relevance of things like an English language test for someone attempting to live in another nation, for instance, if they are an immigrant.

9. The Most Common Approach Utilized in Language Evaluations and Tests

After examining the benefits and drawbacks of employing qualitative and quantitative research approaches, this part focuses on the tried-and-true way of researching language examinations. The quantitative research approach is utilized much more often than the qualitative research technique in language testing and evaluation. According to a study by Jang, *et al.*, (2014), the use of language Evaluation and Inspection has historically been established on a quantitative paradigm. This is because language testers continue to employ statistical approaches to validate tests. Despite this, different methods, in addition to the quantitative research technique, are used in the field of language testing in order to study the test's validity in validation studies. As a direct consequence of this, the validation of tests often uses qualitative research methods such as introspection and retrospection (Alderson & Banerjee, 2002). As a direct result, the area of study that focuses on evaluating languages has gone through a substantial shift. According to Lazaraton (2002), the beginning of the use of qualitative research approaches for explicitly building, characterizing, and validating language examinations is the most significant advancement that has taken place in the field of language testing. Since language testing academics throughout the world mostly use quantitative research methodologies in their work, the quantitative research strategy is still successful in language examination and evaluation research. For instance, the "Language Testing" magazine's January 2016 issue. Publication reveals how almost all investigations used quantitative research methodologies. In addition, statistical research methods were employed in practically all of the studies published in the four issues of the journal Language Assessment Quarterly in 2015. In 2015, the magazine produced a total of four editions. In language testing and evaluation, it is abundantly clear

that the quantitative research approach is still used extensively.

10. Ethical Factors to Take into Account

Hammersley and Traianou (2012) established five principles primarily recognized by the community. These five principles are: reducing damage, maintaining autonomy, safeguarding privacy, encouraging reciprocity, and treating people fairly. While administering language tests, language evaluators are responsible for respecting the civilization and self-respect of each test taker (ILTA, 2016). The applicability of these concepts is up for debate, even though not all of them are. For instance, putting participants in a position where they may decide whether or not to take part in the study puts them in danger of experiencing adverse effects. The fact that ethical norms vary depending on the situation gives rise to several potential issues in terms of ethics. Due to cultural variations, individuals from one society may have the same opinion about a concept, while others from a different culture may hold opposing views. A system of ethical examination is in place for all types of research that include human subjects in some industrialized nations, including the United Kingdom, the United States of America, Canada, and Australia. However, many developing nations still need a methodical process for social scientists to use to assess the ethical norms of such countries (Farrimond, 2013). Another case of an ethical conundrum presented by Loizos (2000) is how unwanted monitoring is handled as an infringement on one's right to privacy. When conducting medical research, he said, the ethics committee has a responsibility to strike a balance between the concerns of individual patients and those of the general public. It is generally agreed upon that ethical considerations play a more significant role in qualitative research than quantitative research.

Because in-depth interviews and observation may be pretty invasive when individuals are questioned about compassionate things, qualitative research may be challenging, particularly in the human services business (Darlington & Scott, 2003). This is especially true in an industry where people are asked questions about personal issues. Furthermore, Blaxter, *et al.*, (2001) pointed out that due to the strong links between the research and the research, ethical difficulties are often predicted to arise with research designs that employ qualitative data-collecting procedures. This is because of the strong linkages between the research and the researched.

According to Crano, Brewer, and Lac (2015), the ultimate purpose of the ethical conundrum is to highlight that no hard-and-fast criteria can be used to determine whether a particular research method or approach is morally acceptable. In addition, not everything that is considered to be "ethical" in one culture or country will necessarily be seen as so in another culture or country. The notions of morality and ethics are understood differently in different cultures. According to Karavas (2013), morality may never be perfect or comprehensive, and ethics cannot be applied to every conceivable circumstance.

11. CONCLUSION

The purpose of the study was to provide a critical analysis of the advantages and disadvantages of the two primary research paradigms in education, namely qualitative and quantitative. Linguistic analysis and evaluation provided the foundation for the study's structure. Because it considers ethical concerns and provides a critical analysis of the benefits and drawbacks of various research approaches and procedures, one may argue that the study has successfully accomplished the goals it set out to accomplish. Utilizing qualitative research methods in language evaluation and testing studies offers many positive implications. These advantages of grading, in addition to the complex components included in language examinations. Additional advantages include having a deeper understanding of the preparation, administration, and analysis of language examinations. However, there are certain limitations, including the following: Government officials may not request qualitative research because of the possibility that conclusions obtained from studies with small sample sizes would be erroneous and cannot be generalized. However, the study's findings revealed a great deal with characteristics that validate the testing research. On the other hand, some drawbacks include the following: The consequences and results of the system being evaluated should be taken into account by quantitative research; for example, a difficult scheme is not troubled with the sum of all the factors. In addition, the research led to the conclusion that academics all around the globe use quantitative approaches while carrying out the vast majority of language testing and assessment studies. The research findings suggest that concerns about ethics associated with the project are now of the utmost relevance. In addition, the many different research approaches each come with their own advantages and disadvantages. On the other hand, it demonstrates that there is no one set of rules that can be taken as gospel when seeking to assess whether or not the conduct of a particular research project or its technique is ethical. It may seem that the inquiry has focused on the topic's most challenging aspects; moreover, research techniques and procedures are extensive areas of study; as a result, it was only feasible to handle some of the challenges in such a limited investigation. Although it may appear that the investigation has centered on the topic's most challenging aspects, this is not the case. As a result, the urgent want for more research will become apparent instantly.

REFERENCES
- Bachman, L. F. (1998). Interfaces between second language acquisition and language testing research. Cambridge: Cambridge University Press.

- Berg, B. L. (2009). Qualitative research methods for the social sciences (7th ed.). Pearson International.
- Blaikie, N. (2007). Approaches to social enquiry (2nd ed.). Cambridge: Polity Press.
- Blaxter, L., Hughes, C., & Tight, M. (2001). How to research (2nd ed.). Buckingham: Open University Press.
- Bryman, A. (2012). Social research methods. New York: Oxford University Press.
- Carr, L. T. (1994). The strengths and weaknesses of quantitative and qualitative research: What method for nursing? *Journal of Advanced Nursing, 20*(4), 716-721. http://dx.doi.org/10.1046/j.1365-2648.1994.20040716.x
- Chalhoub-Deville, M., & Deville, C. (2008). Utilizing psychometric methods in assessment. In E. Shohamy, & N. H. Hornberger (Eds.), Encyclopedia of language and education (2nd ed., Vol. 7, pp. 211-224). New York, NY: Springer Science + Business Media LLC.
- Corbetta, P. (2003). Social research: Theory, methods and techniques. London: SAGE.
- Crano, W. D., Brewer, M. B., & Lac, A. (2015). Principles and methods of social research (3rd ed.). London and New York, NY: Routledge.
- Cumming, A. (2001). ESL/EFL instructors' practices for writing assessment: Specific purposes or general purposes? *Language Testing, 18*(2), 207-224. http://dx.doi.org/10.1177/026553220101800206
- Darlington, Y., & Scott, D. (2003). Qualitative research in practice: Stories from the field. *Social Work Education: The International Journal, 22*(1), 115-118.
- Denzin, N. K. (1989). Interpretive interactionism. Newbury Park, CA: Sage.
- Denzin, N. K., & Lincoln, Y. S. (1994). Handbook of qualitative research. London: Sage Publications.
- Denzin, N. K., & Lincoln, Y. S. (2011). The SAGE handbook of qualitative research. London: SAGE.
- Douglas, D. (2014). Understanding language testing. London: Routledge.
- Farrimond, H. (2013). Doing ethical research. Basingstoke UK, New York, US: Palgrave Macmillan.
- Fidalgo, A. M., Alavi, S. M., & Amirian, S. M. R. (2014). Strategies for testing statistical and practical significance in detecting DIF with logistic regression models. *Language Testing, 31*(4), 433-451. http://dx.doi.org/10.1177/0265532214526748
- Flanagan, T. (2013). The scientific method and why it matters. *C2C Journal*, *7*(1), 4-6.
- Flick, U. (2014). An introduction to qualitative research (5th ed.). London: Sage Publications Ltd.
- Fulcher, G. (2010). Practical language testing. London: Hodder Education.
- Fulcher, G., & Davidson, F. (2007). Language testing and assessment: An advanced resource book. London: Routledge.
- Hammersley, M. (2007). Educational research and teaching: A response to David Hargreaves' TTA Lecture. In M. Hammersley (Eds.), Educational Research and Evidence-based Practice (pp. 18-42). London: SAGE Publications.
- Harry, B., & Lipsky, M. (2014). Qualitative Research on Special Education Teacher Preparation. In M. McCray, T. Brownell, & B. Lignugaris/Kraft (Eds.), Handbook of research on special education teacher preparation (pp. 445-460).
- ILTA. (2016). Codes of Ethics. Retrieved from http://www.iltaonline.com/images/pdfs/ilta_code.pdf
- Jang, E. E., Wagner, M., & Park, G. (2014). Mixed Methods Research in Language Testing and Assessment. *Annual Review of Applied Linguistics*, 34, 123-153. http://dx.doi.org/10.1017/S0267190514000063
- Karavas, K. (2013). Fairness and Ethical Language Testing: The Case of the KPG. Directions in Language Teaching and Testing, 1.
- Katzenberger, I., & Meilijson, S. (2014). Hebrew language assessment measure for preschool children: A comparison between typically developing children and children with specific language impairment. *Language Testing, 31*(1), 19-38. http://dx.doi.org/10.1177/0265532213491961
- Kauber, P. (1986). What's Wrong with a Science of MIS (pp. 572-574)? Proceedings of the 1986 Decision Science Institute, Honolulu, HA.
- Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly, 23*(1), 67-93. http://dx.doi.org/10.2307/249410
- Kroeze, J. H. (2012). Postmodernism, interpretivism, and formal ontologies. In M. Mora. (Eds.), Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems. USA: Information Science Reference.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197. http://dx.doi.org/10.1177/0265532214554321
- Lazaraton, A., & Taylor, L. (2007). Qualitative research methods in language test development and validation. In Fox. (Eds.), Language Testing Reconsidered (pp. 113-130). Ottawa: University of Ottawa Press.
- Loizos, P. (2000). Video, film and photographs as research documents. In M. W. Bauer, & G. Gaskell (Eds.), Qualitative Researching with Text. Image and Sound. London: Sage Publications
- Marcon, T., & Gopal, A. (2005). Uncertain knowledge, uncertain time. Toronto: ASAC.
- Martin, W., & Bridgmon, K. (2012). Quantitative and statistical research methods: from hypothesis to results. New Jersey, USA: Jossey-Bass.
- Maxwell, J. A. (2012). *Qualitative research design: An interactive approach*. Sage publications.
- McNamara, T. (2000). Language Testing. Oxford: Oxford University Press.
- McNamara, T., & Roever, C. (2006). Language testing: The social dimension (Vol. 1). John Wiley & Sons.

- Mohan, B. A. (2012). Qualitative research methods in second language assessment. In L. V. Hedges, R. Coe, & M. Waring (Eds.), Research methods and methodologies in education (pp. 752-767). Thousand Oaks, CA: Sage.
- Payne, G., & Payne, J. (2004). Key concepts in social research. London: Sage.
- Powers, D. E., & Powers, A. (2015). The incremental contribution of TOEIC® Listening, Reading, Speaking, and Writing tests to predicting performance on real-life English language tasks. *Language Testing, 32*(2), 151-167. http://dx.doi.org/10.1177/0265532214551855
- Rasinger, S. M. (2013). Quantitative research in linguistics: An introduction. A & C Black
- Ravitch, D. (2010). The death and life of the great American school system: How testing and choice are undermining education. New York, NY: Basic Books.
- Richards, T. J., & Richards, L. (1994). Using computers in qualitative research. In N. Denzin, & Y. Lincoln (Eds.), Handbook of Qualitative Research (pp. 445-462). London: Sage Publications
- Richardson, A. J. (2012). Paradigms, theory and management accounting practice: A comment on Parker (forthcoming) "Qualitative management accounting research: Assessing deliverables and relevance". *Critical Perspectives on Accounting, 23*(1), 83-88. http://dx.doi.org/10.1016/j.cpa.2011.05.003
- Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (2013). Qualitative research practice: A guide for social science students and researchers. London: Sage.
- Rumsey, M., Thiessen, J., Buchan, J., & Daly, J. (2016). The consequences of English language testing for international health professionals and students: An Australian case study. *International Journal of Nursing Studies*, 54, 95-103. http://dx.doi.org/10.1016/j.ijnurstu.2015.06.001
- Sallee, M. W., & Flood, J. T. (2012). Using qualitative research to bridge research, policy, and practice. *Theory Into Practice, 51*(2), 137-144. http://dx.doi.org/10.1080/00405841.2012.662873
- Schofield, J. W. (2007). Increasing the generalizability of qualitative research. In M. Hammersley (Ed.), Educational Research and Evidence-based Practice (pp. 181-203). London: SAGE Publications
- Silverman, D. (2010). Qualitative research. London: Sage.
- Strauss, A. L., & Corbin, J. M. (1990). Basics of qualitative research (Vol. 15). Newbury Park, CA: Sage.
- Van Maanen, J. (1979). Reclaiming qualitative methods for organizational research: A preface. *Administrative Science Quarterly, 24*(4), 520-526. http://dx.doi.org/10.2307/2392358
- Weber, R. (2004). Editor's comments: The rhetoric of positivism versus interpretivism: A personal view. *MIS Quarterly, 28*(1), iii-xii. Retrieved from http://www.jstor.org/stable/pdf/25148621.pdf?acceptTC=true
- Weir, C. J. (2005). Language Testing and Validation. New York: Palgrave Macmillan.
- Wilson, A. (2014). Being a practitioner: An application of Heidegger's phenomenology. *Nurse Researcher, 21*(6), 28-33. http://dx.doi.org/10.7748/nr.21.6.28.e1251