University of Plymouth

PEARL **https://pearl.plymouth.ac.uk**

Faculty of Science and Engineering School of Engineering, Computing and Mathematics

The Advantages and Disadvantages of Using Qualitative and Quantitative Approaches and Methods in Language "Testing and Assessment" Research: A Literature Review

Rahman, MS

http://hdl.handle.net/10026.1/16598

10.5539/jel.v6n1p102
Journal of Education and Learning
Canadian Center of Science and Education

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Journal of Education and Learning; Vol. 6, No. 1; 2017
ISSN 1927-5250 E-ISSN 1927-5269
Published by Canadian Center of Science and Education

The Advantages and Disadvantages of Using Qualitative and Quantitative Approaches and Methods in Language "Testing and Assessment" Research: A Literature Review

Md Shidur Rahman[1]

[1] School of Social Sciences, Education and Social Work, Queen's University Belfast, Belfast, UK

Correspondence: Md Shidur Rahman, 69-71 University Street, Queen's University Belfast, BT7 1HL, Belfast, UK. Tel: 44-078-9498-7320. E-mail: shidureh@yahoo.co.uk

Received: July 23, 2016 Accepted: October 14, 2016 Online Published: November 10, 2016

doi:10.5539/jel.v6n1p102 URL: http://dx.doi.org/10.5539/jel.v6n1p102

Abstract

The researchers of various disciplines often use qualitative and quantitative research methods and approaches for their studies. Some of these researchers like to be known as qualitative researchers; others like to be regarded as quantitative researchers. The researchers, thus, are sharply polarised; and they involve in a competition of pointing out the benefits of their own preferred methods and approaches. But, both the methods and approaches (qualitative and quantitative) have pros and cons. This study, therefore, aims to discuss the advantages and disadvantages of using qualitative and quantitative research approaches and methods in language testing and assessment research. There is a focus on ethical considerations too. The study found some strengths of using qualitative methods for language "assessment and testing" research—such as, eliciting deeper insights into designing, administering, and interpreting assessment and testing; and exploring test-takers' behaviour, perceptions, feelings, and understanding. Some weaknesses are, for instance, smaller sample size and time consuming. Quantitative research methods, on the other hand, involve a larger sample, and do not require relatively a longer time for data collection. Some limitations are that quantitative research methods take snapshots of a phenomenon: not in-depth, and overlook test-takers' and testers' experiences as well as what they mean by something. Among these two research paradigms, the quantitative one is dominant in the context of language testing and assessment research.

Keywords: qualitative and quantitative research, advantages, disadvantages, testing and assessment

1. Introduction

Qualitative and quantitative research approaches and methods are usually found to be utilised rather frequently in different disciplines of education such as sociology, psychology, history, and so on. Concerning the research approaches, there persist so-called paradigm wars in which researchers belong to the two distinct camps—interpretivism and positivism. The positivistic researchers' belief is that the social world consists of concrete and unchangeable reality which can be quantified objectively. Whereas, the interpretive researchers oppose the positivistic belief of reality; and argue that, instead, the reality is socially constructed by the humans which can be changed and understood subjectively (Corbetta, 2003; Marcon & Gopal, 2005; Kroeze, 2012). As focused on the paradigm wars, there are some arguments about the superiority of research methods: Whether qualitative evidence is superior to quantitative evidence, or vice-versa. In this regard, there is a need to highlight the history of the emergence of paradigms in order to understand the dominant method in the arena of research. For the first half of the twentieth century, the positivist model was dominant in social and educational research by giving importance of using standardised tests and systematic observation, experiment, survey data, and statistical analysis. In other words, the quantitative research method was quite powerful. However, after the 1960s, because of the influence of Kuhn's work, the new approaches—such as, symbolic interactionist ethnography, critical research, feminist, phenomenography, discourse analysis and other forms—came up into practice (Hammersley & Traianou, 2012). That is to say, there was a shift of research methods from quantitative to qualitative, and the superiority of quantitative research was not as powerful as before. The growth of qualitative research was also from the unhappiness with the process of generating knowledge within the positivistic research (Sandberg, 2005). Alongside the research methods and approaches, ethical considerations were also introduced in the research world though it was subsequent to research methods and approaches. Up to

this point, it is obvious that a various types of research techniques and methods were employed in different areas of qualitative and quantitative research.

Likewise, the language testing, as an area of, research has encountered a vast array of methods and approaches (Bachman, 2000), for example, VELC Test® score interpretations technique used by Kumazawa, Shizuka, Mochizuki, and Mizumoto (2016); Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) employed by Sims and Kunnan (2016); action research in Cambridge English Language Test (Borg, 2015; Watkins, 2015; Depieri, 2015).

The study aims at critically discussing the advantages and disadvantages of using quantitative and qualitative approaches and methods for language testing and assessment research. The study begins with an introduction to the background of research methods and approaches (quantitative and qualitative). The introduction is followed by a brief description of language testing and assessment. Then, it presents the pros and cons of using qualitative and quantitative approaches and methods; and evaluates the dominant research methods in language testing and assessment research. Finally, ethical considerations are also pointed out.

2. Language Testing and Assessment

In a general sense, a test is something that demonstrates one's competence-incompetence, ability-inability; and that shows someone's position in the scale consisting of variables such as fail, pass, average, satisfactory, good, and excellent. An academic test also helps taking an important decision of whether or not a student will be allowed to move up to the next step. It can check the progress of a student and suggest whether a student needs more help or not, and allow us to compare the performance between students. The test, furthermore, acts as an important tool of public policy—such as the national examinations are held in the same standard across the country to ensure that only the top performers can get admission to the next level of education (Douglas, 2014). The university admission test (a high stake test) is a tool of this kind. In language testing, the testers are concerned with the extent to which a test can produce scores that reflect a candidate's ability accurately in a specific area, for example, reading, writing a critical essay, vocabulary knowledge, or spoken interaction with peers (Weir, 2005).

Like all other educational assessments, language testing is a complex social phenomenon (Fulcher, 2010). But it is a significant aspect in education which affects people's lives in the society such as—promotion, employment, citizenship, immigration or asylum depends upon passing a language test. Another consequential factor of language testing in education is that it dictates what is to be taught (McNamara & Roever, 2006). So, the discussion so far indicates that the language tests play an important role in many people's lives (McNamara, 2000). However, many have highlighted the validity and reliability of language testing-for example, Fulcher and Davidson (2007) claimed that every book and article relating to language testing addresses the test validity to some extent which is the core concept of testing and assessment. So, a multiple concepts are seen to be involved in language testing and assessment, and it is a good area of research in education.

3. Qualitative Research Approaches and Methods

Defining qualitative research is significant as it is the central focus of this section—but there is a challenge to define this term clearly (Ritchie, Lewis, Nicholls, & Ormston, 2013), since it does not have its theory or paradigm nor an obvious set of methods or practices that are merely of its own (Denzin & Lincoln, 2011). This term also involves a vast array of methods and approaches within the different subjects of research. Hence, the writers have provided the definition of qualitative research distinctively. Strauss and Corbin (1990, p. 11), for example, stated that, "By the term 'qualitative research', we mean any type of research that produces findings not arrived at by statistical procedures or other means of quantification. It can refer to research about persons' lives, lived experiences, behaviours, emotions, and feelings as well as about organisational functioning, social movements, cultural phenomena, and interactions between nations." This means that qualitative research is not statistical and it incorporates multiple realities. Then, Flick (2014, p. 542) claimed that, "Qualitative research interested in analysing subjective meaning or the social production of issues, events, or practices by collecting non-standardised data and analysing texts and images rather than number and statistics." This definition stressed on how people make sense of something in the world. So, the qualitative research is basically associated with multiple aspects. Moreover, Denzin and Lincoln (1994, p. 2) claimed that, "Qualitative research is multi-method in focus, involving an interpretive, naturalistic approach to its subject matter." It is, moreover, apparent that the qualitative research is concerned with multiple perspectives when Van Maanen (1979, p. 520) defines it as, "an umbrella term covering an array of interpretive techniques which seek to describe, decode, translate, and otherwise come to terms with the meaning, not the frequency, of certain more or less naturally occurring

phenomena in the social world." Therefore, qualitative research appears to be an overarching concept under which a variety of issues may be placed, and it has positive and negative perspectives.

3.1 Advantages

There are some benefits of using qualitative research approaches and methods. Firstly, qualitative research approach produces the thick (detailed) description of participants' feelings, opinions, and experiences; and interprets the meanings of their actions (Denzin, 1989). In terms of language testing, for example, Bachman (1998) showed in his study that qualitative research results provide the relationship of information processing with performance specifically and deeply. Chalhoub-Deville and Deville (2008), too, argued that qualitative approaches are employed to achieve deeper insights into issues related to designing, administering, and interpreting language assessment.

Secondly, there are some who argue that qualitative research approach (interpretivism) holistically understands the human experience in specific settings. Denzin and Lincoln (2002), for example, mentioned that qualitative research is an interdisciplinary field which encompasses a wider range of epistemological viewpoints, research methods, and interpretive techniques of understanding human experiences. From the perspective of epistemological position, any language assessment cannot be set apart from context, culture and values of where it was used (McNamara, 2001); and the language assessment researchers began to employ qualitative research methodology for focusing the issues that need an analysis of content-related variables such as the influences of test-takers' characteristics on test performance, the strategies used to respond to assessment tasks, and so on (Tsushima, 2015).

Thirdly, interpretivism research approach is regarded as an ideographic research, the study of individual cases or events (Kelin & Myers, 1999); and it has abilities to understand different people's voices, meanings and events. So the source of knowledge in this approach is the meaning of different events (Richardson, 2012). In language testing, the qualitative research techniques analyse the candidate behaviour, interviewer behaviour, interlocutor behaviour, and cross-cultural influences on behaviour during the speaking tests (Lazaraton & Taylor, 2007).

Fourthly, the qualitative research admits the researchers to discover the participants' inner experience, and to figure out how meanings are shaped through and in culture (Corbin & Strauss, 2008). Such as, in terms of assessing written assignments, "satisfactory", "good", or "60 marks out of 100" are used by the assessors, an investigation might be made in order to understand the meaning of "satisfactory" or "good" or to elicit the features of content or text displayed in the student scripts (Leung, 2012). So, the studies using qualitative approach can help us understand the markers' working assumption about what is to be assessed, and the meaning of the score or grade.

Fifthly, qualitative research methods such as participant-observation, unstructured interviews, direct observation, describing records are most commonly used for collecting data (Cohen, Manion, & Morrison, 2011). During the data collection, the researchers interact with the participants directly such as it happens while data collection through interviews. Consequently, data collection is subjective and detailed. The study of Rumsey, Thiessen, Buchan and Daly (2016) deployed the face-to-face and telephone interview techniques of data collection, and elicited the feelings, perceptions, and views about IELTS tests.

Lastly, qualitative research design (interactive approach) has a flexible structure as the design can be constructed and reconstructed to a greater extent (Maxwell, 2012). Thus, the thorough and appropriate analyses of an issue can be produced by utilising qualitative research methods, and therefore the participants have sufficient freedom to determine what is consistent for them (Flick, 2011). As a result, the complex issues can be understood easily. For example, researchers acknowledged that, because of the nature of classroom dynamics, learners' behaviour may be affected by the numerous factors outside of research focus. In this respect, the qualitative research approach is required to capture these dynamics. As the nature of language assessment practices is also complex, there is a suggestion by Mohan (2012) to employ the qualitative research methodology. Thus, the qualitative research can contribute to the understanding of the complex features of language assessment. The next section will look at the disadvantages of qualitative research approaches and methods.

3.2 Disadvantages

Beyond the above advantages, some limitations are obvious. First, Silverman (2010) argues that qualitative research approaches sometimes leave out contextual sensitivities, and focus more on meanings and experiences. Phenomenological approach, for instance, attempts to uncover, interpret and understand the participants' experience (Wilson, 2014; Tuohy et al., 2013). Similarly, Cumming (2001) focused on the participants' experience rather than any other imperative issues in the context. He engaged six countries (Australia, Canada, New Zealand, Hong Kong, Japan and Thailand) and experienced writing instructors of these countries in his

research. He stated that, "This sampling was selective and purposive, focused on instructors with high levels of expertise in each setting, rather than aiming to be representative of educators in the particular countries or institutions." It seems that the contextual influence on the assessment is not taken into account at all in this research because it has looked at the writing instructors' expertise, and bypassed what impacts of these countries' cultures for example would be on writing assessment.

Second, policy-makers may give low credibility to results from qualitative approach. Sallee and Flood (2012) found that stakeholders frequently use quantitative research when research is called upon. In terms of educational practice in United States, national and state policymakers sought to quantify teacher's and student's performance (Ravitch, 2010), and in many social sciences, quantitative orientations are frequently given more regard (Berg, 2009). In addition, purely qualitative research may neglect the social and cultural constructions of the variables studied (Richards & Richards, 1994).

Third, in terms of research method, smaller sample size raises the issue of generalizability to the whole population of the research (Harry & Lipsky, 2014; Thompson, 2011). Having studied the language testing assessment in Hong Kong qualitatively, Lam (2015) admitted that due to the small sample size the study results do not wish to claim wider generalization to other contexts.

Then, data interpretation and analysis may be more difficult/complex (Richards & Richards, 1994). Berg and Lune (2012, p. 4) also commented that, "Qualitative research is a long hard road, with elusive data on one side and stringent requirements for analysis on the other." Along with the data interpretation and analysis issue, Darlington and Scott (2003) claimed that developing the undeveloped question into a researchable form is harder, and the refining question in qualitative research may be continuous throughout a whole study. In the research paper of Rumsey, Thiessen, Buchan and Daly (2016), it has appeared that the primary research data collection was not focused on the research topic, language testing or language requirement, rather focused on the thematic analysis. But in the secondary data analysis, the collected data were again used for investigating a new question that was not explored in the primary research. It seems that the data analysis and developing the research question using the same data, which is previously collected, is likely to be a harder and continuous process of conducting a qualitative research.

Finally, the analyses of the cases take a considerable amount of time, and one can generalise the results to the larger population in only a very limited way (Flick, 2011). For example, if a legislator needs to vote an issue, she/he cannot wait for three months for a qualitative study to be administered (Sallee & Flood, 2012). Similarly, in language testing and assessment research, in order to devise a new policy within a short period of time, the policy makers may demand quantitative research instead of qualitative research.

However, despite these shortcomings, the qualitative research becomes prominent in language testing research—as it is a regular accompaniment with quantitative data analysis to which reports of standard setting are generally confined (Manias & McNamara, 2015). Moreover, the generalisability seems not to be a problem as Darlington and Scott (2003, p. 18) pointed out that, "If one considers the unit of attention as the phenomenon under investigation, rather than the number of individuals, then the sample is often much larger than first appears." Thus, in the studies of language testing, the number of interactions or contacts investigated would have been infinitely larger than the individuals or families involved. Moreover, Labaree (2004) suggested that no educational research (either quantitative or qualitative) ought to be regarded as generalizable, because too many contextual variables can shape the findings. For example, if a student is unable to demonstrate enough proficiency in reading or math, quantitative measures—test scores may indicate that the teacher is to be condemned. But the test scores fail to consider the classroom environment, students' home life, and other crucial factors. Donmoyer (2012) argued that the qualitative researchers can tell the policy makers what works as the qualitative research provide the thick description. This study will next consider the advantages and disadvantages of quantitative research.

4. Quantitative Research Approaches and Methods

Bryman (2012, p. 35) defined quantitative research as, "A research strategy that emphasises quantification in the collection and analysis of data…" It means quantitative research denotes amounting something. This research method attempts to investigate the answers to the questions starting with how many, how much, to what extent (Rasinger, 2013). In other words, the method lays heavy stress on measuring something or variables existed in the social world. Payne and Payne (2004, p. 180) stated that, "Quantitative methods (normally using deductive logic) seek regularities in human lives, by separating the social world into empirical components called variables which can be represented numerically as frequencies or rate, whose associations with each other can be explored by statistical techniques, and accessed through researcher-introduced stimuli and systematic measurement." The

quantitative research focuses on those aspects of social behaviour which can be quantified and patterned rather than just finding out them and interpreting their meanings the people bring to their own action.

Another aspect visa-à-vis the quantitative research is worth mentioning which is positivism that underpins quantitative research. Weber (2004) mentioned his colleague, Jorgen Sandberg, who has characterised the positivism that in positivism, person and reality are separate; objective reality exists beyond the human mind; research methods are statistics and content analysis; validity-data truly measures reality; reliability-research results can be reproduced; research object has inherent qualities that exist independently of the researchers. Additionally, Bryman (2012) identified positivism as a nomothetic research (which yields law-like or general actions); in positivism, knowledge is obtained from empirical testing (Richardson, 2012). However, interpretivism in qualitative research seems to be opposite to the positivism when the researchers argue that interpretivism is an ideographic research (the study of individual cases or events) (Kelin & Myers, 1999), and knowledge is derived from the meaning of events (Richardson, 2012). From this discussion it is evident again that a paradigm war persists in the research world in which researchers are staying in two different camps, though Weber (2004) contends that the concept of positivism versus interpretivism is vacuous. Furthermore, Bryman (1988) claimed that, like qualitative methods, "quantitative methods" is also an umbrella term covering various types of research. Pros and cons of quantitative research are as follows:

4.1 Advantages

The quantitative findings are likely to be generalised to a whole population or a sub-population because it involves the larger sample which is randomly selected (Carr, 1994). Besides sampling, data analysis is less time consuming as it uses the statistical software such as SPSS (Connolly, 2007). Powers D. and Powers A. (2015), in their research on TOEIC tests, contented that the study sample reflects the larger proportion (2300) of TOEIC test-taking population which helps making the study truth-worthy.

Then, quantitative research is to be based on positivist paradigm of measuring variables (Kauber, 1986). A language proficiency assessment research, for example, conducted by Carroll and Bailey (2016) shows that there are different variables such as EFL student and non-EFL students, and tests in four sub-domains: speaking, writing, reading, and listening. It is also noticed in another study of second language fluency by Préfontaine, Kormos and Johnson (2016) that various variables were used such as class variables: beginning, intermediate and advanced; native speaker variables: British, American, and Canadian.

4.2 Disadvantages

Given the strengths above, the quantitative research has limitations too. The positivism research paradigm leaves out the common meanings of social phenomenon (Denzin & Lincoln, 1998). It also fails to ascertain deeper underlying meanings and explanations. The study of Bouwer, Béguin, Sanders and van den Bergh (2015) measured the effect of genre on writing score involving various types of variance such as person, genre, person by genre, task within genre, person by task within genre, raters who rated tasks within different genres, and random error. But the study has not explained the reasons of the effect and their meanings in that particular context.

Another limitation of quantitative research is that the positivism cannot account for how the social reality is shaped and maintained, or how people interpret their actions and others (Blaikie, 2007). In quantitative language testing research, it is found that the studies just investigate and estimate the language skills, proficiency, scoring, and so on. A study undertaken by Katzenberger and Meilijson (2014) on the assessment of Hebrew language for preschool children in which it is to find out the language-impaired Hebrew-speaking pre-school children and to identify whether the test can distinguish the developing and language-impaired children. However, the study did not explore why some children develop their language learning and why some are impaired. That is to say, the study left over the reality or the children's understanding of learning capacity of Hebrew language.

A further weakness of quantitative research approach is that it has tendencies of taking a snapshot of a phenomenon: It measures variables at a specific moment in time, and disregards whether the photograph happened to catch one looking one's best or looking unusually disarranged (Schofield, 2007), for example, the study of Fidalgo, Alavi and Amirian (2014) engaged a larger number of participants (400) to show the testing statistical significance. Therefore, it has not been possible to take information deeply; rather, it has given the overall picture of the variables.

Then, Hammersley (2007) pointed out that the criticism of unresolved problems in "positivist" research is also apparent, such as-the difficulties of measuring of educational significance. The importance of English language test, for instance, for an immigrant in order to be settled in the UK may not be possible to measure objectively.

Lastly, the quantitative research paradigm overlooks the respondents' experiences and perspectives in highly controlled settings (Ary, Jacobs, Sorensen, & Walker, 2013) because there lacks a direct connection between researchers and the participants when collecting data. As a result, the data obtaining method becomes objective. The study of Gu (2015)—such as—engaged 498 participants in the study of language ability test by distributing test-lets. This shows that the study has been unable to elicit the participants' perceptions as the researchers had no connection with them; just handing out the test-lets.

Although the above drawbacks exist with the quantitative research, a considerable amount of research is being conducted in the arena of language testing and assessment by employing quantitative research methods and approaches. Language assessment has a tradition of employing psychometric procedures and a strong orientation of quantitative research methods focusing on test score validity (Purpura, 2011). Besides, the language testing is taken into account as measurement. It is noted that the psychometric-structural phase of language testing denoted by Spolsky emphasised on test reliability and established the concept of language testing as measurement (Morrow, 2012). So, from this concept, the language testing seems to be quantitative in nature rather than qualitative.

5. Dominant Method in Language Testing and Assessment

Having paid attention to the advantages and disadvantages of using qualitative and quantitative research method, this section stresses on the mainstream research method on language testing. In the field of language testing and assessment, the quantitative research method seems to be used rather than the qualitative research method. Jang, Wagner, and Park (2014) commented that, "The field of Language Testing and Assessment (LTA) has been traditionally dominated by a quantitative paradigm." Because language testers carry on using statistical means of test validation. However, quantitative research technique is not the only method utilised in validation studies, rather multiple other methods are used in language testing for exploring test validity. Qualitative research techniques, therefore, such as introspection and retrospection are extensively used in test validation (Alderson & Banerjee, 2002). Therefore, a significant shift is noticed in the field of language testing research. Lazaraton (2002) stated that the most significant development in language testing is the start of using qualitative research methodologies for designing, describing, and validating language tests in particular.

I would argue that the quantitative research method is still powerful in regard to language testing and assessment research because the researchers of language testing around the world mostly employ the quantitative research techniques for their research. Almost all the studies—for example, in the January 2016 issue of journal called "Language Testing"—involved quantitative research strategies. Furthermore, statistical techniques of research were employed in almost all the published articles of 4 issues in the Language Assessment Quarterly journal, 2015. So, it is evident that the quantitative research is yet a dominant method of research in the field of language assessment and testing research.

6. Ethical Considerations

Ethical issues are paramount important matters not only in the primary research in particular, but also even in terms of using secondary data sets because there are ethical issues relating to fair and unbiased selection of sources and analysis (Farrimond, 2013). So, Hack (1997, p. 37 in Blaxter, Hughes, & Tight, 2001) emphasised that, "It is worth standing back for a moment and considering what effect your actions might have on others as the result can be quite damaging to yourself." The myriad issues are involved in ethical considerations. Such as in "informed consent" of participants, there mention about the research in which participants are involved, about who is conducting the research, and about the results. Maintaining the confidentiality of the results and anonymity of the participants are also crucial along with these issues (Black, 1999). But, there is likely no certainty about the confidentiality of material sent by email, as they can be easily forwarded and copied. Hackers can gain access to the customers' databases of public or private organisations (Blaxter, Hughes, & Tight, 2001). The International Language Testing Association: ILTA (2016, p. 3) also stated that, "Confidentiality cannot be absolute, especially where the records concern students who may be competing for admissions and appointments." In other words, it may not be viable to preserve the confidentiality in every aspect. However, maintaining confidentiality is significant. The ILTA (2016, p. 3) simultaneously stressed on that, "A careful balance must be maintained between preserving confidentiality as a fundamental aspect of language tester's professional duty and the wider responsibility the tester has to society." So, it is the responsibility of language testers to hold all about their test takers in confidence.

Hammersley and Traianou (2012) pointed out five commonly recognised principles: minimising harm, respecting autonomy, protecting privacy, offering reciprocity, treating people equitably. In terms of language testing, language testers shall have respect for the humanity and dignity of each of their test takers (ILTA, 2016).

So, like other professions, a common practice in language testing is to develop two codes such as code of ethics and code of practice. The former highlights morals and ideals of the profession; whereas the latter focuses on the principles of the code of ethics (Karavas, 2013).

However, these principles, though not all, seem controversial, for example, respecting autonomy by allowing participants to decide whether they participate or not in the research means subjecting them to risk of harm. Ethical dilemmas are also noticed as ethics principles are context specific. Due to the cultural variations, a principle may be agreed by the participants in one culture; and disagreed in another. In some developed countries like UK, US, Canada, and Australia; there is a system of ethical review of all research with human subject. But many developing countries do not have any formal ethics review procedure for social scientists (Farrimond, 2013). Another example of ethical dilemma presented by Loizos (2000) is that treating the issue of surveillance without consent as an invasion of privacy. He concurrently stated that in medical research the ethics committee has to balance patient interests against issues of wider public interest.

The effects of not following the above principles are negative in general. Nonetheless, overlooking some principles may be useful. Dane (1990) mentioned Humphreys's tearoom trade project in which it would be better not to inform the participants about the nature of the research because it could cause the psychological harm to them. Besides, Jones (2011) argued that even promising to ensure confidentiality is daunting as it is sometimes difficult to achieve anonymity for a research subject.

It is presumed that ethical issues are highly concerned with qualitative research rather than quantitative research. Qualitative research especially in human services experiences challenges associating with ethical considerations, such as in-depth interview and observation can be widely intrusive because highly personal matters of people are interviewed (Darlington & Scott, 2003). In addition, Blaxter, Hughes and Tight (2001) argued that ethical issues are deemed to appear chiefly with research designs which employ qualitative data collection methods, because there is an intimate relationships between research and researched.

Finally, the ethical dilemmas wish to demonstrate the idea that no simple, absolute rules for deciding whether a particular research practice or method is ethical or not (Crano, Brewer, & Lac, 2015). Beyond this, what is regarded "ethical" in one society or culture may not be always regarded so in another. Different cultures have different concepts of morality or ethics. Morality can never be complete or absolute, and ethical principles cannot be commonly used (Karavas, 2013).

7. Conclusion

The study intended to critically look at the strengths and weaknesses of two crucial research paradigms in education—qualitative and quantitative—on the ground of language testing and assessment. It can be argued that the study has met the purpose as it is concerned with the critical demonstration of benefits and shortfalls of using each research methods and approaches, and of ethical considerations. The benefits of using qualitative research in language testing and assessment research are achieving deeper insight into designing, administering and interpreting language assessment; easy to understand the candidate behaviour, interviewer behaviour, interlocutor behaviour, and cross-cultural influences on behaviour during the speaking tests, meaning of the score or grade, and the complex features of language assessment. Limitations, however, are: Small sample size sometimes makes the results unreliable and ungeneralizable; and policy makers may not ask for qualitative research. On the other hand, the study has figured out the advantages of using quantitative research in the context of language testing and assessment research, such as a larger size of sample and variables that make the testing research trust worthy. The disadvantages, nevertheless, are: Quantitative research leaves out the meanings and effects of a particular system—such as, a testing system is not concerned with the detailed picture of variables. Moreover, the study has found that the dominant method in the language testing and assessment research is quantitative due to basically its rate of use by the researchers around the world. Along with the merits and demerits of research methods and approaches, the study has uncovered that the ethical considerations pertinent to a research project undertaken are very crucial issues now-a-days, but it indicates that there are not any absolute rules for deciding whether a particular research practice or method is ethical or not. It seems that the study has focused on the key issues of the topic, however, it has not been able to cover all the issues in a small scale study of this kind as the research approaches and methods are considerably vast areas of study. So, further research needs to be undertaken in the future.

Acknowledgments

I would like to address special thanks to Prof. Dr. Jannette Elwood (School of Social Sciences, Education and Social Work, Queens University Belfast, UK) for her useful guidance and comments on the draft essay.

References

Alderson, J. C., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, *35*(2), 79-113. http://dx.doi.org/10.1017/S0261444802001751

Ary, D., Jacobs, L., Sorensen, C., & Walker, D. (2013). *Introduction to research in education*. Cengage Learning.

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language testing*, *17*(1), 1-42. http://dx.doi.org/10.1177/026553220001700101

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.

Berg, B. L., & Lune, H. (2012). *Qualitative research methods for the social sciences* (8th ed.). Pearson Education.

Blaikie, N. (2007). *Approaches to social enquiry* (2nd ed.). Cambridge: Polity Press.

Black, T. (1999). *Doing quantitative research in the social sciences*. London: Sage Publications ltd.

Blaxter, L., Hughes, C., & Tight, M. (2001). *How to research* (2nd ed.). Buckingham: Open University Press.

Bouwer, R., Béguin, A., Sanders, T., & van den Bergh, H. (2015). Effect of genre on the generalizability of writing scores. *Language Testing*, *32*(1), 83-100. http://dx.doi.org/10.1177/0265532214542994

Borg, S. (2015). Professional development through the Cambridge English/English UK Action Research Scheme. *Cambridge English Language Assessment: Research Notes*, *61*.

Bryman, A. (1988). *Quantity and quality in social research*. London: Sage.

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

Carroll, P. E., & Bailey, A. L. (2016). Do decision rules matter? A descriptive study of English language proficiency assessment classifications for English-language learners and native English speakers in fifth grade. *Language Testing*, *33*(1), 23-52. http://dx.doi.org/10.1177/0265532215576380

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.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). London: Routledge.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research*. London: SAGE Publications Ltd.

Corbetta, P. (2003). *Social research: Theory, methods and techniques*. London: SAGE.

Connolly, P. (2007). *Quantitative data analysis in education: A critical introduction using SPSS*. London & New York, NY: Routledge.

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.

Dane, F. C. (1990). *Research Methods*. California: Brooks/Cole Publishing Company.

Denzin, N. K., & Lincoln, Y. S. (2002). *The qualitative inquiry reader*. London: Sage Publications.

Denzin, N. K., & Lincoln, Y. S. (1994). *Handbook of qualitative research*. London: Sage Publications.

Denzin, N. K. (1989). *Interpretive interactionism*. Newbury Park, CA: Sage.

Denzin, N. K., & Lincoln, Y. S. (1998). *The landscape of qualitative research: Theories and issues*. London: SAGE Publications.

Denzin, N. K., & Lincoln, Y. S. (2011). *The SAGE handbook of qualitative research*. London: SAGE.

Donmoyer, R. (2012). Can qualitative researchers answer policymakers' what-works question? *Qualitative Inquiry*, *18*(8), 662-673. http://dx.doi.org/10.1177/1077800412454531

Douglas, D. (2014). *Understanding language testing*. London: Routledge.

Depieri, T. (2015). The effects of using correction codes and redrafting on intermediate students' writing. *Cambridge English Language Assessment: Research Notes*, *61*, 15-24. Retrieved from http://www.cambridgeenglish.org/images/253785-research-notes-61.pdf

Farrimond, H. (2013). *Doing ethical research*. Basingstoke UK, New York, US: Palgrave Macmillan.

Flick, U. (2014). *An introduction to qualitative research* (5th ed.). London: Sage Publications Ltd.

Flick, U. (2011). I*ntroducing research methodology: A beginner's guide to doing a research project.* London: SAGE Publications Ltd.

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

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. London: Routledge.

Fulcher, G. (2010). *Practical language testing*. London: Hodder Education.

Gu, L. (2015). Language ability of young English language learners: Definition, configuration, and implications. *Language Testing*, *32*(1), 21-38. http://dx.doi.org/10.1177/0265532214542670

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).

Hammersley, M., & Traianou, A. (2012). *Ethics and educational research*. Retrieved from https://www.bera.ac.uk/wp-content/uploads/2014/03/Ethics-and-Educational-Research.pdf?noredirect=1

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.

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

Jones, C. (2011). *Ethical issues in online research, British Educational Research Association on-line resource*. Retrieved from https://www.bera.ac.uk/researchers-resources/publications/ethical-issues-in-online-research

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

Kauber, P. (1986). *What's Wrong With a Science of MIS* (pp. 572-574)? Proceedings of the 1986 Decision Science Institute, Honolulu, HA.

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

Kroeze, J. H. (2012). Postmodernism, interpretivism, and formal ontologies. In M. Mora et al. (Eds.), *Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems*. USA: Information Science Reference.

Kumazawa, T., Shizuka, T., Mochizuki, M., & Mizumoto, A. (2016). Validity argument for the VELC Test® score interpretations and uses. *Language Testing in Asia*, *6*(1), 1.

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. (2002). *A qualitative approach to the validation of oral language tests*. Cambridge: Cambridge University Press.

Lazaraton, A., & Taylor, L. (2007). Qualitative research methods in language test development and validation. In Fox et al. (Eds.), *Language Testing Reconsidered* (pp. 113-130). Ottawa: University of Ottawa Press.

Leung, C. (2012). Qualitative research in language assessment. *The Encyclopedia of Applied Linguistics*. http://dx.doi.org/10.1002/9781405198431.wbeal0979

Labaree, D. F. (2004). *The trouble with ed schools*. New Haven, CT: Yale University 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.

McNamara, T., & Roever, C. (2006). *Language testing: The social dimension* (Vol. 1). John Wiley & Sons.

McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.

McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing*, *18*(4), 333-349. http://dx.doi.org/10.1177/026553220101800402

Manias, E., & McNamara, T. (2015). Standard setting in specific-purpose language testing: What can a qualitative study add? *Language Testing*, *33*(2), 235-249. http://dx.doi.org/10.1177/0265532215608411

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach*. London: Sage.

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.

Morrow, K. (2012). Communicative language testing. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoynoff (Eds.), *The Cambridge Guide to Second Language Assessment* (pp. 140-146). Cambridge: Cambridge University Press.

Marcon, T., & Gopal, A. (2005). *Uncertain knowledge, uncertain time*. Toronto: ASAC.

Payne, G., & Payne, J. (2004). *Key concepts in social research*. London: Sage.

Préfontaine, Y., Kormos, J., & Johnson, D. E. (2016). How do utterance measures predict raters' perceptions of fluency in French as a second language? *Language Testing*, *33*(1), 53-73. http://dx.doi.org/10.1177/0265532215579530

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

Purpura, J. E. (2011). Quantitative research methods in assessment and testing. *Handbook of Research in Second Language Teaching and Learning*, *2*, 731-751. http://dx.doi.org/10.4324/9780203836507.ch44

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.

Rasinger, S. M. (2013). *Quantitative research in linguistics: An introduction*. A & C Black.

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

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.

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

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.

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

Snadberg, J. (2005). How Do We Justify Knowledge Produced Within Interpretive Approaches? *Organisational Research Methods*, *8*(1), 41-68. http://dx.doi.org/10.1177/1094428104272000

Sims, J., & Kunnan, A. (2016). Developing evidence for a validity arguent for an English placement exam from multi-year test performance data. *Language Testing in Asia*, *6*(1). http://dx.doi.org/10.1186/s40468-016-0024-x

Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research* (Vol. 15). Newbury Park, CA: Sage.

Silverman, D. (2010). *Qualitative research*. London: Sage.

Thomson, S. B. (2011). Qualitative research: Validity. *JOAAG*, *6*(1), 77-82.

Tuohy, D., Cooney, A., Dowling, M., Murphy, K., & Sixsmith, J. (2013). An overview of interpretive phenomenology as a research methodology. *Nurse Researcher*, *20*(6), 17-20. http://dx.doi.org/10.7748/nr2013.07.20.6.17.e315

Tsushima, R. (2015). Methodological diversity in language assessment research: The role of mixed methods in classroom-based language assessment studies. *International Journal of Qualitative Methods*, *14*(2), 104-121. http://dx.doi.org/10.7748/nr.21.6.28.e1251

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

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

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.

Watkins, J. (2015). The effects of an extensive reading programme. *Cambridge English Language Assessment: Research Notes*, *61*.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).