APPLICATION OF CONSTANT COMPARISON METHOD IN SOCIAL SCIENCES: A USEFUL TECHNIQUE TO ANALYZE INTERVIEWS

Dr Shumaila Memon
Dr Sumera Umrani
Dr Habibullah Pathan

ABSTRACT

In Social Sciences, interviews are an important technique for collecting data as they deal with meanings, whereas quantitative data deals with numbers (Dornyei, 2007). As, interviews focus on meanings attributed to this by individuals so, it becomes a matter of high importance to address the question of how do we conceptualize those meanings (Rose 1994) and also how reliable these findings are. Although, these are important issues, many studies give relatively less value to them. This paper is an attempt to demonstrate applicability of Maykut and Morehouse's (1994) 'Constant Comparative Method' as one technique to analyze interviews. It also proposes some useful ways to establish reliability in data. This is aimed to help researchers in social sciences to use it to deal with qualitative data.

_____

Keywords: Interviews; Maykut & Morehouse; Constant Comparative Method.

INTRODUCTION

In Social Sciences, interviews are an important technique for collecting data as they deal with meanings, whereas quantitative data deals with numbers (Dornyei 2007). Moreover, interviews report the depth of the conversation, which moves beyond surface talk to a rich discussion of thoughts and feelings (Maykut & Morehouse 1994: 80). As each classroom setting or a social problem is unique, so to understand the peculiarity of a target situation, interviews give time and space to an interviewee to explain in detail the views and experiences.

As, interviews focus on meanings attributed to this by individuals so, it becomes a matter of high importance to address the question of how do we conceptualize those meanings (Rose 1994) and also how reliable these findings are. Although, these are important issues, many studies give relatively less value to them. It does not mean that such studies do not follow any data analysis pattern systematically. Nevertheless, there is need for paying more attention to the reporting of the procedure in terms of analysis and also reliability. That may make the findings in a study to

come out more rigorously and systematically. In return, this will give a clearer path to other researchers to replicate the procedure.

This article aims twofold:

a) Part-I of the article: It proposes one of the methods to analyze the interviews: constant comparative method. It describes all the steps involved in it to analyze the data. Furthermore, each step in the analysis is demonstrated with particular examples. Such illustration will be particularly useful for the researchers who have to deal with the analysis of semi-structured interviews.

b) Part-II of the article: It takes the issue of reliability on board. It describes some of the ways through which reliability may be ensued in a qualitative study. And for the purpose of illustration, some examples are given from a study of the researchers.

PART-I
WHAT IS CONSTANT COMPARATIVE METHOD

The constant comparative method is one way to conduct an inductive analysis of qualitative data (Glaser & Strauss, 1967; Lincolin & Guba, 1985). The constant comparative method (CCM) of analyzing qualitative data combines inductive category coding with a simultaneous comparison of all units of meaning obtained (Glaser & Strauss, 1967, cited in Maykut & Morehouse, 1994:134).

Lincolin and Guba give a detailed description of the procedure involved in this method:

'The essential tasks of categorizing are to bring together into provisional categories those cards [data cards] that apparently relate to the same content; to devise rules that describe category properties and that can, ultimately, be used to justify the inclusion of each card that remains to be assigned to the category as well as to provide a basis for later tests of reliability; and to render the category internally consistent' (Lincolin and Guba, 1985:347).

> **Inductive Category**
> **Coding and simultaneous comparing of units of meaning across categories**

> **Refinement of categories**

> **Exploration of relationships and patterns across categories**

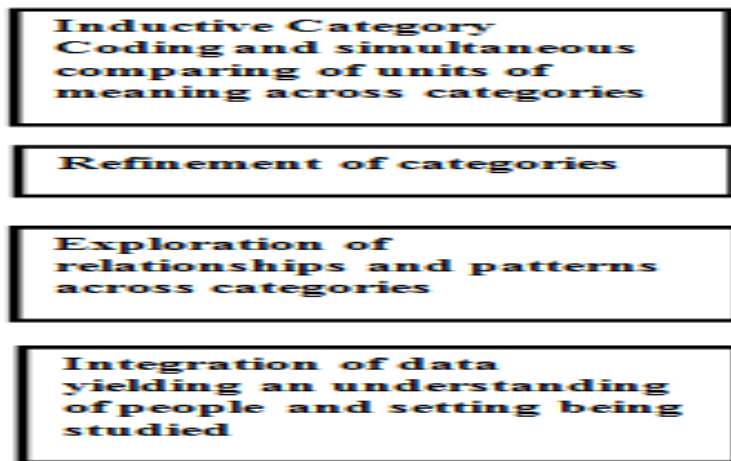> **Integration of data yielding an understanding of people and setting being studied**

Figure-1: Maykut & Morehouse's (1994) Constant Comparative Method

Recently, Maykut and Morehouse (1994) proposed CCM as one of the rigorous way of dealing with interview data. They based their method on the above guideline given by Lincolin and Guba (1985) with some adoptions. They described it in four major steps. These steps may be seen in figure 1.

Each step given in the above figure will be explained in this study with an illustration of an example. The examples are given from a study conducted by researchers themselves on reading attitudes. The study conducted 22 interviews and used CCM for analysis. The details of each step is given next with relevant examples.

STAGE 1: INDUCTIVE CATEGORIZATION

This stage involves reading the interview transcripts carefully to identify an initial identification of recurring themes and concepts. This results in a tentative list of themes. In the second step, any ideas/themes, which overlap with one another, should be combined and subsequently be assigned some provisional codes. That is to create some temporary codes with the aim of grouping themes. For creating these codes, Maykut and Morehouse (1994) suggest using the '*look alike and feel alike*' criteria in checking whether a unit of meaning is very similar to another unit of meaning (Maykut & Morehouse, 1994: 136).

Example From a Study: Using this guideline, the researchers took one interview transcript and created provisional codes for a unit of

meaning found in a sentence, paragraph or in a section. Those specific sentence, paragraph or a complete section were underlined. When it comes to organize qualitative data while analysis, different authors suggest different methods such as color coded paper pieces. The researchers used Microsoft word document in computer to keep the data handy. The provisional code was added in the review note. Figure 2 (Microsoft snapshot: An example of a general rule of inclusion for a unit of meaning in an interview transcript).
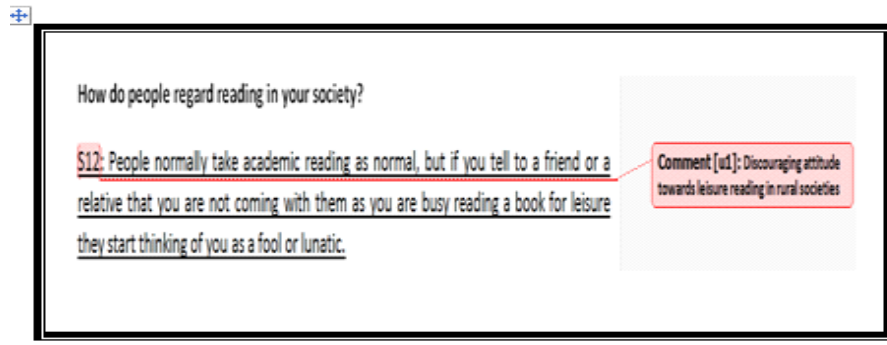

Figure-2
Microsoft snapshot: An example of a general rule of inclusion for a unit of meaning in an interview transcript demonstrates this

After assigning the codes, the researchers checked back and forth between various codes, using '*look alike and feel alike*' criteria, and shifted a few units of meaning from one code to another until a provisional list was made. Same procedure was repeated with all the interview transcripts. Thus, stage 1 ended by providing a modified working list of 34 codes from 22 interview transcripts.

STAGE 2: REFINEMENT OF CATEGORIES
This stage advances from the look/feel alike approach to the propositional rule statement (Maykut & Morehouse 1994:141). A researcher has to work on the provisional list of codes and has to plan the specific rules which would later serve as the criteria to include or exclude any theme from a code. This may be performed this into three steps:
First step involves bringing together all the similar themes, grouped under one provisional category. For this, all the interviews will be scanned and kept together in one place. Next, the researcher needs now to re-read all the pieces and draw some general rules. Lastly, having

prepared the provisional rules, they need to undergo the process of comparing and revising them. Taking the rules on board, all the units of meaning should be rechecked to see whether or not they fit their assigned categories. This results in confirming that each code is placed rightly in the matching category.

Example From the Study: The researchers made a table in a Microsoft file. All the categories and their rule for inclusion are to be recorded in a separate Microsoft Word file. We copied the relevant parts from all the interviews and pasted them under its category. Table 1 illustrates an example of it.

TABLE-1
AN EXAMPLE OF A CODE WITH RELEVANT QUOTES FROM THE INTERVIEW TRANSCRIPTS

Code	Example
Role Modelling by Parents in reading	**S4:** My father himself reads a lot. He reads mostly books on politics. I myself do not like politics a lot and therefore I do not read his books, but his habit has inspired me to follow him and read. **S7:** My father is a professor. I often see him sitting and reading till late night. I am so very much inspired by him. **S3:** Reading is in my blood. If you meet with my father, grandfather, you will hear them discussing about books and authors.

Once these tables for each code with its relevant bits from the interviews are made. The researcher needs now to reread all the pieces and draw some general rules. For example, for, 'Role modeling by parents in reading' the rule of inclusion proposed in researchers' study was: 'Those meaningful units, in which participates are exposed to see their parents engaged in reading tasks. Shall be coded as *Role modeling by parents* in reading'!

Another example from researcher's study is For example, for 'Reading as a way for broadening knowledge about multicultural values', the rule of inclusion was sought to be: '*Those meaningful units, in which participants regard reading as a resource to learn about different cultures, shall be coded as 'Reading as a way for broadening knowledge about multicultural values'.*'

Having prepared the provisional rules, they need to undergo the process of comparing and revising them. Taking the rules on board, all the units of meaning should be rechecked to see whether or not they fit

their assigned categories. This results in confirming them. In researchers' study, 31 categories were finalized from the original 34.

STAGE-3: EXPLORATION OF RELATIONSHIPS ACROSS CATEGORIES
　　　　This stage refines the categories further by grouping them under an umbrella. Categories with common elements are grouped to make broader groups.

　　　　Example From the Study: For example: in researcher's study, overall, some codes dealt with aspects described in the educational set-ups in rural areas. They were: *shortcomings in lack of infrastructure, lack of accountability in teachers, teachers focused on personal tuition centres, problems in commuting to schools, dearth of subject specialists and lack of co-education*. Therefore, they were grouped under the heading 'Shortcomings in the rural educational institutions'.

　　　　This step ends here as suggested by Maykut and Morehouse (1998). But, in the researchers' study, the ratio of occurrence of each unit of meaning in a code was also quantified, with the aim of finding out the maximum and the minimum ratio of participants for any response. Although, Maykut and Morehouse (1994) do not suggest this, some researchers suggest that counting how often codes occur is helpful in clarifying whether reality is in accordance with the overall impressions gained by the researcher (Morgan, 199; cited in Taylor, 2001). Therefore the responses were quantified as it lets us find out the maximum and the minimum ratio of participants for any response. It can enable us to see the strength of a code by giving us the exact number of responses. To perform this, a file was created in Microsoft Excel. Table 2 demonstrates an excerpt of it as an example of this.

TABLE-2
AN EXAMPLE OF THE FINAL CATEGORIES, CODES IN THEM ALONG WITH PERCENTAGE

Category	Codes	Frequency	%
Shortcomings in the rural institutions	Lack of infrastructure	6	46%
	Lack of accountability in teachers	7	54%
	Teachers focused on personal tuition centres	4	31%
	problems in commuting to Schools	5	38%
	Dearth of subject specialists	8	62%
	Lack of co-education	6	46%

This marks the conclusion of this stage.

STAGE 4: INTEGRATION OF DATA

This stage involves the synthesis. The propositions built so far are to be examined carefully to help understand the meaning in the data. One way to understand the meaning in data can be to cluster the relevant categories around the research question they could answer. It would create a complete understanding and patterns of meaning. Thus, the analysis phase ends completely. The summary of this process can be given below in table 3.

TABLE-3
SUMMARY TABLE OF THE QUALITATIVE DATA ANALYSIS

Stage	What needs to be done	Outcome
Start of the process	Transcribing audio tapes Identifying reoccurring ideas	Interview transcripts ready for analysis
Stage#1: Inductive category coding	Making provisional categories based on '*Look alike and feel alike criteria'*	Initial list of categories
Stage#2: Refinement of categories	Making rules of inclusion as propositional statements	List of categories with propositional statements
Stage #3: Exploration of relationships across categories	Combining categories into a few broad ones	Final groups of categories Stand-alone propositions Outcome propositions
Satge#4: Integration of Data	Searching for meaning in the data	Synthesis

PART-II

Apart from reporting of the procedural details of analysis, the question of reliability is also important. It is just as important for qualitative research as it is for quantitative research Reliability determines whether qualitative measurement instruments provide consistent results across different coders, raters or observers (Popping 2010:1068; Tashakkori and Teddlie 1998:75-85). However, the issue of reliability is discussed more often in research studies based on quantitative rather than qualitative data. Nevertheless, some researchers emphasize on the issue of reporting the reliability measures very explicitly in qualitative work (Popping 2010:1068; Tashakkori & Teddlie 1998:75-85).

A number of authors have suggested alternative ways to evaluate qualitative research in terms of objectivity and consistency which proved

valuable to the present study. Yin (2003) usefully suggests documenting as many steps of the analysis procedures in detail as possible. Gibbs (2007) gives a detailed suggestion on the issue of reliability: to ensure reliability by checking transcripts for errors, creating consistent coding by constantly comparing data with codes and writing memos about them, documenting the meetings if working in groups, cross checking codes developed by different researchers by comparing independently derived results.

Following the above criterion, reliability can be ensured in qualitative studies. Next, an example of a study is given from researchers' study.

EXAMPLE OF RELIABILITY CHECK

In one of the studies, the researchers ensured reliability in the following ways:

a) **Checking Transcripts for Errors:** This step aims to confirm that the interviews are transcribed with low inference and present only what the respondents have said. Firstly, in line with Gibbs (2007) discussed above, the English version of the transcripts were read and re-read several times to check for the errors, inferences, and interpretations written instead of the plain translation.

b) Consistent Coding: Secondly, the present study followed the guideline of Maykut and Morehouse (1998) in systematically and constantly comparing the data to ensure the consistency and thus making reliable coding patterns.

c) Inter Coder Reliability or Cross Checking: The concern of inter coder reliability remains to develop coding schemes that are reproducible (Campbell *et.al.,* 2013). It deals with the question of whether different coders would code the same data the same way. This process actually ensures that a single knowledgeable coder may be reasonably confident that his or her coding would be reproducible by other equally knowledgeable coders if they were available (Campbell et al., 2013). Therefore, evaluation of inter coder reliability and agreement should be part of the development of coding schemes for qualitative data in order to satisfy other researchers that the data are sound (Hruschka I., 2004; Krippendorff 2004; Miles and Huberman 1984, Weber, 1994).

Although inter coding reliability is very much emphasized, relatively few resources report the exact way of determining it. For example, Lombard, Snyder-Duch, and Bracken (2002, cited in Campbell 2013) reviewed 137 research articles based on content analysis of various

sorts including interviews. Only 69 percent of these articles contained any information on inter coder reliability. Such omission has been noticed by others too (Fahy 2001; Riffe and Freitag 1997).

Moreover, there is not much guidance in the literature for researchers concerned with establishing reliable coding of in-depth semi structured interview transcripts particularly. The literature available focuses on establishing coding schemes for other types of data like field notes (Miles and Huberman 1984), documents (Krippendorff 2004; Weber 1990), conference discussion transcripts (Fahy 2001; Garrison *et.al.,* 2006), ethnographies (Hodson 1999) and observations of behaviour recorded on videotape (Rosenthal 1987).

In a study by Campbell *et.al.,* (2013), they have given a very detailed and systematic account of inter coder reliability analysis, which is the reason why their approach is reproduced in the present study. The details of the procedure followed for establishing reliability are given below along with illustration from a study.

STEP # 1: FAMILIARITY WITH THE STUDY AND CODING PROCEDURE

To begin with, the researchers contacted a colleague to help develop the coding scheme. Coding semi structured interviews often involves interpreting what respondents mean in their answers to questions (Campbell *et.al.,* 2013). Therefore, it is necessary to familiarize the second coder with the study in as much detail as possible. Keeping this in view, in the first place, a detailed report of the study was sent to the second coder. In the next phase, we arranged to meet in person. The aims of the meeting were twofold:

a) To talk through the report and discuss in detail the study.
b) To explain the purpose and the procedure of coding.

Having discussed the study with the other coder, a transcribed and coded transcript was used to clearly display the coding scheme development process. The researchers explained to him how they identified the units of meaning in that transcript. This was followed by a discussion on the initial pack of codes. Next, we moved towards the crux of the inter coder reliability: coding and comparing the data, which can now be discussed in detail.

STEP# 2: CODING AND COMPARING THE DATA

Following Campbell et al (2013), the second coder and the researcher took the same full length interview transcript to code independently. After they finished coding, they compared the coding

generated. This process brought forward some challenges. They were negotiated and solved consequently.

One of these challenges concerned determining the '*discriminant capability of the coding scheme*' (Campbell et al (2013). This involves determining how well coders can readily and unambiguously categorize text content (Fahy 2001; Kurasaki 2000). In the present study, there were instances where several codes could be applied to the same section of text. For example, We coded the following string from interview no 8 of the rural participants as anxiety *in university classroom*. But the second coder coded it as *Lack of exposure to co-education*.

It is a new experience for me to study with girls; I mean this is so new for me. I feel so conscious of their presence while participating in the class. I feel, what if I say something wrong. I will look like a bumpkin before them. I just shut down my expression.

Rural Participant 8

In terms of how to deal with this kind of coding difference, some scholars (e.g. Garrison *et.al.,* 2006; Campbell *et.al.,* 2013) have adopted a 'negotiated agreement' approach for assessing inter coder reliability where two or more researchers code a transcript, compare codings, and then discuss their disagreements in an effort to reconcile them and arrive at a final version in which as many discrepancies as possible have been resolved. I also adopted this approach. We discussed this case and finally agreed upon coding it as *Lack of exposure to co-education* since we agreed that this anxiety is the outcome of a rural learner's not being exposed to studying in a co-educational system. The presence of girls makes them feel shy. They are reluctant because of lack of co-education in their surroundings. Other differences of this type between the coders were resolved in the same way.

The second issue in establishing inter coder reliability is related to the breaking-up of text sections/units of meaning called 'Unitization' (Krippendorff 1995). Researchers have been debating whether clearly demarcated parts of text, such as sentence, paragraph are the appropriate units of analysis (Garrison *et.al.,* 2006). This may cause problems insofar as different coders may unitize the text differently because they may disagree on which segments of text contain a particular meaning (Kurasaki, 2000). For example, two coders may identify a string of text for the same code but the length of text may differ. One coder includes text providing background information that helps information to establish the context for the code in question but the other coder does not (Fahy,

2001; Krippendoff, 1995). And if coders do not unitize a text exactly the same way, it may become difficult to determine whether their coding is the same.

In the present study, this did not occur frequently. The nature of the semi-structured interview was not very complex. The participants did not use very lengthy answers and the interviewer probed to keep them away from switching the conversation in another direction. However, the problem in unitization - deciding the length of the text segment -did occur occasionally. For example, the following segment:

"My elder brother had a personal collection of books. He always used to keep them in his suitcase. I remember, when we all kids slept, my brother would read books in the light of the lantern. It inspired me always. Slowly and gradually, I started joining him for shorter duration and then I also accompanied him in reading until he slept. That gave me love for reading and it is still there".

Rural Participant 9

Although I kept the length of the text as given above, the second rater unitized it as follows:

I remember, when we all kids slept, my brother would read books in the light of the lantern. It inspired me always. Slowly and gradually, I started joining him for shorter duration and then we I also accompanied him reading untill he slept. That gave me love for reading and it is still there.

Rural Participant 9

With regard to how to deal with such problems, Campbell et al (2013) dealt with the problem of unitization by giving the second coder a copy of the interview transcript in which the units of meaning were already identified by the main researcher and the second coder was asked to code only. But this method may bias the level of inter coder reliability. Therefore, in the present study, already mapped out transcripts were not used. We asked the second coder to identify the units of meaning by himself and then to code them. And the problems such as above were solved through negotiated agreement.

We explained to the second rater that the following part: 'My elder brother had a personal collection of books. He always used to keep them in his suitcase' which he did not include in the text segment actually reflects the life style of the rural people. It mirrors the lack of resources. His brother used a suitcase instead of the book shelf or cupboard because

that is not available to everyone in rural areas. And where it is available, it may not be allowed to use for protecting books as a large family shares it. Therefore, this background information needs to be kept. We agreed upon this and included it.

Thus, the second rater and I kept on coding transcripts independently in this way, then discussed any discrepancies, until towards the end of the process we found our coding was becoming more or less similar. Thus, the inter coder reliability process helped in removing the coding discrepancies, refining codes and improving code definitions. Concrete and precise definitions of codes were developed. In addition, it also helped standardizing the meaningful units.

CONCLUSION

This article is not a rulebook rather it gives an illustration of one of the ways which could be adopted to analyze meanings in qualitative data systematically. This article's contribution lies in showing application of constant comparison method and reporting of the procedure in terms of analysis and reliability. This article contributes by giving a clearer path to other researchers to replicate the procedure. It may make the findings in a study to come out more rigorously and systematically as there is not much guidance in the literature for researchers concerned with establishing reliable coding of in-depth semi structured interview transcripts particularly.

REFERENCES

Bourque, L. & Fielder, E. P. (1995). How to conduct self–administered and mail surveys, Thousand Oaks, CA: SAGE Publishers.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth Semi structured interviews problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42 (3), 294-320.

Dörnyei, Z. (2003). Questionnaires in second language research. London: Lawrence Associates Publishers.

Dörnyei, Z. (2007). Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies. Oxford: Oxford University Press.

Fahy, P. J. (2001). Addressing some common problems in transcript analysis. *The International Review of Research in Open and Distributed Learning,* 1(2).

Garrison, D. R., Cleveland-Innes, M., Koole, M., & Kappelman, J. (2006). Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *The Internet and Higher Education*, 9 (1), 1-8.

Gibbs, G.R (2007) Analyzing qualitative Data. In U Flick (Eds.) *The Sage Qualitative Research Kit.* London: Sage.

Glaser, B. G. & Strauss, A. (1967). The Discovery of grounded theory: Strategies for qualitative research. New York: Aldine.

Hodson, R (1999). Analyzing Documentary Accounts. Thousand Oaks, CA, Sage.

Krippendorff, K (1995). On the Reliability of unitizing continuous data. *Sociological Methodology*, 25, 47-76.

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, 30 (3), 411-433.

Lincolin, Y. and Guba, E (1985). Naturalistic Inquiry, Beverly Hills, CA: Sage.

Liska, A. E. (1984). A Critical examination of the causal structure of the Fishbein/Ajzen attitude behaviour model. *Social Psychology Quarterly.* 47, (1), 61-74.

Logan, S & Johnston, R (2009) Sex differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading*, 32 (4), 199-214.

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587-604

Maykut, P., & Morehouse, R. (1994). Beginning qualitative research: A philosophical and practical guide. London: Falmer Press.

Miles, M. B., Huberman, A. M. (1984). Qualitative Data Analysis: A source book of new methods. Beverly Hills, CA, Sage.

Morgan, D. L. (1993). Qualitative content analysis: A guide to paths not taken. *Qualitative Health Research.* 3,(1), 112-121.

Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research*, 3, 362-376.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research.* 1(1), 48-74.

Popping, R (2010). Some Views on Agreement to be used in content analysis, *Quality and Quantity,* 44, 1067-1078.

Riffe, D. & Freitag, A. (1997). A content Analysis of content analysis: Twenty five years of Journalism Quarterly'. *Journalism and Mass Communication Quarterly*, 74, 873-882.

Rosenthal, R (1987) Judgment Studies: Design, analysis and meta-analysis. New York: Cambridge University Press.

Tashakkori, A. & Teddlie, C. (1998). Mixed Methodology: Combining Qualitative and Quantitative Approaches', Thousand Oaks, CA, Sage

Tashakkori, A., & Teddlie, C. (Eds.) (2009). Foundations of Mixed Methods Research United States of America, *Sage* Publication Inc.

Weber, R.P. (1990). Basic Content Analysis. 2nd ed. Newbury Park, CA, Sage.

Yin, R.K (2003). Case Study research: Design and methods (3rd ed.). Thousand Oaks, CA: Sage

Yoon, J.-C. (2002). Effectiveness of sustained silent reading on reading attitude and reading comprehension of fourth grade Korean students (Unpublished PhD Thesis), University of Georgia.

_____

165