What Makes A Case Study Really Qualitative?: Show Me Your Evidence, Please!

Chongwon Park (Pukyong National University)**
Hye-Won Lee (Korea Institute for Curriculum and Evaluation)


This paper examines 38 qualitative case studies, published from 2007 to 2009, to critically review whether the core notions of qualitative inquiry have been realized in the articles published within last three years. Based on the integrated criteria of evaluation from several researchers, one can conclude that most of the participating case studies in this study lack negative case analyses, which challenge researcher's hypotheses or beliefs. As a result, the credibility of the researchers’ interpretations is seriously threatened. In addition, the notion of triangulation has still not been represented, although participating researchers claim that they did collect data from a variety of sources in their studies. To reach a credible conclusion, all researchers would agree that negative case analysis and data triangulation are not options but necessary conditions, especially in case studies. To fill the gap between the qualitative researchers’ awareness of these standards and their failure to represent the results in incorporated and systematic ways, the researcher of this study postulates that adopting a computer assisted qualitative data analysis system (CAQDAS) may not only enhance the quality of analysis, but make the writing process less complicated.

I. INTRODUCTION

Since Farrell (1994) first introduced and used qualitative inquiry (QI) as the main tool of investigation to explore English as a Foreign Language (EFL) classrooms at a Korean university, QI has not only received attention but gained popularity in Korea as well. As of December 2009, studies using QI have numbered 128 out of 4354 articles, making up

---

* This work was supported by Pukyong National University Research Fund in 2006 (PKS-2006-018).
** Corresponding Author (chongwonpark@pknu.ac.kr)
three percent of all papers published among eight English teaching related journals in Korea. By simply looking at the number, one may conclude that QI in a Korean context is still in its infancy but examination of 128 qualitative-orientated papers shows that half of these (N=59) were published between 2007 and 2009. Therefore, one can conclude that, quite recently, QI has begun rising up as the main tool of investigation in the context of teaching English in Korea. Despite such an apparent increase in numbers of QI articles that call for extensions in scopes of enquiry, whether current practice actually reflects the intrinsic logics-particularly in terms of representation—is a question that remains hitherto unanswered. This paper then deals with this issue, providing critical reviews of previous QI studies as well as providing insights regarding the possible direction of future applications of QI. One major concern here is the requirement of special attention within QI-in order for it to attain the highest possible results in qualitative findings and transferability. Rather than arriving at a state of generalizability, QI lays out all results together with the researcher’s interpretations and leaves the reader to reach precisely the same conclusion. Mindful of this, it shows how the importance of presentation of QI-based data is on equal footing with the task of data collection and interpretation.

Therefore, the purposes of this study are two-fold. The first aim is to delimit critical factors contributing to quality writing in case study reports. Second, by critically reviewing the current practices, this study will set a milestone inscribing on what we have missed from the previous case studies and what should be represented in future QI oriented research. By doing so, the researcher hopes that graduate students and beginning researchers might gain some insights on what to write and appreciate in case studies to ensure the quality.

II. LITERATURE REVIEW

Originating from the traditions of ethnography, grounded theory, phenomenology, case study, life history, action research, and conversation analysis, QI initiated the discussion of focusing on the quality of qualitative research. However, recently, this focus has mostly shifted from fundamental, epistemological, and philosophical levels to more concrete and practical levels of research (Flick, 2008). Compared to its counterpart, quantitative inquiry, qualitative researchers take different stances on the issues of truth, reality, subjectivity, and hypothesis, although tension still exists among different QI camps (Ellingson, 2009). Qualitative oriented researchers assert that positivist approaches that seek to discover ‘truth’ have missed the mark largely due to the non-existence of a universal ‘Truth’ that can be known objectively, and with certainty (Savin-Baden & Major, 2010). To those who take a qualitative approach, multiple realities will always be left over, and interpretation of
these realities might vary according to the observer and the observed. In that context, truth is not out there and predetermined but mutually co-constructed through the interaction and negotiation between the observer and the observed. Based upon this logic, when collecting data from the field, qualitative researchers prefer bottom up, inductive, and grounded approaches to top down and deductive ones. If the data are collected from the field, and theories evolve from the raw data, the researchers can fill the lacuna between theory and reality. Thus, the goal of QI is not to test hypotheses formulated from readings of previous studies but to establish them based on field experience and offer a new theory by leading further studies (Saldana, 2009). This unique approach of QI may have significant implications for TESOL researchers where theory and reality are interacting with each other on a daily basis.

Also, in light of respecting participants’ personal experiences and their emic perspectives, subjectivity is accepted as a key factor in QI, as opposed to the notion of objectivity in quantitative inquiry. In the end, through the description of the participants’ worldviews, reflexivity becomes the hallmark of qualitative inquiry. Since the focus of QI is mostly on individual differences and their uniqueness, its goal is not necessarily to generalize its findings to the whole population. Instead, QI seeks the possibility of transferring its results to the readers. Here, the role of the researcher is not that of omnipotent or all-knowing presence but that of an intersubjective interpreter of what she or he has learned from the field. QI usually aims at a holistic understanding of people’s cultural behavior through an extended period of fieldwork and what Geertz (2003) called thick description: description and interpretation that play close attention to cultural contexts and insider’s perspectives.

As for the notion of generalizing the findings of the study to the whole population, qualitative researchers cast doubt on its applicability because they tend to believe that there is no such thing as ‘fixed reality’ or ‘truth’ out there (Holliday, 2002; Richards, 2003). Rather, they propose logical transferability of the findings of the study to the reader. According to this approach, if the findings of a study make sense to the reader, transferability of a study has been realized, and thereby it has served its purpose as research (Wolcott, 1994). In comparison with the statistical notion of generalization, Seale (1999) emphasizes theoretical generalizations, which emphasize logic rather than probability. Citing Mitchell (1983), Seale advocates “…choosing the case for its power to explain rather than for its typicality” arguing that “idiosyncratic cases can throw general principles into sharp relief” (p.109). Bogdan and Biklen (2007) also support this notion by stating that, “…if qualitative researchers carefully document a given setting or group of subjects, it is then someone else’s job to see how it fits into the general scheme of findings. Even a description of a deviant type is of value because theories have to account for all types” (p. 36).
In the initial stage of investigation, QI emphasizes the relationships between a researcher and his informants. This emphasis mainly stems from the belief that relationships ultimately affect the quality of the data (Seidman, 2006). Enough rapport, equitability, and thoughtful ethical considerations with the informants will definitely guarantee desirable results. So the importance of the relationships is also applicable to the representation and the relationships with QI writers and readers.

Rather than imposing a standardized or universalized format of writing on the researcher, by extending the format, QI stresses diversity, which captures the realities of the field under investigation (Ellingson, 2009). However, the diversity of writing standards does not necessarily mean that QI representations lack rigor. QI researchers have choices on how to represent their qualitative works originating from one’s epistemological and ontological stance: positivist, postpositivist, and interpretivist (Denzin & Lincoln, 2000). Some researchers have also brought up the strong need to see the field of methodology not as an art/science or qualitative/quantitative dichotomy but existing along a continuum from positivism (i.e., scientific research that claims objectivity) to radical interpretivism (i.e., scholarship as art) (Duff, 2008; Ellingson, 2009). From this viewpoint, art and science do not oppose one another; they anchor ends of a continuum of methodology, and most of researchers situate themselves somewhere in the vast middle ground. What really matters is the researcher’s credible conclusion supported by multiple sources of data, methods, and researchers. To make it happen, some qualitative researchers even further suggest the need for extending the notion of triangulation to crystallization. Therefore, it is no wonder to say that transferability does not happen automatically, and researchers’ efforts to realize it in the representation is essential. Without triangulation or crystallization, chances are slim that the readers will reach the same conclusions with equal clarity as the researcher claims in a specific study. Logically speaking, if a research work uses all types of data collection methods, it will eventually enhance the data crystallization. As a result, it is no wonder to say that it will increase the transferability of the study, which is the ultimate goal of conducting qualitative research (Lincoln & Guba, 1985). According to Richards (2003), however, caution is needed in proposing that all qualitative researches require data triangulation or crystallization. Table 1 illustrates this point.
As illustrated in the above table, ethnography, grounded theory, case study, and action research require data triangulation whereas phenomenology, life history, and conversation analysis call for only one source of data collection. Therefore, whether to triangulate or not depends on the focus and tradition of QI that one is implementing in a specific study. In other words, if one conducts a case study with no data triangulation, it might seriously violate the logic of case studies in general and thereby kills the validity of the study (Maxwell, 1996).

QI has also started to receive wide attention from Teachers of English to Speakers of Other Languages (TESOL) researchers for whom issues of cross-cultural communication have particular salience (Holliday, 2002; Park, 2000, 2006; Richards, 2003). Although which method to choose in a specific study purely depends on the types of research questions and the researchers' purposes (Locke, Spirduso, & Silverman, 1993; Maxwell, 1996; Nunan, 1994), QI surely meets the TESOL researchers' diverse research needs. For example, interactionists hypothesized that input, interaction, and output should form an optimal environment where language acquisition takes place (Krashen, 1985; Long, 1982;
Swain, 1985). To prove their hypotheses, they used psycholinguistically based experimental studies. At this point, based on the results of their studies, there seems to be no way of contradicting the fact that the three above mentioned conditions obviously accelerate the language acquisition process and product. However, anyone who has had the experience of observing an EFL or English as a Second Language (ESL) classrooms knows that this is not always the case. It might be quite hard to find a student or students to speak up voluntarily if they don’t understand what a teacher says. Not only input but when it comes to output, the situation seems to be getting worse. Therefore, in this specific case, one can say that a huge gap exists between the literature and the reality.

Regarding the gap between the literature and the reality, Allwright and Bailey (1991) deplore the fact that we know very little from the research on teaching methodology because the results of the studies are contradicting. As a result of it, they propose that EFL or ESL teachers should not prescribe what should be taught in classrooms. Rather, they need to observe and describe what is going on in the classroom. Based upon observation and description, Allwright and Bailey believe that ideal methods can be constructed which not only reflect learner needs but also ultimately guarantee teaching and learning success. According to van Lier (1988), these methods of observing and asking should be the core of classroom-oriented research.

The logics and approaches of QI seem to offer some suggestions for handling the issues of the gap between theories and realities and the notion of how to generalize the findings of the study to the whole population. As for the gap between theories and realities, especially in interaction-oriented studies, interactionists offer some insights on the ideal model where language learning or acquisition takes places rapidly. However, they fail to explain why such an ideal environment is not possible or even desirable for certain groups of learners. In the end, what they created was the conflict between theory and reality.

As discussed above, QI’s unique logics seem to offer fruitful benefits to both social research in general and TESOL specifically. However, little attention has been paid to how these concepts are incorporated and represented in a desirable way. From the previous studies focusing on methodological issues, it becomes apparent that QI has usually been positioned in a quantitative and qualitative dichotomy and viewed as a package lacking its nature of diversity (Joo, 2006; Kim, 2006). Considering each qualitative study reflects its own unique stances, developing criteria, which entail its methodological diversity, is inevitable. This is what generally lacked in the previous studies, and further in-depth discussion is needed. One might argue that caution is also needed to reduce the complexity and “messiness” of qualitative research by way of a series of simplified “how-to” offerings (Bloomberg, 2008). Although the researcher’s intention is to demystify the process of writing a case study, the researcher does not sacrifice intellectual rigor for the sake of simplification. This paper is not intended to be a quick-fix, nor does it offers an easy recipe
for success. Rather, it offers some preliminary considerations for enhancing quality in case study representations. Merriam (1998) defined case studies as processes as opposed to outcomes, in contexts rather than specific variables, discoveries rather than confirmation. Although case studies do share many commonalities with other types of qualitative research, several particulars pertaining to writing up case studies is that they are richly descriptive giving readers vicarious experiences of having been there. According to Merriam (1998), the experience of ‘having been there’ is done through writing a vividly descriptive narrative of the setting and situation.

III. METHOD

1. Participants and the Procedure of the Study

The participants of this study were two primary investigators and four participants who were willing to take part. All six participants had obtained, or were due to obtain, doctoral degrees in TESOL and have trained to critically review and write QI works in their doctoral programs. In the first meeting with the participants of the study, the evaluation criteria—constructed from items under scrutiny, especially from part-two which addressed quality analysis—were distributed. After discussing the evaluation criteria, each participant was given the same article and mandated to evaluate it within a week. One week later, from the second meeting, the participants cross-checked their evaluation on the co-assigned article. The focus was to discuss items that showed more than two scale discrepancies among participants. The second part of the survey applied usages of scale from one to five—ranging from strongly agree to strongly disagree. Overall, the participants expressed their understandings and agreement on the given criteria and there were no instances of any items showing discrepancies of more than two scales in reactions from all participants. Again, after reading the title and abstract of the articles, they were free to choose six or seven articles that drew their interests or were related with their expertise in TESOL. Then, the participants were given three months to critically review and evaluate each of the articles assigned to them. Three months later, the third meeting was held for collection and discussion of individually assigned articles. Specific items of interest were opened up and shared between each evaluator, and additional three weeks were given to ensure each member checked each item. Following a period of seven months, allowing for data analysis and writing-up, a final meeting was held to enable a collective check on all interpretations of research finding. Although the exercise was propelled by the willingness of the participants, gifts cards were awarded as tokens of gratitude to each participant.
2. Sample Studies

This study evaluated papers from eight publishers, written between 2007 and 2009: *English Language & Literature Teaching, English Language Teaching, English Teaching, Foreign Languages Education, Journal of Applied Linguistics, Multimedia Assisted Language Learning, Primary English Education*, and *The Society for Teaching English through Media*. Information on English language teaching related articles was stored on EndNote X4—a bibliographical software system—which found thirty-eight case studies from eight out of a total of ten publishers. The search was conducted in such a way that excluded quantitative and mixed-method modes of research—focusing on qualitative methods only—arriving at a final thirty-eight case studies as the focus of investigation.

3. Evaluation Criteria

The evaluation criteria of the qualitative works for this study are integrated from *TESOL Quarterly* submission guidelines, Holliday (2002), Park (2006), and Richards (2003). Park (2006) merged ideas from the above cited researchers or institutions. In analyzing the qualitative works in this study, these ideas serve a central framework of judgment. The analysis consists of two parts. The details are expounded as follows:

Part I. Demographic Data

1) Primary Locations of the Investigation

The primary location of the investigation was divided into four areas: (1) the classroom, (2) naturalistic environment, (3) mixed, and (4) textbook. The ‘classroom’ refers to the contexts where students, teachers, and formal class hours are the focus of the investigation. In contrast, a ‘naturalistic environment’ refers to the contexts where no formal instruction is administered. Instead, in a naturalistic environment, a researcher collects the data in a certain period of time. ‘Mixed’ entails a setting where both in and out of class contexts and investigations are conducted simultaneously. Finally, ‘textbook’ was also included as a location of investigation for the sake of convenience.

2) The Types of Data Collection

Creswell (2003) proposed research strategies where diverse types of data can be integrated and investigated in a concurrent or integrative manner. He argues that these methods will deepen the level of interpretation of the findings of a study. He
stated, “mixing at the stage of data analysis and interpretation might involve transforming qualitative themes or codes into quantitative numbers and comparing that information with quantitative results in an “interpretation section of a study” (p. 212). According to Creswell, in sequential strategy design, there are explanatory (starting from quantitative to qualitative), exploratory (starting from qualitative to quantitative), and transformative (either explanatory or exploratory) designs. In contrast, in the concurrent strategy, there are triangulation, nested, and transformative designs. Thus, in this paper, instead of creating a qualitative vs. quantitative dichotomy, the types of data collections are classified in more elaborate ways, such as, 1) data elicitation, 2) interview, 3) transcript, 4) survey, 5) diary, journal entry, or field notes, 6) observation, 7) document, and 8) manual vs. computer assisted qualitative data analysis system (CAQDAS). Although there was a total of thirty-eight separate studies, it was possible for each researcher to utilize at least one extra method of data collection: for example; a combination of observations, interviews, and surveys. In such cases, all methods’ outcomes were counted separately in a single study. These methods of data collection will be operationalized in detail as follows.

1) Data elicitation: This refers to the case where a researcher uses figures, tables, or standardized tests to collect the data (Nunan, 1994).

2) Interviews: Compared to the data elicitation, in an interview, without using a specific instrument, a researcher collects the data through the form of dialogue. Types of interview questions are divided into open ended, semi-structured, and structured questions.

3) Transcripts: Without eliciting the data, the researcher transcribes the naturally occurring contexts.

4) Surveys: The collection of data (usually related to attitudes, beliefs, or intentions) from subjects without attempting to manipulate the phenomena or variables under investigation.

5) Diaries, journal entries, or field notes: This method of data collection includes learner diaries, reflectional memos, or field notes written by a researcher.

6) Observations: In reality, almost all of research falls into the category of observation. However, if a researcher writes observation journals or focuses on some form of investigative analysis, it can also be classified as observation. In addition, two different types of observations are identified in this study, that is, participant or non-participant observation. These different types of observations have something to do with the level of involvement of a researcher in a specific study. For example, if a researcher observes the class and writes memos at the back of the classroom, it could be a typical example of non-participant observation. If a researcher participates in the classroom activities by performing tasks with the students, this could be an example of participant observation.
(7) Document: Rather than inventing an investigating tool, the researcher provides evidence which is directly or indirectly related to the research and supports the researcher's claim, for example, a literature review.

(8) Methods of data analysis - Manual vs. computer aided qualitative data analysis system (CAQDAS): If a researcher uses CAQDAS as the main tool of investigation, he needs to mention the names of CAQDAS, for example, Nvivo 2, 7, or 8, NUD*IST 4 or 6, The Ethnograph 5.08, Atlasti, Hyperresearch 2.6, or MAXqda (Park, 2004). Unless the name of a specific software is mentioned, it is assumed that a researcher analyzed his data manually. In research, all of the above-mentioned types of data can be collected either exclusively or complementary.

3) The Length of Investigation Period

The length of the investigation period refers to the time spent in a study. They were divided into four ranges. These are (1) under one day, (2) several days, (3) one semester, and (4) more than one semester.

4) Participants of 38 Articles

The participants of the study were divided according to the age or social status, as follows:
(1) kindergarten, (2) primary, (3) middle school, (4) high school, (5) college or university, (6) adult including graduate students, (7) teachers, (8) parents, and (9) others.

5) The Topic of the Investigation

The topic of the investigation deals with the five language skills. These are (1) speaking, (2) writing, (3) listening, (4) reading, and (5) culture. In addition to these five major domains of TESOL, other topics were also included.

Part II. Quality Analysis

In this part, unlike part one where categories seem to be obvious, a more in-depth analysis will be administered. While reading the participating papers in this study, participants' highly critical and integrative thinking skills were strongly demanded. To differentiate between the participants' judgmental levels for each item, five point likert scales were administered. Guided by previously discussed studies of pertinence, researchers of this study developed specific items to survey in order to elicit information
and conduct quality analysis in a case study representation. Items addressed include relationships between researcher and informant; negative cases to support the validity of the researcher’s interpretation; the rationale of using a qualitative method; emerging themes; contribution to understanding second-language cultures and schools; educational implications of the findings of the study; thick descriptions and transferability; data triangulation and actual representation; and ensuring anonymity. Full descriptions of survey items are displayed in the appendix section of this study.

IV. RESULTS

In this section, the results of the study will be reported according to the order of the survey.

1. Demographic Data

1) Primary Locations of the Investigation

Among the 38 articles evaluated, 61 percent of the articles’ (N=23) primary locations of investigation were carried out in the classroom. Other locations are naturalistic (N=11), mixed (N=2), textbook (N=1), and on-line (N=1) respectively.

2) The Types of Data Collection

The most frequently used sources of data collection were document (22%) followed by observation (20%), interview (18%), transcript (16%), survey (10%), diary (10%), and data elicitation (5%). Regarding interview methods, researchers used semi-structured interviews more frequently (41%) than open-ended (30%) or structured ones (19%). Among the papers evaluated, eleven percent of researchers did not report their interview methods explicitly. Also unknown or unclear is the method of data analysis except one in article where Statistical Package for the Social Science (SPSS) was used and reported. Therefore, it can be assumed that most of the researchers analyzed their qualitative data by hand.

3) The Length of Investigation Period

Half of the studies (55%) took one semester investigation period whereas 21 percent of articles did more than one semester of investigation. Eight percent of the studies took
several days, and three percent of the studies invested less than a day. However, thirteen percent of the studies did not explicitly report their investigation period.

4) Participants

Twenty four percent of the studies involved teachers, and primary (18 percent), college or university students (18 percent), and 16 percent were adults. Few studies involved kindergarten (6 percent), high school (6 percent), middle school (4 percent), and others (8 percent).

5) The Topic of the Investigation

Among the sample studies, fifteen different topics including four English language skills were investigated. Table 2 capitalizes each topic.

<table>
<thead>
<tr>
<th>Topic and Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>speaking (23%)</td>
</tr>
<tr>
<td>reading (21%)</td>
</tr>
<tr>
<td>writing (19%)</td>
</tr>
<tr>
<td>listening (12%)</td>
</tr>
<tr>
<td>culture (4%)</td>
</tr>
<tr>
<td>immersion (4%)</td>
</tr>
<tr>
<td>teacher training (2%)</td>
</tr>
<tr>
<td>teacher reflection (2%)</td>
</tr>
<tr>
<td>home literacy development (2%)</td>
</tr>
<tr>
<td>emotional issues (2%)</td>
</tr>
<tr>
<td>textbook (2%)</td>
</tr>
<tr>
<td>story telling (2%)</td>
</tr>
<tr>
<td>on-line graduate course (2%)</td>
</tr>
<tr>
<td>curricular reform (2%)</td>
</tr>
<tr>
<td>native speaker fallacy (2%)</td>
</tr>
</tbody>
</table>

As one can see from the table, besides the four skills of language learning which compose 75 percent of the total studies, although minuscule in number, diverse topics (25 percent) were investigated as well. Among the four skills, a lot of researchers used QI when investigating speaking. Regarding other topics, it seems that diverse topics emerged by responding to researchers’ various needs.

2. The Results of Quality Analysis

Starting with five distinctive ranges on the likert scale survey, the researcher merged them into a more simplified scope: agree or disagree. In the survey, ‘strongly agree’ and ‘agree’ were merged as ‘agree’ and ‘strongly disagree’ and ‘disagree’ were identified as
What Makes A Case Study Really Qualitative?: Show Me Your Evidence, Please!

‘disagree’. Neutral responses were discarded as an object of analysis and therefore, the total numbers of responses do not necessarily add up to thirty-eight due to some responses being discarded. Table 3 reports the results of the analysis according to the order of disagreement for each item.

TABLE 3  
The Order of Disagreement among 11 Survey Items

<table>
<thead>
<tr>
<th>Item</th>
<th>Agree</th>
<th>Disagree</th>
<th>Difference (Agree - Disagree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Validity of researcher's interpretation</td>
<td>5</td>
<td>33</td>
<td>-28</td>
</tr>
<tr>
<td>2. Negative case analysis</td>
<td>5</td>
<td>31</td>
<td>-26</td>
</tr>
<tr>
<td>10. Illustration of data triangulation</td>
<td>10</td>
<td>28</td>
<td>-18</td>
</tr>
<tr>
<td>1. Researcher and participant relationship</td>
<td>16</td>
<td>19</td>
<td>-3</td>
</tr>
<tr>
<td>6. Smoothly emerged theme</td>
<td>17</td>
<td>19</td>
<td>-2</td>
</tr>
<tr>
<td>4. Rationale of using QI</td>
<td>19</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>5. Validity of the researcher's observation</td>
<td>20</td>
<td>17</td>
<td>3</td>
</tr>
<tr>
<td>9. Thick description</td>
<td>22</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>11. Anonymity</td>
<td>25</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>7. Contribution of a study</td>
<td>19</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>8. Explicitly stated implication</td>
<td>34</td>
<td>4</td>
<td>30</td>
</tr>
</tbody>
</table>

As illustrated in table 3, items which most of the participants disagreed with were validity of the researcher’s interpretation (33 out of 38), negative case analysis (31 out of 36), and illustration of data triangulation (28 out of 38). In contrast, the most agreeable items were explicitly stated implication, using anonymity, thick description on the context of a study, and stating contribution of a study.

V. DISCUSSION

This paper examines 38 case studies to critically review whether the core notions of qualitative inquiry have been sufficiently represented in the articles (published within last three years, from 2007 to 2009). By doing so, the researcher hopes that novice researchers in QI can get some insight as to what should be included in QI writing and consequently grasp the concept of what to write in their future QI based investigations. The summary of findings is as follows. First, in terms of primary locations of the investigation, there exists balance between locations in and out of classroom, and even diverse settings like on-line learning are included. Second, there is also a balance of many different kinds of data collection methods, although how the data are analyzed is not explicitly stated except for...
one article. Third, although it is an ideal situation to conduct long-term investigation in QI, only twenty one percent of the sample studies conducted investigations of more than one semester in length. Fourth, in reference to the participants of the study, twenty four percent of the studies involved teachers. In addition, other participants like primary, college or university, or adults are somewhat balanced. Relatively fewer participants are involved from kindergarten, middle or high school settings. Fifth, sample papers entail 15 different topics although four skills of language are dominant issues.

The quality analysis revealed that the participating studies only partly reflected the notions of QI in their writings. Three notions that received negative responses from the participants of this study need to be addressed in this section. These are actual representations of data triangulation and negative case analysis, which eventually threaten the validity of the researcher’s interpretation (Lincoln & Guba, 1985; Maxwell, 1996). Responses from the participants of the study showed how the majority of sample studies, thirty-three out of a possible thirty-eight, actually failed to confirm the validity of the researcher’s interpretation. In other words, there is no way of knowing the possibility of whether the results of their studies are right or wrong. In order to assess the validity of the researcher’s interpretation, it is generally agreed that data triangulation is the prerequisite. As a means of enhancing the validity of QI, Maxwell (1996) emphasizes data, method, and theory triangulation, a relevant literature review, and feedback and member check. Although most of the participating studies did a very good job of relating their results to the relevant literature review, a few studies did not provide the results of the data or method triangulation, feedback, and member check. Lincoln and Guba (1985) also emphasize that these are essential elements in enhancing the credibility of QI. It is undeniable that a credible study will have more possibilities of transferring the results to other contexts.

As a way of enhancing credibility of the studies, sample papers’ researchers explicitly mentioned the data triangulation process in the data collection section of the studies. However, they failed to represent the integration of different sources of data collection in the findings sections of the studies, including twenty-eight out of thirty-eight articles. Regarding other triangulation techniques, only one researcher showed an excellent job of describing and showing member check in her study. Nonetheless, no study used researcher triangulation. To make matters worse, in the worst cases, even a single source of data was not reported holistically and the findings showed only a part of the theme.

In other words, the choice of which themes to represent was predetermined by the researcher, and the analytic process is hidden or mysterious. Therefore, readers might ponder whether the themes that a researcher introduces in his study emerged from the raw data, or were imposed and selected by the researcher’s impression of the data. To positivists, this veiled process has made it impossible for them to replicate the study and
What Makes A Case Study Really Qualitative?: Show Me Your Evidence, Please! 93

check the validity of the findings. The disappearance of triangulated data in the representation might be partially attributable to the nature of the data handling process of QI. When conducting qualitative research, researchers usually encounter enormous numbers of unstructured data. If one conducts it manually, it is very hard for a researcher to manage different sources of data. As a result, this difficulty can also affect the writing process of QI especially where data triangulation is the norm as in case studies. In addition to the inadequacy of displaying data triangulation, for some extreme cases, researchers provided their interpretations of statistical testing or survey results without even providing the raw data and tables! In that case, it can be hard to avoid the criticism that their definition of case study was somewhat misunderstood or abused. Considering the fact that almost all of the participating papers of this study do not show any evidence of using CAQDAS, one can speculate that the researchers of the 28 articles might have difficulty in incorporating and reporting different sources of data in the finding sections of the studies. No one will deny that providing different sources of data is a way of showing evidence, and failing to do so is a lack of evidence. Without evidence, no matter what the researcher’s claim is, it is doomed to be void. If this is the case, the researcher of this study proposes to adopt CAQDAS to enhance both the systematicity of data management and analysis and the efficiency of the writing process in QI in a balanced manner.

Among CAQDAS programs, NVivo 2 or 8 is optimal for managing and representing data effectively (Park, 2004). Following figure illustrates the point.

![Matrix Coding Query Results](image)

In the above figure, the left column shows five codings assigned in an exemplified study. The themes identified are not desirable way of handling group discussion, not ready for class-no reading, etc. Although five codings are illustrated for the sake of demonstration, the possible number of codings in NVivo are limitless. In addition, the right column shows different sources of data collected in the study, such as diary, survey, classroom observation memos, or field notes. Contrary to SPSS, from NVivo, one can simply click on the numbers, return the original document, and modify the codes. From the figure, a researcher can see the total numbers of codings and their sources, whether they are single or triangulated. For example, coding number 1, not desirable way of handling group
discussion has three codings from the single source (diary). However, coding number two, not ready for class-no reading, has thirteen codings from diary, survey, and classroom observations. In drawing the conclusion of the study, coding two might give more plausible and rich explanations on the data. Also, in the beginning stage of representing QI results, it will help researchers to decide on what to write. The recent version of NVivo 9 also makes QI researchers even visualize multiple sources of data (Richards, 2010).

In spite of CAQDAS’s potent to improve QI processes and products, Seror (2005) points out that not all researchers are comfortable with using it in their data analysis mainly because of the predetermined structure of the programs and cost and time effectiveness. The structure of the data analysis tool is intertwined with the writing format of QI because different types of QI not only require different coding techniques but affect analysis and representation as well. If one is involved in phenomenology, life history, or conversation analysis, where handling multiple sources of data is not required, and the identified themes are more than words or phrase levels, CAQDAS may not be an essential tool. However, for other types of QI where triangulation is the priority, using CAQDAS is a necessity because it will definitely reduce time spent on the analysis and representation. Within seconds, the researcher can jump from a key concept in the analysis to the original data, to a memo recording the thoughts that helped develop the concept (Saldana, 2009). The CAQDAS therefore enhances closeness to the data, helping the researcher create thick descriptions (Geertz, 2003) that are crucial to meaningful accounts of the data.

Among the evaluated studies, another almost equally neglected method of enhancing the credibility of study results is negative case analysis. In addressing the importance of including negative case in a study, Lincoln and Guba (1985) emphasized that the technique of negative case analysis does provide a useful means to make data more credible by reducing the number of exceptional cases to zero. By not showing exceptional cases most of the researchers did not persuade readers well enough in accepting their conclusions. Also few sample studies reported feedback and member check in their representations. Through feedback and member check, one can reveal differences as well as similarities in the participants’ and researchers’ perspectives. To Saldana (2009), tightly woven interpretations that show little sign of human differences are less interesting and believable than those that are less impeccably portrayed. At least for the case of sample studies, these notions were not applied at a satisfactory level, and in some extreme cases, participants’ voices were devoiced and only researchers’ interpretations were prevalent throughout the representations. As Saldana (2009) emphasizes, QI oriented researchers should prioritize and honor the participant’s voice and represent not the researcher’s own summary but what he or she has heard. Yet, beyond the sample studies, in too many write-ups of qualitative findings of studies where the researcher’s stated purpose was to showcase the participants,
What Makes A Case Study Really Qualitative?: Show Me Your Evidence, Please!

the views or voices of the researcher still prevail, or the views or voices of different participants are quickly reduced to one (Sandelowski, 1998).

This study is limited in the light of its scope under QI and the period of sample studies. As defined earlier, QI indicates a research paradigm, and under it, there are diverse research methodologies, which reflect the nature of their paradigms. Among seven research methodologies of QI, this study exclusively covered case studies. Also, since the sample studies were conducted in relatively short-term periods of time, the results may not yield a very solid conclusion at this moment. Therefore, these limitations call for further investigation by extending the scopes and periods of sample studies. By doing so, further research will broaden and deepen our level of understanding of QI as a whole picture.

VI. CONCLUSION

Methodological diversity in TESOL and heated discussion from different perspectives should be accepted as a welcome sign in that it will complement different perspectives and offer valuable, unique insights to the field. However, this complementation will only be possible on the condition that different methods do not sacrifice their own rigor. If so, the results of the study will not reflect anyone’s reality. To avoid these flaws, this study questioned what the most important considerations are in conducting and representing case studies. Despite the increase in numbers, the violations of the rigor of case study in the process of conducting them are predominant at this point. In addition to this, lack of diversity in terms of representing case study results suggests future experimental writing-ups in QI studies.

In an EFL context, the question of how to represent case studies has rarely received attention and has been taken for granted under the influence of positivists’ standard form of writing up the research report. If the emphasis is not on the simple report but sharing the results to co-construct meanings with the readers, how to represent them also should be a major concern. Unfortunately, most of the sample studies lacked diversity in representation, and experimental write-ups in case studies using different genres are highly recommended. The diversity in representation is especially urgent and relevant in an EFL context if the consumers of research reports should not be limited to expert groups. No one would deny that teaching and learning English is everyone’s concern in Korea. If so, it becomes the researcher’s responsibility to share and embrace people in general, by providing reader friendly, jargon free, and interesting story telling with different kinds of genres to make transferability feasible.
REFERENCES


Above mentioned references indicate the sample studies discussed in this study.

APPENDIX
Survey items undergoing quality analysis in a case study representation

1. Are the relationships between researcher and informant explicitly stated?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

2. Were negative cases that negate researcher's hypotheses or beliefs found and reported?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

3. Does the negative case that the researcher showed support the validity of the researcher's interpretations?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

4. Does the researcher state the rationale of using a qualitative method in his or her study?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

5. Is the interpretation based on the observation valid? Other studies or theories related to the current study are mentioned?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

6. Does the theme smoothly evolve from the raw data? Is the theme systematically organized and reported?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

7. In terms of understanding second language culture, do the results of this study make a great contribution to understanding second language culture and school?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

8. Does the researcher explicitly state the educational implications of the findings of the study?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree
9. Were the research contexts or settings explained in-depth? That is to say, does the researcher thickly describe the study so that the results of the study can be transferred to other settings?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

10. Did the researcher not only describe the sources of data triangulation, but also integrate and actually report them in the finding section of the paper?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

11. Ethical concerns: Was anonymity ensured?
   ① strongly agree  ② agree  ③ neutral  ④ disagree  ⑤ strongly disagree

Application levels: graduate, adult education
Key words: research methods, Computer Assisted Qualitative Data Analysis System (CAQDAS), data triangulation, qualitative representation

Chongwon Park
Dept. of English Language and Literature
Pukyong National University
599-1 Daeyon Dong, Nam Gu
Pusan 608-737, Korea
Tel: (051) 629-5386/H.P.: 010-3044-0769
Fax: (051) 629-5373
E-mail: chongwonpark@pknu.ac.kr

Hye-Won Lee
Division of English Education Research
Korea Institute for Curriculum and Evaluation
Jeongdong Bldg. 15-5 Jeong Dong, Jung Gu
Seoul 100- 784, Korea
Tel: (02) 3704-3641/H.P.: 010-3947-4112
Fax: (02) 3704-3954
E-mail: jhwlee@kice.re.kr

Received in August, 2010
Reviewed in August, 2010
Revised version received in November, 2010