



## The Quality of Teacher-Made Test in EFL Classroom at the Elementary School and Its Washback in the Learning

### AUTHORS INFO

#### Desrin Lebagi

Universitas Sebelas Maret  
[desrinlebagi@yahoo.co.id](mailto:desrinlebagi@yahoo.co.id)  
+6282293742423

#### Sumardi

Universitas Sebelas Maret  
[arif\\_sumardi74@yahoo.co.id](mailto:arif_sumardi74@yahoo.co.id)

#### Sudjoko

Universitas Sebelas Maret  
[desrinlebagi@yahoo.co.id](mailto:desrinlebagi@yahoo.co.id)

### ARTICLE INFO

o-ISSN: 2502-6909

p-ISSN: 2502-9207

Vol. 2, No. 2, December 2017

URL: <http://usnsj.com/index.php/JEE/article/view/2.2.97-104>

Received	:	12 June 2017
Reviewed	:	27 August 2017
Published	:	13 December 2017

© 2017 JEE All rights reserved

### **Suggestion for the Citation and Bibliography**

#### *Citation in text:*

Lebagi, Sumardi, & Sudjoko (2017)

#### *Bibliography:*

Lebagi, D., Sumardi, & Sudjoko. (2017). The Quality of Teacher-Made Test in EFL Classroom at the Elementary School and Its Washback in the Learning. *Journal of English Education*, 2(2, Dec.), 97-104.

### Abstract

One of essential phases in language learning is measurement. Test as a tool of measurement process must then be well constructed. The quality of test itself can be determined through test item analysis. However, in some occasions, teachers tend to ignore test item analysis because of time limitation and other responsibilities. Referring to this problem, this research aimed to describe the quality of test items including the difficulty index, the discrimination index, the distractor index, and the reliability of the test and the Washback of teacher-made test on students' motivation in learning English. It was conducted at Gamaliel Elementary School in academic year of 2016-2017. This case study utilized purposive sampling. In collecting the data, the researcher used interview, observation, and document analysis as the techniques of collecting data. The informants were an English teacher and students of Gamaliel Elementary School. The documents were students' answer sheets. In analyzing test items, the researchers used ITEMAN program. The result of this study shows that the teacher-made test can be classified in good test. The test brings both positive and negative Washback in students' motivation in learning. Therefore, it is recommended for the teacher to conduct test analysis as a way of evaluating and improving his teaching and learning and test itself as well as to encourage the students to study even though they are not confronted with a test.

**Keywords:** teacher-made test, washback on teaching, washback on learning

## A. Introduction

In measurement, a test plays important role as the tool in measuring the success of teaching and learning activities in which the result of the test itself can be valuable information for the teacher and the students. Again, Gronlund & Linn (1990) define that test is an instrument or systematic procedure for measuring a sample of behavior. In relation to the benefit of measurement, Peraturan Pemerintah R.I No. 19/2005 tentang Standar Nasional Pendidikan Pasal 64 Ayat 1 dan 2 states:

*Ayat (1): penilaian hasil belajar oleh pendidik sebagaimana dimaksud dalam pasal 63 ayat (1) butir (a) dilakukan secara berkesinambungan untuk memantau proses, kemajuan, dan perbaikan hasil dalam bentuk ulangan harian, ulangan tengah semester, ulangan akhir semester, dan ulangan kenaikan kelas. Ayat (2) penilaian sebagaimana dimaksud pada ayat (1) digunakan untuk: 1) Menilai pencapaian kompetensi peserta didik; 2) Bahan penyusunan laporan kemajuan hasil belajar; dan 3) Memperbaiki proses pembelajaran.*

"Government Regulation No. R.I. 19/2005 on National Education Standards Article 64 Paragraphs 1 and 2 states, "Paragraph (1): assessment of learning outcomes by educators referred to in article 63 paragraph (1) point (a) is carried out continuously to monitor the process, progress, and results of the improvement in the form of daily repeat, semester midterm, semester final repetition, and grade repetition. Paragraph (2): The assessment referred to in paragraph (1) is used to: 1) Assess the achievement of the competence of learners; 2) Material preparation of progress reports on learning outcomes; and 3) Improving the learning process."

This regulation includes national education standard dealing with educational assessment. It is written that educational measurement is conducted simultaneously to monitor the process, progress and improvement in form of daily test, mid-test, and final test. Furthermore, assessment as stated in verse 1 is used to assess students' achievement as well as the basic in reporting students' improvement. Also, it is used as the stepping stone in improving learning process. By looking at the rule, it is clearly written the expected function of test result for the improvement of teaching and learning. Furthermore, Purwanto (2009) states that the result of test is often used as a reflection of how far students master material taught. In addition, Schritchfeld in Yusuf (2015) argues that assessment is a systematic collection, review, and the use of information about educational programs undertaken for the purpose of improving students' learning and development. Furthermore, Remmer in Arifin (2016) points out that test result can be used to help the students to understand themselves better, explain pupil growth and development to their parents and assist the teacher in planning instruction. Without administering test, having good teaching method, technique or media is useless because teacher cannot see how far the students mastered the material delivered as well as the improvement of his or her students.

Ideally, a good teacher should not only be able to provide teaching material and to test the students' comprehension about it, but also be able to analyze the quality of his or her own test. Remmers, Gage, & Rummel (1965) point out that since a classroom test can be used as a teaching-learning experience, item analysis data can have value in class discussions as aids in clarifying the intended meaning of items misunderstood by some students. Simultaneously, Malcolm (2009) argues his point of view about the importance of test analysis. He states that by analyzing the test, you can gain information on how well you taught the concepts to your students. Test items can tell you many things about student performance. Similarly, Arifin (2016) states that test items must be constructed in appropriate principle and procedure in order after administering the test, it can be judged whether or not it has good quality. To know the quality of the test items, then, it is needed to conduct test items analysis. That is to say, test item analysis is an essential process that provides information about students' strength and weakness, and also the quality of the item itself.

However, in several occasions there are many teachers who can provide good teaching material and test it to the students, but they do not analyze the quality of their own test that contain the reliability coefficient, the difficulty index, the discrimination power, and the distractor index. They still lack of awareness about the importance of test analysis. Another reason, time constraint and other responsibilities force the teachers not to analyze manually their test items referring to several characteristics of good test item. Whereas, the result of test

analysis can be a stepping stone for them in improving their test items as well as informing teacher about the quality of their teaching and learning. Arifin (2016) states his point of view about teacher's awareness on the quality of the test. He states that practically, in evaluation conducted at school, teachers tend to ignore about the quality of their test items.

Consequently, this research investigated the teacher made test using Item Analysis (ITEMAN) to see the quality of the test without consuming time as well as investigated the Washback of the test specifically to the learning improvement. ITEMAN stands for Item Analysis. It is a device or program that helps the teacher to analyze item and test easily. This program is included in one package with MicroCat developed by Assessment System Corporation began in 1982 and revised in 1984, 1986, 1988, and 1993: from version 2.00 to version 3.50. Washback is a technical term about how testing influences teaching and learning. It exists in any type of assessments in which test result affects both test-taker and test-maker futures. Knowing the result of test will enable teacher and students in determining what to do for future teaching and learning. In line with the role of test Washback, Pearson in Fulcher & Davidson (2007) argues that public examination influences the attitudes, behavior, and motivation of teachers, learners and parents. In addition, Cheng & Watanabe (2004: 7) claim that it is essential that the education community work together to understand and evaluate the effects of the use of testing on all of the interconnected aspects of teaching and learning. Therefore, in this research, the researchers investigated the Washback of test in teaching and learning, specifically on students' motivation in learning English. Commonly, there are several ways to analyze the quality of test items. The ancient way of test item analysis is by computing the criteria of good test manually following particular formula that is extremely consuming time. Recently, there are some programs that can help the test-makers in analyzing their test items such as, Pearson correlation in SPSS 11.5, Epinifo 7 software and ITEMAN program. From those programs, in this research, ITEMAN program was utilized to compute the quality of test. In conducting this study, the researcher analyzed teacher made-test of an English teacher (Mr. I) that was administered in the final examination of the first semester in the academic year of 2016-2017 at Gamaliel Elementary school, Palu. Several aspects of test item analysis analyzed in this study were reliability, difficulty index, discrimination index, and distractors index. Those criterion were computed statistically by using ITEMAN and interpreted descriptively. Furthermore, to gain data about the Washback of test on learning, the researcher addressed open-ended interview questions to the English teacher as the test-maker, and some students as the test-takers.

## B. Literature review

Test is an essential instrument in measurement process to gain information about pupil growth. Allen & Yen (1979) argue that a test may be defined as a tool to obtain the data about students' behavior such as attitudes, achievement and interest. In addition, Gronlund & Linn (2000) define that test is an instrument or systematic procedure for measuring a sample of behavior by posing a set of questions in a uniform manner. Similarly, Arikunto (1996) states that test is a process of collecting information about students' achievement. Furthermore, Brown (2003) defines that a test is a method of measuring a person's ability, knowledge, or performance in a given domain. Regarding those definitions, it can be concluded that a test is an important tool of measurement process consisting several question items aims to gather information about students' achievement, attitude and interest.

Dealing with the construction of test itself, test is divided into standardized test and teacher made test. Standardized test is constructed by eligible people, in Indonesia it is called Badan Standar Nasional Pendidikan (BSNP), while teacher-made test, as its name, is a test that is constructed by the teacher. Ahman & Glock (1971) state that standardized test is a test that is carefully constructed almost invariably by teams of individuals rather than by a single person. Moreover, Popham (1995) points out that a standardized test is a test, either norm-referenced or criterion-referenced, that is administered, scored, and interpreted in a standard manner. In addition, Basuki & Heriyanto (2014) argue that standardized test is a test with high standard including reliability and validity. To sum up, standardized test is a test that is carefully constructed by competent people, being administered and interpreted in standard manner, and has good quality. Moreover, Ahman & Glock (1971) state that classroom tests are sometimes called teacher-made test or "informal" test that is constructed by a classroom teacher for use in

his particular classes under condition of his choosing. Teacher-made test is considered more applicable because the one who knows classroom situation such as, individual pupils' ability and growth is the teacher. Additionally, Arifin (2016) argues that teacher-made test is a test constructed by teacher who is going to utilize the test itself and it aims to measure students' mastery on material taught. Commonly, it is administered in daily test, formative test and summative test. By underlying the name of test itself, the writer can conclude that teacher-made test is a test that is constructed by the teacher and will be administered to measure students' mastery after being taught in particular period.

Washback or backwash is a term that is commonly used in applied linguistics and educational evaluation. It refers to the influence of test on teaching and learning. It is supported by some definitions from different experts. Firstly, Hughes (1989) argues that the effect of testing in language teaching and learning is called as Washback. In line with Hughes, Spolsky in Ahmad & Rao (2012) claims that Washback deals with the effects of examination in language teaching. Additionally, Messick in Fulcher & Davidson (2007) states that Washback refers to the extent to which the introduction and use of a test influences language teachers and learners to do things that they would not otherwise do that promote or inhibit language learning. Furthermore, Bailey in Pan (2009) defines Washback as the influence of test on teaching and learning. In short, it can be concluded that Washback is the effects of test in language teaching and learning.

Besides doing a measurement, one of procedures that should be done at the end of learning is test item analysis. In this phase, all test items administered in measurement process are analyzed to see the quality of the test itself in which the result of the analysis might be a valuable information for the teacher, as the maker to improve his learning as well as improve the quality of his test items. Ahman & Glock (1971) state that reexamining each test item to discover its strengths and flaws is known as item analysis. Item analysis is a statistic procedure in judging the quality of test items in which its numerical result can be used as to improve test itself. In addition, Thompson & Levitov in McCowan & McCowan (1999) argue that item analysis uses statistics and expert judgment to evaluate tests based on the quality of individual items, item sets, and entire sets of items, as well as the relationship of each item to other items. It investigates the performance of items considered individually either in relation to some external criterion or in relation to the remaining items on the test.

In relation to ITEMAN program, Hadi (2008) explains that ITEMAN program requires that input data be formatted in ASCII (text-only) files. This can be accomplished through the use of the ITEMAN for Windows text editor, Notepad, a word-processing editor that produces true ASCII output, or a program written specifically to format your data. The result of ITEMAN analysis includes numerical information about the quality of each item. Firstly, Seq. no and Scale-Item columns that refer to the number of item. Secondly, Prop-Correct in third column refers to difficulty index of the item. Furthermore, numerical result in Point Biser column interprets the discrimination index of the item. Then, Alt. column refer to the optional answers of each item, for example, A, B, C, and D. Furthermore, Prop-Endorsing and point Biser columns that explain the distracters quality of each item. If the point Biser value is positive, the distractor does not function well. Meanwhile, if the distracter is negative, it functions well. Moreover, key column refers to the correct option of each item that is noted by icon "\*".the reliability of a test is presented in statistic scale as shown in figure 2. 4. Lastly, the reliability can be identified by comparing Alpha index with coefficient of reliability. Since the test was teacher-made test, the coefficient of reliability used is 0.7. If the Alpha is  $\geq 0.7$  (coefficient of reliability), the test is reliable. If the Alpha is  $< 0.7$ , the test is unreliable.

## C. Methodology

### 1. Research design

The design of this study was case study. Yin (2003) defines case study as qualitative approach which investigates a contemporary phenomenon within its real-life context, specifically when the boundaries between phenomenon and context are not clearly seen. In addition, Yin (2003) explains three types of case study categorized based on the purpose namely, exploratory, descriptive and explanatory. This research used descriptive type of case study. Descriptive case study presents a complete description of phenomenon within its context. In this case, the researcher analyzed the quality of teacher-made test of an English teacher at

Gamaliel Elementary school using ITEMAN program and investigated the test Washback to the teacher's way on teaching and students' motivation in learning English.

### 2. *Source of Data*

Glesne & Peshkin in Puspitasari (2015) suggest three kinds of data resources, (1) events, a set of related activities that people carry out, (2) informants, are indispensable partners in the conduct of qualitative inquiry and play a variety of roles, and (3) documents that provide both historical and contextual dimension to the observation and interview to enrich what is seen and heard by supporting, expanding, or challenging portrayal and perceptions. For the purpose of this study, the data were gathered from teaching and learning activities as events, an English teacher and ten students as the informants, and students' answer sheet as the documents.

### 3. *Technique of Data Collection*

In collecting the data, the author conducted observation, interview and documents analysis. Observation was conducted to gain data related to the test Washback on learning. In addition, interviews were conducted to gather data about the Washback of test on teaching and learning and to gain supporting data related to the quality of teacher-made test administered in final examination. The author also conducted documents analysis to get data about the quality of teacher-made test.

### 4. *Instrument*

The instruments used to collect the data of this research were observation checklist, field-note, interview manual, students' answer sheets as the documents and the author themselves.

### 5. *Technique of Data analysis*

In analyzing the data of this research, the researcher applied an interactive model proposed by Miles & Huberman (1984). They were data collection, data reduction, data display, and drawing conclusion and verification. In addition, to analyze quantitative data, the researcher applied quantitative descriptive analysis using descriptive statistics.

## **D. Findings and Discussion**

The findings are presented based on the problems of the study which are explained in three points.

### 1. *The Quality of the Teacher-made Test*

The quality of teacher-made test administered in final examination at Gamaliel Elementary school was computed using ITEMAN program. Furthermore, the result of ITEMAN is spread out into several parts including difficulty index, discrimination index, distractor index, and the reliability of the test.

- a) Difficulty index of an item is identified by paying attention on prop. Correct columns and compare it with the criteria of difficulty level. The findings show that from 40 items, only 5% items are in difficult category, 27.5% item are in moderate category, and 67.5% items are in easy category. Considering this result, it can be concluded that the English test was an easy test. It is supported by the result of interview with the students. Most of them argue that the English test accomplished in final examination was easy for them. However, after interviewing the teacher, the teacher claimed that most of Gamaliel's students attend English course. Consequently, the test was claimed as an easy test because they had learned the material both in school and in English courses.
- b) Discrimination index of each test item can be determined based on point Biser value in the fifth column on the result of ITEMAN analysis and be interpreted based on some criteria. The findings shows that from 40 items, 82.5 % items have very good discrimination index and they need to be retained, 10 % items have good discrimination index, and only 7.5 % items that have marginal discrimination index and they need to be improved or revised. To sum up, 92.5% test items can discriminate students' ability.
- c) The quality of distractor index of each test item can be determined based on prop. endorsing and the second point Biser values on the result of ITEMAN analysis. The findings show that from 160 distracters, there are 34 or 21. 25% distracters do not

function well. Nine of them are in positive, while 25 others were not chosen by the students.

The reliability of the test was determined by looking at alpha value in the result of scale statistics on ITEMAN analysis result. Since, the test is teacher-made test, the coefficient of reliability used is 0.7. The statistic scale of ITEMAN result show that Alpha index of this test is 0.920 that is higher than 0.7. Compared to the reliability coefficient, the Alpha index is higher than reliability coefficient indicating that the test is reliable.

## 2. *The Washback of Test on Students' Motivation on Learning*

The result of this study shows that test brought two influences on students' motivation on learning. Test affected what students learn and how students learn. However, the findings show that test was not only brought positive Washback, but also brought negative Washback. It is positive Washback because facing examination motivated the students to work harder. In addition, the test result can be a reflection for the students in improving their ability. Those who get satisfied score will be more motivated to maintain their achievement. Meanwhile, those who get unsatisfied score will be more motivated to work harder and to get reward from school. On the other hand, test also brought negative Washback on students' motivation on learning because they will temporary stop their language course to focus on their test.

Considering the data above, the discussion which elaborated the research findings on research questions is presented. In line with the difficulty indexes of the test items, the teacher claimed that his made test had been constructed based on the indicators of learning and teaching material. Referring to the analysis result and students' opinion about difficulty level of the test, he stated that it is acceptable if the students claimed the test as an easy test because they attend English course. It is also proved by Budiyo (2015) who argues that the difficulty level of a test also depending on the group of testes being tested. If the group includes high competent students, the difficulty index tends to be higher which means that the test is an easy test.

By looking at the result of ITEMAN analysis, there are 33 items have discrimination index up to 0.30. Therefore, they are classified as good items. It relates to Budiyo's assumption (2015). He states that if a test item has discrimination index equals to or higher than 0.3 ( $D \geq 0.3$ ), it can be assumed that it has good discrimination. In addition, based on the criteria of developed by Ebel, those 33 items do not need any revision. Meanwhile, they should be retained. Furthermore, there are only three items; item number 20, 30, 33 included in marginal criterion. They possibly cannot distinguish students in high and low group. Therefore, the test-maker should revise those items or replace them with other items.

In preparing distractors in each item, the test maker should pay attention on the content and word classification. It means that the alternative options should be able to distract testes' attention from the correct answer. In line with distractor in an item, Budiyo (2015) points out that to distract the testes, the distractors in each item should have the same attractiveness as the correct option has. It means that the incorrect options should be in the same category with the correct option. After interviewing the English teacher, the teacher stated that he had constructed the distractors with the same category with the correct option. Furthermore, Budiyo (2015) argues that one of requirements to determine the distractors function is that the options should be chosen by at least 5% of testes. Additionally, if the point Biser of the distractor is in negative, it indicates that the distractor function well. However, if the point Biser is in positive, it means that the distractor does not function well. Related to Budiyo's opinion, 9 of distractors do not function well because they are in positive. Meanwhile, 25 others do not function well because none of students chose them as the option. Lastly, 78.75% distractors of English test items constructed by an English teacher of Gamaliel Elementary school had functioned well.

In ITEMAN program the reliability index is computed automatically using Alpha Cronbach's technique. In addition, the reliability of the test can be determined by looking at scale statistic in last part of ITEMAN analysis result. According to the information shown in the scale statistics in the result of ITEMAN analysis, the Alpha index is 0.920 which is higher than 0.7 (reliability coefficient for classroom exam). It means that the teacher-made test administered in final examination at Gamaliel Elementary school was reliable. In addition, Sukardi (2008) argues that the higher Alpha coefficient that the test has, the more consistent the test is. Dealing with this

assumption, it can be concluded that the teacher-made test of Gamaliel Elementary School administered in final examination is reliable. In the other words, the test has high consistency.

### E. Conclusion and Suggestion

Related to the findings of this research, the researcher draws several conclusions. Firstly, the teacher-made test administered in final examination overall can be classified as a good test. The test was analyzed based on several characteristics of good test. Secondly, the teacher-made test administered in final examination brought positive Washback on teacher's way of teaching including what teacher teaches and how teacher teaches. Lastly, the teacher made-test distributed in final examination brought both positive Washback and negative Washback on students' motivation on learning. It influences what students learn and how students learn.

The researcher proposed several suggestions for the teacher, students, school, and future researchers. Firstly, for the teacher, it is recommended to increase the awareness about the importance of conducting test analysis to know the quality of his made test. Secondly, for the students, they should not only study when they have examination, but also study when they do not have any examination. Since, learning is not talking about preparing self to face examination, but it is talking about how students get new knowledge. Furthermore, for the school, it recommended to instruct all teachers to analyze the quality of their test and the Washback of test itself in their way of teaching. In addition, school as the policy maker, needs to pay attention to the result of test analysis which can be a part of considerations in determining the next policy. Lastly, for the other researchers, considering the limitation of this study, the researcher proposed suggestion to conduct deeper study related to the Washback of test. The analysis of test Washback on material development and school policy can be the interesting problems for the next studies.

### F. References

- Ahmad, S. & Rao. (2012). A Review of the Pedagogical Implications of Examination Washback. *Research on Humanities and Social Sciences*, 2(7), 11- 20.
- Ahman, J.S. & Glock, M.D. (1971). *Evaluation Pupil Growth, Principle of Tests and Measurements*, 4th Ed. Boston: Allyn and Bacon. Inc.
- Allen, M.S. & Yen. (1979). *Introduction to Measurement Theory*. Monterey: brooks/Cole.s
- Arifin, Z. (2016). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya
- Arikunto, S. (1996). *Dasar-Dasar Evaluasi*. Jakarta: Bumi Aksara
- Basuki, I. & Heriyanto. (2014). *Asesmen Pembelajaran*. Bandung: PT Remaja Rosdakarya
- Brown, H.D. (2003). *Language Assessment: Principles and Classroom Practices*. California: Longman
- Budiyono. (2015). *Pengantar Penilaian Hasil Belajar*. Surakarta: UNS Press
- Cheng & Watanabe. (2004). *Washback in Language Testing*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Fulcher & Davidson. (2007). *Language Testing and Assessment*. New York: Roudledge
- Gronlund, N.E & Linn, R.L. (1990). *Measurement and Assessment in Teaching*, 6th Ed. New York: Macmillan Publishing Company
- Gronlund, N.E & Linn, R.L. (2000). *Measurement and Assessment in Teaching*, 8th Ed. New Jersey: prentice-Hall, Inc.
- Hadi, S. (2008). *Analisis Tes/Instrumen dengan ITEMAN*. Yogyakarta: UNY
- Hughes, A. (1989). *Testing for Language Teachers*. United Kingdom: Cambridge University Press.
- Malcolm, L.V.B. (2009). *Measurement and Statistics for Teachers*. New York: Roudledge
- McCowan, R.J & McCowan, S.C. (1999). *Item Analysis for Criterion-Referenced Test*. New York: SUNY
- Miles, M.B. & Huberman, A.M. (1984). *Qualitative Data Analysis*. Beverly Hills: SAGE Publications.
- Pan, Yi-Ching. (2009). A Review of Washback and Its Pedagogical Implications. *VNU Journal of Science, Foreign Languages*, Volume 25, 257-263
- Peraturan Pemerintah R.I. No. 19/2005 tentang Standar Nasional Pendidikan. Bandung: Fokusmedia
- Popham, W.J. (1995). *Classroom Assessment. What Teacher Need to Know*. Boston: Allyn and Bacon
- Purwanto. (2009). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar
- Lebagi, D., Sumardi, & Sudjoko. (2017). The Quality of Teacher-Made Test in EFL Classroom at the Elementary School and Its Washback in the Learning. *Journal of English Education*, 2(2, Dec.), 97-104.

- Puspitasari, A. (2015). *A Study on the Implementation of Attitude Competence of 2013 Curriculum in Teaching English*. Scholar Thesis at Sebelas Maret University. Surakarta: Unpublished
- Remmers, H.H., Gage, N.L., & Rummel, J.F. (1965). *A Practical Introduction to Measurement and Evaluation*. New York: Harper and Row Publishers
- Sukardi. (2008). *Evaluasi Pendidikan: Prinsip dan Operasionalnya*. Jakarta: PT Bumi Aksara
- Yusuf, A.M. (2015). *Asesmen dan Evaluasi Pendidikan*. Jakarta: Prenada Media Group
- Yin, R.K. (2003). *Case Study Research: Design and Methods*. London: SAGE Publication, Inc.