Application of Teachers' Knowledge of Socratic Questioning in Developing EFL Critical Thinking Skills among Omani Post-Basic Learners

Moosa Ahmed Hassan Bait Ali Sulaiman

Department of Education
Dhofar University, Salalah, Sultanate of Oman

Author: Moosa Ahmed Hassan Bait Ali Sulaiman

Thesis Title: Application of Teachers' Knowledge of Socratic Questioning in Developing EFL Critical Thinking Skills among Omani Post-Basic Learners

Institution: University Malaysia Sabah, Faculty Psychology and Education, Malaysia

Major: Education

Degree: Ph.D.

Year of award: September, 2019

Supervisor: Assoc. Prof. Dr. Suyansah Swanto & Assoc. Prof. Dr. Wardatul Akmam Din

KeyWords: Socratic questioning, students' critical thinking, teachers' knowledge, Oman

Abstract

With the premise that effective use of Socratic Questioning in instructional practices is of vital importance in EFL/ESL classrooms, this study was undertaken to investigate application of Omani EFL teachers' knowledge of Socratic Questioning (SQ) on Students' Critical thinking (CT) in post basic schools. This study is conducted in two phases of investigation. Phase I examined the correlation between teachers' knowledge and actual use of SQ. Phase II study investigated the application of Omani EFL teachers' knowledge of SQ on students' CT in post-basic schools. The present study made an attempt to (a) Determine the relationship between teachers' knowledge of SQ and their actual use of SQ. (b) Ascertain whether there is any statistically significant difference between mean scores of those who are taught through SQ and those who are taught CT skills in a normal setting. (c) Identify CT strategies, the students were able to develop and apply at the end of the intervention phase. A total of 230 EFL teachers, which comprises a 100% of the total population of all EFL post-basic female teachers in Dhofar region of Oman, participated in this study. The multi-method procedures and data analyses showed that (a) There is a strong positive relationship between teachers' perceived knowledge and their actual use of SQ; (b) There are significant differences between mean scores of those who were taught CT through SQ and those who were taught CT skills in a normal setting; (c) There is a clear evidence to suggest that students in the experimental group were able to develop effective CT strategies during the intervention phase. The findings of this study offer a number of implications: (a) For instructional practice that involves the teachers and students; (b) For policy and decision makers; and (c) For syllabus designers and testing and evaluation. In addition, it identifies and proposes certain areas related to pedagogy for future research.

Cite as: Sulaiman, M.A.H.B.A. (2019). Application of Teachers' Knowledge of Socratic Questioning in Developing EFL Critical Thinking Skills among Omani Post-Basic Learners. University Malaysia Sabah, Faculty Psychology and Education. Malaysia. (Ph.D.Thesis). Retrieved from *Arab World English Journal* (ID Number: 257 August 2020, 1- 187.

DOI: <https://dx.doi.org/10.24093/awej/th.257>

**APPLICATION OF TEACHERS' KNOWLEDGE OF
SOCRATIC QUESTIONING IN DEVELOPING EFL
CRITICAL THINKING SKILLS AMONG
OMANI POST-BASIC LEARNERS**

MOOSA AHMED HASSAN BAIT ALI SULAIMAN

**FACULTY PSYCHOLOGY AND EDUCATION
UNIVERSITY MALAYSIA SABAH
2019**

**APPLICATION OF TEACHERS' KNOWLEDGE OF
SOCRATIC QUESTIONING IN DEVELOPING EFL
CRITICAL THINKING SKILLS AMONG
OMANI POST-BASIC LEARNERS**

MOOSA AHMED HASSAN BAIT ALI SULAIMAN

**THESIS SUBMITTED IN FULFILLMENT OF
THE REQUIREMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

**FACULTY OF PSYCHOLOGY AND EDUCATION
UNIVERSITY MALAYSIA SABAH**

2019

DECLARATION

I hereby declare that the material in this thesis is entirely my own, except for quotations and summaries sources of which have been duly acknowledged.

09 Ogos 2019

Moosa Ahmed Hassan Bait Ali Sulaiman
DT1411007A

CERTIFICATION

NAME : **MOOSA AHMED HASSAN BAIT ALI SULAIMAN**
MATRIC NO : **DT1411007A**
TITLE : **APPLICATION OF TEACHERS' KNOWLEDGE OF
SOCRATIC QUESTIONING IN DEVELOPING EFL
CRITICAL THINKING SKILLS AMONG OMANI
POST-BASIC LEARNERS**
DEGREE : **DOCTOR OF PHILOSOPHY
(TESL)**
TARIKH VIVA : **12 SEPTEMBER 2019**

CERTIFIED BY ;

SUPERVISOR

1. Assoc. Prof. Dr. Suyansah Swanto

Signature

.....

CO-SUPERVISOR

2. Assoc. Prof. Dr. Wardatul Akmam Din

.....

ACKNOWLEDGEMENTS

Conducting this doctoral research and writing the dissertation has been one of the most enriching experiences of my life, which put me in touch with many learned scholars and perceptive people. Those who contributed significantly towards the success of this project include my esteemed supervisors, my dear family members, and my close friends and colleagues.

First, I would like to express my sincere gratitude for my research guides Assoc. Prof. Dr. Suyansah Swanto and Assoc. Prof. Dr. Wardatul Akmam Din. This work has taken a desired shape and improved immeasurably by their dedication, heightened intellectual capacity and personal care. It has been a stimulating experience to receive and share with them their inspiring enthusiasm and deep discussions in a series of long meetings for every aspect of the work. My supervisors have been a constant source of ideas, insights and encouragement, which always instilled confidence and motivation in me. I personally owe a lot to them. Then, I have to offer my undying gratitude for my family, for their endless help, blessings, prayers and wishes that inspired me during the difficult times. I really appreciate all of you more than words can express for compromising with my family commitments and presence with you on several occasions, and lending me emotional support which kept me focused and continue with the ongoing work. Next, I really appreciate a meaningful, support, and encouragement of some of my close colleagues and friends who have helped me, in different ways, to make this dissertation a better piece of work than it would otherwise have been. I am grateful to Dr. Khalid Musallem Al Mashikhi, Dr. Sobhy Sulaiman, Dr. Nasser Abdull Rashid, and Dr. Abdulkader Al Sayid. I really appreciate your support, assistance, and help. Your fellowship always provided me with good spirit and determination to endure this interesting work. I must single out two of my committed friends- Dr. Vijay Thakur and Hussain Ba'abood- who have always been and continue to be a continual source of inspiration, motivation, and all kinds of support. Without you this work wouldn't have been this work and I am in deed grateful to you for everything. This note of thanksgiving would be incomplete if I failed to appreciate a great help and cooperation that I received from the two universities. First, I would like to place on recorded my sincere gratitude to Universiti Malaysia Sabah for giving me the opportunity to do my PhD as a fulltime student and providing me with all kinds of academic support services. Second, I also, extend the same feeling of gratefulness to the Vice Chancellor Professor Hassan Said Kashoob, and the Deputy Vice Chancellor Professor Mohammed Limam of my parent university- Dhofar University. I am deeply indebted to them for their continuous help, support, and encouragement.

Lastly, I owe a great debt to all the scholars, researchers, and critiques in the field of my work whose contributions have lent me clear focus and direction in order to develop desired perspective of research in my dissertation. As it is not always possible to acknowledge every single person who has contributed to the success of this kind of voluminous work, I hope they will forgive me if they find an idea or expression or help borrowed without particular acknowledgement.

Moosa Ahmed Hassan Bait Ali Sulaiman
26 September 2019

ABSTRACT

With the premise that effective use of Socratic Questioning in instructional practices is of vital importance in EFL/ESL classrooms, this study was undertaken to investigate application of Omani EFL teachers' knowledge of Socratic Questioning (SQ) on Students' Critical thinking (CT) in post basic schools. This study is conducted in two phases of investigation. Phase I examined the correlation between teachers' knowledge and actual use of SQ. Phase II study investigated the application of Omani EFL teachers' knowledge of SQ on students' CT in post-basic schools. The present study made an attempt to (a) Determine the relationship between teachers' knowledge of SQ and their actual use of SQ. (b) Ascertain whether there is any statistically significant difference between mean scores of those who are taught through SQ and those who are taught CT skills in a normal setting. (c) Identify CT strategies, the students were able to develop and apply at the end of the intervention phase. A total of 230 EFL teachers, which comprises a 100% of the total population of all EFL post-basic female teachers in Dhofar region of Oman, participated in this study. The multi-method procedures and data analyses showed that (a) There is a strong positive relationship between teachers' perceived knowledge and their actual use of SQ; (b) There are significant differences between mean scores of those who were taught CT through SQ and those who were taught CT skills in a normal setting; (c) There is a clear evidence to suggest that students in the experimental group were able to develop effective CT strategies during the intervention phase. The findings of this study offer a number of implications: (a) For instructional practice that involves the teachers and students; (b) For policy and decision makers; and (c) For syllabus designers and testing and evaluation. In addition, it identifies and proposes certain areas related to pedagogy for future research.

Keywords : Socratic questioning, students' critical thinking, teachers' knowledge, Oman

ABSTRAK

APLIKASI PENGETAHUAN GURU DALAM PEMIKIRAN SOCRATIC DALAM MEMBENTUK KEMAHIRAN BERFIKIR ARAS TINGGI DALAM KALANGAN PELAJAR PASCA OMANI

Melalui premis penggunaan teknik pertanyaan Sokratis secara berkesan dalam pengajaran adalah penting dalam bilik darjah Bahasa Inggeris sebagai Bahasa Asing (EFL)/Bahasa Kedua (ESL), kajian ini meneliti pengaplikasian pengetahuan teknik menyoal Sokratis (SQ) dalam kalangan guru-guru EFL/ESL di Oman terhadap kemahiran berfikir secara kritis (CT) dalam kalangan pelajar di peringkat pendidikan post-basic. Terdapat dua tahap kajian. Tahap 1 meneliti korrelasi antara pengetahuan guru dan penggunaan sebenar SQ. Tahap 2 pula mengkaji pengaplikasian pengetahuan teknik menyoal Sokratis (SQ) bagi guru-guru ESL/EFL di Oman terhadap kemahiran berfikir secara kritis (CT) dalam kalangan pelajar di tahap pendidikan post-basic. Kajian ini bertujuan untuk a) menilai perhubungan antara pengetahuan SQ guru-guru dan penggunaan sebenar SQ. b) menentukan sama ada terdapat perbezaan yang ketara secara statistik antara skor min pelajar yang belajar melalui teknik SQ dan mereka yang mempelajari CT tanpa SQ. c) mengetahui strategi-strategi SQ yang membantu membina kemahiran CT pelajar. Responden kajian ini terdiri daripada sejumlah 230 guru EFL (mewakili kesemua guru-guru EFL Post-Basic di Dhofar, Oman). Kajian kualitatif dan kuantitatif, serta analisis data telah menunjukkan bahawa a) terdapat hubungan negatif yang ketara antara pengetahuan yang dirasakan ada (perceived knowledge) dengan pengetahuan sebenar (actual knowledge) mengenai SQ b) terdapat perbezaan ketara antara skor min mereka yang mempelajari CT menerusi teknik SQ dengan kumpulan pelajar yang sebaliknya c) terdapat bukti yang jelas menunjukkan bahawa pelajar dalam kumpulan eksperimen telah mempelajari strategi CT yang berkesan semasa fasa intervensi. Perbincangan dapatan kajian juga telah diteliti daripada aspek implikasi pedagogi di mana saranan-saranan yang dibuat adalah (a) terdapat hubungkait positif antara pengetahuan SQ guru-guru dan penggunaan SQ dalam kalangan guru-guru (b) terdapat bukti jelas bahawa SQ telah menambahbaik kemahiran CT pelajar dalam pembelajaran pemahaman c) pelajar menggunakan kemahiran CT mereka sendiri semasa pembelajaran pemahaman. Selain itu, kajian ini mengenalpasti dan mencadangkan beberapa topik berkaitan dengan pengajaran untuk kajian masa depan.

Kata kunci: teknik menyoal Sokratis, pemikiran kritis pelajar pengetahuan guru, Omani

TABLE OF CONTENTS

	Page
TITLE	i
DECLARATION	ii
CERTIFICATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
<i>ABSTRAK</i>	vi
TABLE OF CONTENTS	vii
LIST OF TABLES	x
LIST OF FIGURES	xiii
LIST OF ABBREVIATIONS	xiv
LIST OF APPENDICES	xv

CHAPTER 1 : INTRODUCTION

1.1	Overview	1
1.2	Background of the Study	2
1.3	Problem Statement	3
1.4	Purpose of the Study	9
1.5	Research Objectives	9
1.6	Research Questions	10
1.7	Scope and Limitation of the Study	10
1.8	Significance of the Study	11
1.9	Theoretical Framework	12
1.10	Operational Definition	10
1.11	Summary and Conclusion	17

CHAPTER 2 : LITERATURE REVIEW

2.1	Introduction	19
2.2	Theories and Conceptual Framework	20
2.3	Past Studies	22
2.4	Research Conceptual Framework	25
2.5	The Role of Display and Referential Questions in SLA	27
2.6	Critical Thinking	31
2.6.1	Critical Perspective on Critical Thinking	32
2.6.2	Philosophy-Based Theories and Definitions	33
2.6.3	Psychology-Based Theories and Definitions	35
2.6.4	Teaching for Critical Thinking	39
2.6.5	Assessment of Critical Thinking	43
2.7	Higher Order Thinking Strategies	46
2.7.1	Socratic Questioning	47
2.7.2	Previous Studies on Classroom Questions and CT Queries	49
2.7.3	The Role of Cognitive Questions	55
2.7.4	The Role of Affective Questions	58
2.8	Summary and Conclusion	59

CHAPTER 3 : RESEARCH METHODOLOGY

3.1	Introduction	63
3.2	Research Design	64
3.3	Population and Research Sample	65
3.4	Research Instrument	68
3.5	Research Ethic	69
3.6	Research Procedures	73
3.7	Pilot Study	73
3.7.1	Phase I: The Pilot Study, Instruments, and Procedural	74
3.7.2	Aspects	

3.7.3	Screening of Questionnaire	77
3.7.4	Missing Data	78
3.7.5	Normality Test	79
3.7.6	Outliers Assessment	82
3.7.7	Non-Responses Bias Test	82
3.7.8	Reliability and Validity of the Questionnaire	83
3.7.9	Construct Validity of the Questionnaire	89
3.7.10	Questionnaire Modifications for the Main Study	89
3.7.11	Classroom Observation Checklist	90
3.7.12	Reliability and Validity of the Classroom Observation Checklist	92
3.7.13	Preliminary Analysis of Correlation	92
3.7.14	Outliers Detection	94
3.7.15	Normality Test for the Main Study	96
3.7.16	Linearity and Homoscedasticity	97
3.8	Phase II: Data Collection, Instruments, and Procedural Aspects	97
3.8.1	The Watson-Galsser CT Appraisal Form A Test	97
3.8.2	Pedagogical Strengths of Watson-Glasser CT Test	99
3.8.3	Interviews and Quizzes	99
3.8.4	Statistical Analysis Procedures	101
3.8.5	Ethical Considerations of Research	103
3.9	Summary and Conclusion	103

CHAPTER 4 : RESEARCH FINDINGS

4.1	Introduction	105
4.2	Phase I: Analysis and Findings	105
4.2.1	Main Findings of RQ 1	109
4.3	Phase II: Analysis and Findings	111
4.3.1	Normality and Homogeneity Measures	111
4.3.2	Phase II: Analysis and Findings of RQ 2	115
4.3.3	Phase II: Analysis and Findings of RQ3	117
4.4	Summary and Conclusion	142

CHAPTER 5 : DISCUSSION & CONCLUSION

5.1	Introduction	144
5.2	The Relationship between Teachers' Knowledge and Actual Use of SQ	144
5.3	The Effect of SQ as a Teaching Strategy on Students' CT	149
5.4	The CT Skills developed from SQ Strategies as a Teaching Method	151
5.5	Dliimitations of the Study and Recommendations for Further Research	160
5.6	Implications of The Study	161
5.7	Summary and Conclusion	164

BIBLIOGRAPHY	167
---------------------	-----

APPENDICES	187
-------------------	-----

LIST OF TABLES

	Page
Table 3.1 : The names of schools and the volunteered teachers participated in the interviews	67
Table 3.2 : Examples of Socratic questions	70
Table 3.3 : The Summary of the eight weeks' intervention	72
Table 3.4 : The general daily lesson plan guide	73
Table 3.5 : Weighted mean values	76
Table 3.6 : The distribution of data in the independent variable	79
Table 3.7 : The saturation values for factor one; knowledge of probing and illustrating questions	84
Table 3.8 : The saturation values for factor two; knowledge of analytical questions	85
Table 3.9 : The saturation values for factor three; knowledge of recognition of assumptions	86
Table 3.10 : The saturation values for factor four; knowledge of deduction questions	87
Table 3.11 : The saturation values for factor five; knowledge of evaluative questions	88
Table 3.12 : The final structure of the questionnaire for the main study	89

Table 3.13	:	The test results of normality	96
Table 3.14	:	The quiz plan during the intervention stage	101
Table 3.15	:	The summary of data analysis procedures	102
Table 4.1	:	The teachers' responses to their knowledge in different domains of SQ questioning n=230	106
Table 4.2	:	The teachers' responses to their knowledge in different domains of SQ questioning n=33	107
Table 4.3	:	The levels of teachers' use of SQ questioning in different domains of SQ questioning (n=33)	108
Table 4.4	:	The correlations between the two variables of questionnaire & observation	111
Table 4.5	:	The significance of differences between the two groups in the external variables	113
Table 4.6	:	The significance of differences in CT test between the two groups in the pre-test	114
Table 4.7	:	The significance of differences in CT test between the two groups in the post-test	115
Table 4.8	:	The mean scores of CT skills in four quizzes (n=30)	119
Table 4.9	:	Students' results of the four quizzes in each skill (n=5)	120
Table 4.10	:	Students' quiz scores in inference skill (n=5)	121
Table 4.11	:	Students' quiz scores in recognition of assumptions skill (n=5)	122
Table 4.12	:	Students' quiz scores in deduction skill (n=5)	123
Table 4.13	:	Students' quiz scores in interpretation skill (n=5)	123

Table 4.14	: Students' quiz scores in evaluation of arguments skill (n=5)	124
Table 4.15	: Students' CT responses to the interview questions (n=5) in quiz 1	126
Table 4.16	: Students' CT responses to the interview question in quiz 4 (n=5)	128
Table 4.17	: Students' responses to inference skill in quiz 1 & 4	134
Table 4.18	: Students' responses to recognition of assumptions skill in quiz 1 & 4	135
Table 4.19	: Students' responses to deduction skill in quiz 1 & 4	137
Table 4.20	: Students' responses to interpretation skill in quiz 1 & 4	138
Table 4.21	: Students' responses to evaluation of arguments skill in quiz 1 & 4	139

LIST OF FIGURES

	Page
Figure 1.1 : Theoretical framework	15
Figure 2.1 : The Socratic questioning model: Moor and Rudd (2002)	43
Figure 3.1 : Normality distribution of the sample	78
Figure 3.2 : Outliers before detection I	80
Figure 3.3 : Distribution of the outliers before detection II	80
Figure 3.4 : Outliers after detection	81
Figure 3.5 : Distribution of the outliers after detection	81
Figure 3.6 : Extreme values and outliers in the questionnaire	93
Figure 3.7 : Extreme values and outliers in the observation	93
Figure 3.8 : Sample distribution of the independent variable	95
Figure 3.9 : Sample distribution of the dependent variable	95
Figure 3.10 : Linearity and homoscedasticity of independent and dependent variables	96

LIST OF ABBREVIATIONS

SQ	Socratic questioning
CT	Critical thinking
ELAS	English language achievement score
TGPA	Total grade point average
CTSS	Critical thinking skills score
IQS	Intelligence quotient score
KMO	The Kaiser-Meyer-Olkin Measure Sampling Adequacy

LIST OF APPENDICES

	page
Appendix A: Permission from Ministry of Education to conduct the current study	187
Appendix B : Referees' further information and background	189
Appendix C : Intervention program phase II	190
Appendix D : Pre-study questionnaire phase I	211
Appendix E : Rotated component matrix table phase I	217
Appendix F : Main study questionnaire phase I	219
Appendix G : Pre-study classroom observation-checklist phase I	225
Appendix H : Main study classroom observation-checklist phase I	230
Appendix I : Quizzes phase II	235
Appendix J : Interviews phase II	262

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

The concept of “Critical Thinking” (CT) is commonly indicated and used in educational procedures and strategies, especially in curriculum design and instruction goals. Questions reflect the plan of our thinking. They define what information we look for. Most studies seem to agree that educationists considered CT skills as a significant factor that leads to effective thinking and critical questioning skills. For example, Angelo and Cross (1993) believe that when teachers construct their teaching goals they should be able to create a learning context which urges students to get involved in CT and evaluate teaching materials by collecting and analyzing evidence. Lombard and Grosser (2004) assert that the main goal of learning is to teach learners different CT skills. A central issue in CT process is that teachers should have the knowledge and the ability to use questioning skills that would help students to construct the target skills needed to think critically and analytically. Although considerable research has been devoted to the area of CT, rather less attention has been paid to train teachers to apply and use different question types that activate students CT skills and higher cognitive process Carl (2010).

More recently in Oman, there has been a growing concern that CT, particularly related to teaching and learning process, may not be implemented and applied purposefully and effectively by teachers (Ministry of Education, Oman, 2008). It seems that there is a mismatch between what Omani teachers know about questioning skills and how they apply it while teaching. Moosa, Suyansah and Wardatul Akmam (2015) found that there is a negative relationship between teachers’ knowledge and actual use of question types. Despite this

occurrence, precise comprehension between teachers of what CT really means is required.

The main aim of this study is to investigate more complicated variables like teachers' knowledge of Socratic Questioning (SQ) and actual use of (SQ) in relation to students' CT skills that might be considered as strong variables in teaching EFL. Also, the current study implemented an intervention program that was based on Socratic questioning as a teaching strategy and examine the result of teachers' knowledge of (SQ) and use of (SQ) through students CT skills. The large scale of sample and the advanced statistical procedures could reflect conclusions that might contribute to the literature of teaching CT.

1.2 Background of the Study

The Sultanate of Oman is facing the challenge of educating its youth for life and work in the new priorities made by the current global economy. These priorities need a high degree of adaptability and a concrete background in school education particularly in the English language, in order to deal with the modern economy, and capture international business opportunities. Consistent with this understanding, the Ministry of Education of the Sultanate of Oman has specified that in the Omani education system Omani students acquire knowledge and skills in all areas of curriculum including skills in questioning, investigating, critical thinking, problem solving, and decision making (Ministry of Education, Oman, 2008). In addition, the following specific oral language objectives for higher classes are considered as fundamental productive skills that are related to CT:

1. To initiate and participate in longer conversations and interactions.
2. To recognize and produce common idiomatic and conversational expressions.
3. To use English to carry out practical transactions in everyday life, using a largely predictable and restricted set of language and functions.
4. To use English for social communication.

Although it is stated in the national curriculum specifications that “the new English Language curriculum is being designed to equip learners with higher cognitive abilities and skills, and attitudes that Omani learners will need to succeed in this rapidly changing society” (Ministry of Education, 2008, p.7). Teachers’ procedures and techniques of questioning may prevent learners from attaining this objective. This is based on the understanding that teacher’s knowledge and actual use of questions could influence the way they apply national education programs in their teaching, which in turn may affect the use of classroom questions. This will therefore hamper the nation’s effort in achieving its objectives as stated in its educational system. Furthermore, the studies conducted on the status of the pedagogy of CT skills in Omani post-basic schools reflect an unfavorable scenario. The study on post-basic EFL teachers’ use of behaviors nurturing CT skills conducted by Al Kindi and Al Mekhlafi (2017) on 12 post-basic schools of Oman reports that teachers working in those schools rarely use the behavior that nurture CT skills.

Similarly, a study conducted on the first year students of a public university in Oman by Al Mahrooqi, Denman, and Al Aghbari (2018) on whether CT skills were adequately developed in the schools they studied. The findings of this study revealed that the CT skills were not developed adequately and the researchers of this study strongly argued in favor of a better integration of CT in the curriculum in the educational system of Omani schools in order to adequately prepare learners for the demands of the university studies and the workforce. Thus, it is in this background, discussed above, that the present study was conceptualized and initiated as a Ph. D. project. What follows next in the sequence is the discussion of problem statement.

1.3 Problem Statement

A central argument developed through research on CT is a means to transform learning and society. Social practice is one of the indispensable components of CT (Benesch, 1993; Atkinson, 1997; Oster, 1989; Brookfield, 1987; Shor & Freire, 1987; Fox, 1994; etc.) that caught the researcher’s attention as a language teacher and led him to explore its relevance in English

language teaching. Upon further exploration, Dewey (1933) remarks that teaching and thinking skills are the essential goal of education. To Lombard and Grosser (2004, p. 2), "the main goal of education is to train learners on critical thinking." Despite the recommendations of developing learners' CT, researchers have observed that most educational institutions are neither challenging learners to think critically about various subjects nor encouraging them to acquire the reasoning abilities that are required to deal with the difficulties of modern life effectively (Parvize and Marzieh, 2010; Paul and Elder, 2006; Nunan, 2003; and Kip, 2002).

In this regard, Kip (2002) has summarized a number of significant reasons as to why further research in language teaching in the context of CT is vital. First, modern communities are challenged with difficulties partly made by scientific development that will not be resolved through additional scientific gains. Biological destruction, local wars, cultural and religious clashes, are complicated issues requiring logical thinking and decisions. Therefore, English language provides a real communication between nations and individuals that might help in making balanced choices and common agreements. Second, increasing learners' knowledge and understanding through English language materials is a necessity for understanding current situations in the world today. Third, reasoning and thinking are required to have a wider understanding of social context activities and the degree to which they are the same or different from one area to another.

In the social context of language use, Halliday (1999, cited in Bardhan 2013, p. 94) rightly argues that language, as a semogenic system, is a resource for meaning making and the meaning is understood in functional terms, i.e. in relation to the social contexts in which it has evolved with the human species. By logical implication, therefore, language teaching and learning needs to be interactive for the reason that, as Long (1996) argues, interaction and negotiation of meaning facilitate acquisition because they involve input modification, internal learner capacities, and output in productive ways.

However, the vital question arises, and remains to be answered, is that to what extent are language learning activities in school curricula genuinely authentic and interactive? Close examination of pedagogical environment in EFL schools suggests that it is predominantly dominated by outcome-bound approaches and textbooks-based instructional practices which involve language activities, tasks, and tests that mostly require one right answer or response.

Pedagogical practices led by such an approach limit students' ability to be original and skeptical in reflecting upon various issues based on their own thinking and experiences. As Sivasubramaniam (2015) and Nunn and Sivasubramaniam (2011) strongly argue, such a focus has entirely centered on bureaucratic efficiency aimed at having a uniform curriculum for the majority of the students and a scheme of teaching and evaluation based on recalls, cloze texts, and multiple-choice questions, etc. in standardized texts. Such an approach leads students towards the practice of one right answer.

Thus, however, in much needed socially-oriented view of competence, spontaneity, flexibility, and diversity accrues only through a process-centered pedagogy of voice, agency and response (these are further clarified in section 1.9) which seems to be seriously lacking in the curriculum of EFL schools. This problem calls for an in-depth research-based investigation and exploration of scientific pedagogical solutions.

As Carter (1997) argues for a strong stance against a unilateral view of one right reading is a non-negotiable precondition for democratizing reading and freeing our students' emotional drives and democratic potentialities as a way of assigning renewed and fresh relevance to reading in an ideological world.

This suggests, in line with Mackenzie (2002, p. 47), that the meanings students propose and construct signpost the inferential processes of recovering propositional attitudes and explicatures; deducing implicatures; attempting figurative interpretations and creating a

context (including a social context) in the attempt to make a text optimally relevant.

Conversely, if the students are frequently asked to memorize the text mechanically, their memorization will not constitute knowledge and such a reading of the text will neither result in real reading nor in the knowledge of the objective realities which the text refers to (Freire and Macedo, 1987; McCormick, 1994). In spite of these important facts that are crucial for developing English language and much needed critical competence, there are limited studies that investigate students' CT in relation to second language learning in the Middle East, particularly in Oman.

According to Kuhn (2005) and Rebecca, Berkes, Mosley, Hui, and Joseph (2005), there is an inappropriate daily justification, reasoning and poor argumentation abilities applied by different individuals. Even higher education system seems to have inadequate effect on college students' CT abilities, including, providing logical interpretations of reading texts and formulating balanced and well-reasoned opinions (Evans, Harkins, Mary, Young, and Jeffrey, 2008).

The California Commission on Teacher Credentialing conducts a research project on university lecturers and professors. They conclude that in spite of a wide agreements, CT was an essential goal of their teaching (89%), only few (19%) were able to clarify what they meant by CT, and very limited participants (9%) actually taught CT in a normal teaching classroom (Paul and Elder, 2006). According to Kenzic (2009), instructors state that they implemented discussion 53.6% of the time and SQ 17.6%. However, he also concludes that discussion appeared only 19.5% of the time and Socratic questioning only 2.2%. These results show that while concern about CT is commonly shared, real teaching for CT is not implemented on a broad scale.

Another factor that contributes to CT challenges is materials of teaching English as a second language. They offer various opportunities to improve abilities and practices needed for developing students' CT. However, English language teachers who participated in a study by Tomlinson (2005) were unsuccessful to challenge learners explicitly to enhance reasoning skills. Training in the reasoning tasks and synthesizing ideas may support students to construct skills and strategies needed in their daily life situations. In many English language courses, teaching materials, textbooks, tasks and activities are used inappropriately.

Teachers tend to focus on lower cognitive skills and teach in traditional methods using top-down techniques where students tend to be only receivers of knowledge. As a result, students will memorize to pass the exam rather than developing higher cognitive skills that will help them become good thinkers (Moosa, Suyansah, and Wardatul Akmam, 2015). This is because teachers may not have sufficient knowledge of higher cognitive questions that promotes CT skills. Shulman (1987) classifies teacher knowledge into 7 categories. The most related one to this study is the Knowledge of teaching methods. It is very important that teachers should have the ability to use different teaching methods and decide when and how to use each method. Therefore, teacher's knowledge of SQ is considered as an essential method that reflects critical thinking skills. It is found that 90 % of the Omani teachers who participated in a case-study claimed that they had the knowledge of different types of classroom questions (Moosa, Suyansah, and Wardatul Akmam, 2015). However, the findings of the study indicate that only 5 % were able to apply different types of classroom questions.

EFL teachers should know how to deal with different challenges such as developing their language abilities, teaching within global context, applying classroom action research, ability to have an access to professional development activities and education technology practices (González, 2000). Such abilities challenge those teachers to develop language knowledge, pedagogical applications, different theories of learning and teaching contexts.

In the Omani context, to use Al-Ani's (2016) argument, Oman is still grappling with entrenched, traditional methods of education, which are incapable of meeting diverse learning needs of producing young people who are equipped to readily take their places in the workforce.

Critical thinking, considered by educationists to be one of the most essential skills in the classroom across all levels, is actually being included, taught, and evaluated in Omani schools remains largely unexplored. Much of the research available on this subject in Oman indicates that critical thinking in the English language classrooms remains a distant outcome in middle and high school teaching and learning (Al-Busaidi, 1995; Al-Issa, 2010; Al-Issa and Al-Bulushi, 2012; Al-Mahrooqi, 2012; Al-Gardani, 2012; Samarsinghe, 2014).

Omani pre-service EFL teacher education programs should focus on what constitutes the required professional teaching knowledge so that EFL teachers can be critical practitioners and effective language teachers. As Shulman (1986) states, EFL teacher education programs should focus on and include the required knowledge that will help EFL teachers to become more effective and critical.

On a separate note, spurred by the findings and recommendations of the international conference on secondary education reform, Omani educational institutions continue to graduate learners who do not have reasoning skills or CT abilities. There is a call for a "significant increase in the amount of post-basic education diploma graduates who have advanced communication skills, self-learning abilities, critical thinkers and deal with modern sciences and technology issues" (Ministry of Education, Oman, 2008, p. 3).

In view of the importance of CT and the pedagogical gaps and discrepancies involved in the teaching and learning of CT skills, as realized from the discussion of above studies and further examination of reasonably comprehensive literature in the field suggest that though researchers investigated different variables related to students' CT, they were able to measure only limited variables and could not take into account issues such as Teacher Knowledge and

use of Socratic Questioning (Oman Al-Kindi and Al-Mekhlafi, 2017).

Furthermore, most of the studies have developed a macro level perspective of research (Oman Al-Kindi and Al-Mekhlafi, 2017). They also state that there is hardly any study conducted in Oman which focuses on CT skills at the micro level. Given the importance of CT skills in the intellectual growth and development of the students as well as in enhancing their employment prospects in general, aims to identify the gaps that lie in the existing research on CT skills in general and in Oman in particular. It also attempts to broaden the scope of research on the students' knowledge and use of CT skills both at the macro and micro level. Incidentally, macro level CT skills focus on the main CT skill as a product, e. g. inference and deduction whereas micro level CT skills, by targeting the strategies that help achieve the main CT skill as a product.

Strongly motivated by the gravity of the pedagogical scenario of CT skills in Oman, the researcher is keen to find out the application of teachers' perceived knowledge and actual use of Socratic questioning on students' CT skills. The study aims to investigate the relationship between teachers' knowledge of SQ and teachers' use of SQ. In addition, the study intends to examine the application of SQ strategies on students' CT skills. Furthermore, this study attempts to add to the body of knowledge on how students learn English language through different tasks that involve higher cognitive questions, and how English language courses can be used to build up students' CT abilities by using SQ based on the taxonomy of Paul and Elder's CT model (2006).

1.4 Purpose of the Study

Driven by the gaps identified in the literature relevant to English language pedagogy in schools in the ESL/EFL contexts in general and in the Omani context in particular, the main aim determined in this study is to investigate the application of teachers' knowledge of SQ on grade 12 students' CT in post-basic schools in Dhofar region in the academic year 2016-17. Though, the new educational reform in Oman is heavily student-centered and based on communicative

language teaching approach, it has changed in theory; however, the situation on the ground has been largely otherwise in terms of practice. The disparity between theory and practice still exists and persists, as argued by Al-Issa and Al-Bulushi (2012, p.170). Therefore, it requires teachers with real and hands-on understanding of different methodological and pedagogical aspects and considerations of questioning strategies and its application that would help to enhance students' language abilities and CT skills.

1.5 Research Objectives

The main objective of the current study is to examine the result of applying teachers' knowledge of SQ on students' CT skills. In so doing, the study seeks:

1. To determine the relationship between teachers' knowledge of SQ and actual use of SQ.
2. To ascertain whether there is any statistically significant difference between mean scores of those who are taught using SQ and those taught using CT skills in a normal setting.
3. To identify CT strategies, the students were able to develop and apply at the end of the intervention phase.

1.6 Research Questions

The current study is based on two phases of investigation. Phase I is descriptive and correlative in that it investigates the relationship between variables like teachers' knowledge of SQ, and actual use of SQ. While phase II is experimental and the researcher proposed a null hypothesis regarding the statistical significant differences between the two groups. It aims to investigate the application of teachers' knowledge of SQ on students' CT. Thus, the researcher used the two-phased research design to answer the following questions:

1. What is the relationship between teachers' knowledge of SQ and their actual use of SQ?
2. Are there any statistically significant differences between mean scores of students

who are taught through SQ and mean scores obtained by students who are taught CT skills in a normal setting?

3. What CT strategies the students were able to develop and apply at the end of the intervention phase?

1.7 Scope and Limitation of the Study

The purpose of this study is to investigate the application teachers' knowledge of SQ on students' CT in Omani post-basic schools (grade 12) during school year 2016-2017. This study was conducted in two phases. Phase I aims to find the relationship between teachers' knowledge of SQ and actual use of Socratic questioning. Phase II implements eight weeks teaching program based on Paul and Elder's model (1996) that can be considered as teachers' guide of using Socratic questioning strategies to develop students CT skills in English language classes. The aspects that looked into were the teachers' knowledge of SQ, use of SQ, and students CT skills.

However, the limitations of this study can be presented as the following: a) this study covers only the EFL teachers who teach Post-Basic Certificate level in Omani schools. This is generally based on the supposition that the question types that are examined in this study would be clearer and practically viable at this level of instructional practices, and that presumably students would use the English language better than those in the lower classes. Therefore, the findings of the study would be relevant only to EFL teachers who teach students in Post-Basic Certificate level only; b) the target population of this study is limited only to female students in the target schools. This is because students gender is not mixed in higher classes according to the education policy in Oman. Even male teachers are not allowed to teach in girls' higher schools— only female teachers can teach in these schools; c) the majority of English teachers are Omani female teachers, which is due to the Omanization policy in the country, and also teaching English is not a preferred job for Omani male graduates. Therefore, English language teaching in male schools is taught by a large section of non-Omani male teachers. As

a result, the findings of the study would be relevant only to EFL female teachers who teach female students in post-basic level only.

1.8 Significance of the Study

This study is significant in three contexts: theoretical context, research context, and practical context. First, from the theoretical perspective, this study augments the existing literature on Socratic Questioning and cognitive questions in relation to critical thinking. This would act as a point of reference for future research works. The study adds to the available literature and tries to fill the research gaps identified in the areas of education in general and in TEFL in particular. The current research design is descriptive and experimental; which produced quantitative and qualitative data sets. Therefore, the method used in this research was a multi-method approach (M. Teresa, Angel, Jose, Pedro, and Anthony, 2018). Both quantitative and qualitative data were used and triangulated to the relevance and purpose of specific research questions. From the research perspective, both quantitative and qualitative data used for the study provide credence to the notion of using the right and appropriate methods for various levels of analysis.

Additionally, this study involves the adoption and development of a questionnaire that measured teachers' knowledge of Socratic Questioning. It also includes, a classroom checklist that measured the actual use of Socratic Questioning in relation to students' CT. Furthermore, this study adopted Paul and Elder's CT model (1996) to examine the result of the intervention program on students' CT skills. This is expected to contribute to the existing body of knowledge and scholarship through the adopted conceptual model. The practical significance of the results of the study stems from the understanding that Socratic Questioning and cognitive questions enabled teachers to examine their way of teaching that allows or restricts the development of critical thinking in learners. As such, they would be better prepared to meet the needs of learners who are capable of thinking critically in various situations. Therefore, the study claimed to be useful for the Sultanate of Oman in its initiatives to have an education system that produces well-rounded learners with CT skills such as investigating, critical thinking, problem solving, and decision making. In this

way, the study proposed solutions and alternatives to Omani education system to assist in producing better graduates who can cope with the labor market requirements and work efficiently.

In addition to the practical implications towards the individual teachers and learners, nation (the Sultanate of Oman) and organizations, theoretical contributions towards literature are important part of the impact that the study makes. Of greater relevance to this thesis is the fact that this study conceptualises a model that investigates deeper into the approaches of Socratic Questioning and cognitive questions for critical thinking, particularly in the Sultanate of Oman. This is practically the same as drawing the attention of the syllabus designers to the importance of Socratic Questioning that may be found to be neglected in the current textbooks; raising the teachers' awareness about the importance of Socratic Questioning and different question types; suggesting support systems (such as training courses / supervision practices) for teachers to enable them to use SQ with other types of question in appropriate ways; and above all, helping teachers to identify the most appropriate ways of using questions.

1.9 Theoretical Framework

This study is conceptualized and framed in the educational philosophy and principles of 'Deep Thinking' via 'Socratic Questioning' (Keng, 1996; Kenzik, Wubbels & Hajer (2010); of 'Engagement with the Text' via 'Reader Response Theory' (Rosenblatt, 1978; Rosenblatt, 1995); of 'Making Sense of Experiences by Constructing one's own Mental Models' via 'Constructivist Approach to Language Competence' (Piaget, 1966, 1972, 1974; Vygotsky, 1962; Burden 1997; Seifert, 1997; Reagan, 1999; Osborn 2000; Pelech and Pieper, 2010; Onuf, 2013); and Paul and Elder's (2006) 'CT Skills Model to Study the Knowledge and Practice of Socratic Questioning in the ESL/EFL Context'. The selection of principles underlying the construction and implementation of the current study is mainly based on the foundations of Social Constructivist Theoretical Framework. Social Constructivism as a current theory of cognitive construction has its grounding in Jean Piaget's theory of constructivist learning (Piaget, 1974). It proposed that

from early stages, individuals are engaged in building their own individual understanding of the world. This learning appears through an engagement of personal meaning-making of their involvement in the surrounding environment. While supporters of modern social constructivism still support this fundamental view of Piagetian constructivism, their aim is to put less importance on the cognitive practices taking place within the learners' mind and focus instead on the function that interpersonal relations take part in the construction of personal meaning.

A central component of a social constructive approach of Vygotsky (1962) applied to this study is that learning is considered as a dynamic process (Sutherland, 1992) in which learners are not passive receivers of knowledge but are dynamically concerned in its construction by involving external input in acts that are limited to them in making their own personal understanding of the social practices that form the ground of their learning. In addition, the received input is directly linked with the idea of personal meaning-making is the hypothesis based on that every individual is a unique learner with a unique background and history and has different personal needs, and that a learner's uniqueness is essential to construct knowledge that is personal and internal to the learner: As Williams and Burden (1999, p. 193) pointed out "different learners will have different comprehension and they construct their own understanding that is personal to them". Learner uniqueness consequently needs to be supported and encouraged by learning context that is as personal to individual's experience as possible. It is because of the deeply involved personal ability of learning, and the strong mediating function of contextual aspects on the practices of personal meaning-making that teachers are seen as supporters, aiming to support the learner to conclude his or her own thoughts and comprehension of the content (Bauersfeld, 1995). Eventually, according to Glasersfeld (1987), the accountability for learning exists increasingly with the learner as teachers principally guide, facilitate and support, appropriate learning environments. Figure 1.1 below shows the theoretical relationship between teachers' knowledge of SQ and actual use of SQ an effective teaching qualities that would enhance students' CT. Figure 1.1, consists of 7 squares, which reflect what the current study proposes to look at the following three variables that are examined in the current study; teachers' knowledge of SQ as an independent variable,

teachers' actual use of SQ and students' critical thinking both dependent variables in this study.

Social Constructivism is basically a paradigm for thinking development because it is simply adjustable to conceptualizing the function of question types in relation to CT in learning context (Sivan, 1986). One of the features of such a paradigm that makes it mostly appropriate as a framework is that it provides discussion opportunities for individuals to construct their thoughts and that basically will reflect on their language learning. From a constructivist point of view, CT is considered as dependent upon context, but the function of context is reflected beyond that of simply concluding how CT appears (Sivan, 1986). Social Constructivism considers CT as a mediating role of interactions with individuals in the learners' context (especially language teachers that, who act as learning supporters by becoming CT helpers through different types of questions) and this is based in the center of the proposed theoretical model of this study. It also, emphasizes the value of the learner's own voice and agency, specially the fact that, learners are not depending on external conditions that construct their CT skills through procedures over which they have no control. Voice and agency, according to literature refers to independent thinking and its application. According to Crookes and Lehner (1998), voice and agency are joint goals relevant to the simultaneous development of English communicative abilities together with the ability to apply them to developing critical awareness of the world and the ability to act on it to improve matters (p. 320).

Figure 1.1: Theoretical Framework

As dynamic co-constructors of their beliefs and attitudes, learners always keep a certain amount of options on how these come to influence their CT abilities and their learning behavior. Thus, from a social constructivist point of view, enquiry-based learning indicates that every learner thinks differently. As Williams and Burden (1997, p.197) asserted, Individuals will make their own understanding of the different external conditions which are considered areas that are personal to them, and they will proceed on their personal nature and apply their personal understanding in different ways. As a result, what makes individuals to learn a foreign language

and keeps that individual learning until he or she has gained a level of proficiency with which he or she is satisfied will be different from one to another.

To sum up, a Social Constructivist vision can involve a helpful framework for comprehending the CT implications of self-concept on language learning. Social constructivists consider learners' conception of themselves at the core of the learning situation because these conceptions significantly affect the perceptions of practice and therefore play an essential role in the way learners build up new knowledge (Williams and Burden, 1999). Seifert (1997) considers classroom setting as an opportunity where students can construct their language learning by testing their understanding through classroom tasks and learning situations. In this context, the possible implications of aspects of one's self-concept are obvious: how they consider themselves is possibly to affect not only their motives for learning a language (and therefore the aims that direct their learning), but also what aspects of their environment encourage or obstruct their thinking abilities, which in turn representing a major determinant in the alternatives they choose in relation to their learning.

1.10 Operational Definitions of Key Terms

The following main terms are used in this study. Those terms are fully defined in the literature review. However, below is a brief operational definition of each one.

1. **Socratic questioning** is the type of questions that enhance thinking process through discussion of thoughts, ideas, examples, and views to develop the dialogue to a maximum advantage. SQ could activate and evoke students' knowledge and lead the discussion to the target knowledge.
2. **Teachers' actual use of SQ** refers to the real application of teachers' knowledge of SQ. It will be obvious during classroom observations if the teachers are able to apply SQ. Teachers show their ability to ask questions that reflect deep thoughts, clarifications, and reasoning that help students to respond more logically and use their analyses, interpretation, justifications and conclusions that reflect their comprehension responses in a logical act.

3. **Teachers' knowledge of SQ** refers to teachers' understanding that reflects deep considerations, explanations, and cognition that help students to think more analytically in order to build comprehensible justifications and conclusions. Teachers' knowledge in this study refers to the knowledge of Socratic questioning that they have or they will construct through the proposed training program in which teachers might be able to apply SQ in order to develop Omani students CT skills.
4. **Students' critical thinking skills** refer to those skills that students need in their daily life situations either in school or in real life environment. They need skills like decision making, infer ideas, justify actions and thoughts, and, evaluate claims and conclusions. Independent learners can be self-regulators where they can use skills like problem solving, researching target knowledge, making their decisions, looking for other alternatives, and think logically.

1.11 Summary and Conclusion

This introductory chapter was devoted to the description, discussion, and explanation of the background and rational of the study, the problem statement and related literature review, the purpose of the study, research objectives and research questions, research methodology, the scope and significance of the study, and the theoretical foundations used for the study. This study is based on the premise that develops from the findings of the studies conducted by Allen, Butler-Mader, and Smith (2010); Cheng, Cheng, and Tang, (2010); and Wideen, Mayer-Smith, and Moon (1998) who suggest that there is a disparity between the theory taught to the students in teacher education programs and the subsequent practice of these student teachers in the classrooms. What has become clear from these studies is that the idea of simply transmitting important pedagogical knowledge to teachers, hoping that they will apply this knowledge in their practices, does not really work.

Led by the premise the researcher undertook this study to investigate the actual result of the application of teachers' knowledge of SQ on students' CT in grade 12 of post-basic

Certificate in the Omani public schools located in Dhofar region. The researcher developed three research objectives from which the three research questions were developed focusing on; (a) the relationship between teachers' knowledge of SQ and their actual use, (b) statistically significant differences between the mean scores of the experimental and control groups, and (c) the SQ strategies that helped the students to develop CT skills.

The study was divided into two phases. Phase I focused on collecting data to measure teachers' knowledge of SQ through a questionnaire and teachers' actual use of SQ through observation checklist. Phase II was based on investigating the application of teachers' knowledge of SQ on CT students' skills in reading classes via pre- and post-tests in order to measure the actual result of the eight-week long intervention program.

This study is significant in terms of its theoretical as well as practical perspectives. The findings of both quantitative and qualitative data have its useful implications in advancing and/or adding the knowledge and scholarship to the existing literature and also, in providing with practical insights to improve the teaching and learning situations in the Omani schools and also, the schools in similar context beyond Oman. Having discussed the introductory information relevant to this study, the next chapter is devoted to the discussion of the critical review of literature related to the influence, focus and direction in the current study.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

It is important to lay the structure of the theoretical body of this study by defining and reflecting upon the context of the main concepts which must be obvious in this study before discussing issues related to the influence, focus and direction in the current study. These concepts are introduced mainly in terms of themes; therefore, this review will not discuss the concepts in isolation. The researcher reviewed the related literature by concentrating on key concepts that reflect the nature of research questions in the current study. The researcher organized these concepts around themes because the components and the nature of the concepts are interrelated and present the effect of teachers' knowledge and use of Socratic Questioning on EFL students' CT.

Many researchers found that there is lack of consensus on how to define CT, on what kind of skills are required to teach CT, and constructing the best procedures and strategies for this teaching. Generally, even now educational bodies and systems have not reached common implementations to teach CT. This chapter reflects on these different views in order to construct a clear understanding of different aspects related to CT.

There are three major sections in this chapter on literature review. The first section covers the literature available on the teachers' knowledge of questioning strategies in relation to students' CT. The second section discusses the literature available on different concepts related to CT, like philosophy-based theories and definitions, psychology-based theories and definitions, teaching CT, and assessment of CT. The third section discusses the literature on concepts that are related to students' higher order thinking strategies, like Socratic Questioning,

inquiry learning, and cognitive and affective questions. What follows next is a review of literature relevant to the teachers' knowledge of questioning strategies.

2.2 Teachers' Knowledge of Questioning Strategies

The recent worldwide movement in the reform of education concentrates on many issues, such as, qualities, principles, and teacher training. Teachers are the keystone of communities and their development is one of the more significant features of community development as they are entrusted with the implementation of the reform strategies, and the overall education of our children. Recently, the Educational Reform Project in Oman emphasizes on students' attainment of learning outcomes and encourages students for lifelong learning. However, even the good imagined framework applied in the most perfect conditions, using the most updated system will not ensure effective implementation (Eggen and Kauchak, 2002). Successful and effective implementation is primarily based on teachers and their effective roles in the classroom. Though, there are many studies on effective teaching, there are few that focus on teachers' knowledge of SQ that are related to students' critical thinking which is important to achieve learning outcomes. This deficit of research on Teachers' knowledge of SQ linked with students' CT provides a strong rational and need for this study.

Teachers' knowledge is the most recent pedagogical issue under discussion in education circles. It seems almost everyone has an opinion on what constitutes and construct a teacher's knowledge, how teacher's knowledge can be developed and extended, and how much impact teacher's knowledge has on student learning outcomes. Obviously, what happens in classrooms is central to the educational development and the desired outcomes. However, to focus attention exclusively on classroom observation does not provide the complete picture of the realities of teaching and learning.

Teacher's' knowledge begins at the planning stage, and the teacher perception of different pedagogical and methodological aspects will provide insights into the quality that will be reflected in classroom contexts. Susana, Franco, Yolanda, Padrón, Hersh, and Waxman

(2015) investigated EFL teachers' knowledge and professional development, in which the conclusions indicate that many EFL teachers are not receiving adequate training, mainly in areas related to EFL teaching and knowledge related to second-language learning. In quantitative investigation of EFL teacher knowledge on Australian adults, Yazdanpanah (2015) found that teachers were more interested in teaching activities for the language skills rather than using effective methodologies, approaches, and needed techniques while teaching. As teachers' knowledge provides a clear picture of how learning outcomes are achieved in the classroom, this study looks at teachers' knowledge of Socratic questioning and cognitive questions that are related to students' critical thinking and how the teachers' knowledge of SQ can help teachers to improve and develop their pedagogical knowledge in order to implement more effective strategies that are meant to enhance students' CT.

Educational researchers classified teachers' knowledge into three categories; content knowledge, pedagogical knowledge and pedagogical content knowledge (Shulman, 1968b, and Eggen and Kauchak, 2002). The current study aims to investigate and focus on the pedagogical content knowledge. Pedagogical content knowledge is the knowledge of how to teach specific contents (Watzke, 2007), which includes knowledge of how to explain specific ideas such as how to demonstrate and rationalize procedures and methods and how to correct students' theories and misconceptions about subject-matter (Hammond, 2006). Bigelow and Ranney (2005) investigated pre-service teachers' knowledge about language and its transfer to lesson planning. They found that twelve out of nineteen participants expressed some degree of anxiety or confusion about new materials because they lacked pedagogical knowledge.

Pedagogical content knowledge plays a significant role in identifying the need for teachers to transform content knowledge into forms that are accessible to students and reflecting on how to accomplish this transformation (Backett and Slater, 2005). While all the participants in the current study have basic professional qualification to teach the subject matter, it is important to investigate their perception of SQ and cognitive questions, and whether gender and nationality have any differences on their actual use. However, before

getting into the further discussion of this investigation, there is a need to present findings of some classroom researchers on teachers' knowledge in order to provide a clear perceptive on the matter under this investigation. After examining the literature on teachers' knowledge of questioning strategies, it is in the right scope and sequence to discuss another dimension of the teacher's knowledge. The following section is devoted a discussion of literature available on classroom research on teachers' knowledge.

2.3 Classroom Research on Teachers' Knowledge

Contemporary constructivist teaching practices evolved from an intellectual perspective regarded knowledge as constructed rather than received, and uncovered rather than retrieved accumulated ideas (Backett and Slater, 2005). Building on Dewey's propositions, Yazdanpanah, (2015) maintained that through processes of questioning, debating and discussing as well as comparison of practices and conceptions in regard to others, teachers would gradually enhance both their stock of knowledge and its pedagogical delivery. Within this paradigm, educational research is no longer undertaken with the sole objective of guiding teachers' actions, however, it rather stresses that the teaching courses are geared towards enlightening teachers about both possible applications and the positive attitudes that should consolidate their experience (Watzke, (2007).

As Chaudron (1988) remarks, research in both education and applied linguistics has witnessed a subtle shift of emphasis from research focusing on students' learning outcomes and academic achievements to classroom applications and learning processes, leading to "teacher research". As a result of this shift, classroom research is carried out with the goal of addressing the vital pedagogical issues of concern in teaching and learning. Furthermore, Hammond, (2006) views teacher research as the systematic and planned process whereby teachers intentionally carry out their pedagogical investigations.

Nunan, (2003), in his description of the scope of classroom research focused on the issues that comprise the developmental features of learners' language and the learning

techniques and styles employed by students. The research interest also extends to the kind of language necessitated by the variety of teaching materials and pedagogical activities in conjunction with the interaction between learners and how it has its bearings and consequences upon the development of learner language. A substantial body of second/foreign language research is conducted in the subfield of Applied Linguistics identified as Second Language Acquisition (SLA).

The main premise of SLA is the description and various configurations of the phases that learners undergo in acquiring a language as well as the identification of the techniques the learners develop to attain take-off level proficiency in the target language under consideration. The objective of in-classroom research, then, logically is to demarcate the pedagogical inputs that is likely to facilitate or hamper language learning. Such inputs may be linked to the instructional treatment and /or context, such as the teacher, the learner, or, indeed, any relevant manifestation of interactions between these factors (Bakett and Slater, 2005).

Classroom research can alternatively focus on the learner, the interaction between the teacher and the student. Teacher-oriented research by contrast investigates such concerns as classroom decision-making processes on the part of teachers and related teacher talk. The latter encompasses the modifications effected by teachers when engaging with the second language learners as well as the nature and amount of such talk combined with the kind of error correction tips and feedback provided with and the questions posed by teachers (Chaudron, 1988). The general consensus on the potential value of teacher talk as a comprehensible input (Krashen, 1985) for the learners has led to the recognition of the need to identify and highlight the characteristics and requirements for language acquisition in classrooms.

Four strands of research in second language acquisition are identified by Chaudron (1988), namely psychometric research, discourse analysis, interaction analysis, and ethnographic research. Psychometric research is normally characterized by the use of

designated experimental methods. Interaction and discourse analysis approaches strongly favor analytical observation processes and strategies. Yet, the former concentrates on the social meaning which manifests itself during the classroom interaction, while the latter is limited to the analysis of the linguistic aspects alone in the interaction. The fourth tradition is ethnography of communication which, according to Chaudron (1988), is largely a subjective paradigm that provides with interpretive insights into various phenomena occurring in the classroom. Yet, as Nunnan (2003) points out, whether to regard interaction analysis and discourse analysis as distinct traditions is a widely contested issue. The authors conceive of them merely as data collection methods rather than as discrete traditions. An acceptance of such a view would reduce Chaudron's taxonomy belonging to two strands. Indeed, this is consistent with the widely held differentiation within the educational literature between qualitative and quantitative methods (Nunnan, 2003).

Much greater significance has recently been given to the integrated and interactive aspects of the emerging patterns of the participants' classroom behavior, including turn-taking, questioning and answering, negotiation of meaning and the mutual exchange of feedback. This shift is justified by the fact that learning a second language is by its nature a highly interactive process, and that the nature of such interaction is deemed to have a substantial effect on students' learning (Tsui, 2002). Tsui also assumes that the increasing recognition of teachers' talk in pedagogical research is ascribed to its potential effect on learners' comprehension which is regarded as an important prerequisite to language acquisition.

According to Nunan (2003), teachers' classroom questions constitute an integral part of teachers' talk and they are deemed by to play a crucial role in the language acquisition process. In fact, studies on questioning behaviors confirm that they consist of 20 to 40 per cent of the total teachers' talking time (Chaudron, 1988). In the majority of classroom settings, interaction is triggered by teachers' questions and questioning techniques. A study by Tsui, (2002) into teachers' talk in Hong Kong has demonstrated that 70 per cent of such interaction is composed of the teacher posing a question and then assigning a student to provide an answer, and the

answer itself to which the teacher responds with a feedback. It can, then, be concluded that questions constitute an important element of classroom talk. These questions are typically employed to verify students' comprehension, and to ensure that they have acquired the knowledge presented to them. Questions are equally useful in capturing students' attention, in involving them as well as in advancing the lesson and enforcing discipline in the classroom. Finally, questions are instrumental in directing students to practice certain linguistic items and use the target language as a normal channel of communication (Tsui, 2002). Having discussed the classroom research and the findings on teachers' knowledge, the following section focuses on the role of classroom questions played in SLA. After reviewing the literature on different aspects of teacher's knowledge related to questioning strategies and CT, the following section focusses on the discussion of literature related to the role of classroom questions in second language acquisition.

2.4 The Role of Classroom Questions in SLA

A considerable number of SLA interactionist theories and research were heavily influenced by psychology-related learning theories. Masatoshi and Susan, (2016) have claimed that much second language is acquired through interaction when conversing with others. This is close to the first language acquisition theory which considers child-directed speech as important in the first language acquisition process. Long and Sato's (1983) claims are based on the observations of authentic interactions between EFL/ESL students and the native speakers. Long and Sato (1983) argue with Krashen (1985) who advocates that comprehensible input is of utmost importance in language acquisition. However, Long and Sato (1983) are more concerned with "how" input is made comprehensible during the process of language acquisition. They view the modified interaction as an important mechanism for this to be realized. Long and Sato's (1983) view, on what learners really require is not essentially the explanation of the linguistic forms, but rather on occasion to interact with other speakers in communicative manner which leads learners to modify what they are saying until the learner reflects on the responses of the understanding to what is being said. According to Masatoshi and Susan (2016), beginner language learner acquires language through interaction with native-speaker which has been

adjusted in some way. In fact, Long (2007) claims that native speakers steadily adjust their language in sustained conversations with non-native speakers to meet the demands of the communication in question.

Modified interaction must be needed for language learning and acquisition. This connection has been concluded as follows (Long, 2007):

1. Modified interaction makes input more comprehensible;
2. Comprehensible input supports language acquisition;
3. Modified interaction fosters language acquisition.

Lightbown and Spada (1999) believe that modified interaction is not only involving language clarification, but it also provides explanation, relaxed speech level, gesture, and additional contextual clues. Here are some examples of conversational modifications introduced by Lightbown and Spada, (1999):

1. Comprehension checks : Efforts made by the native speaker to confirm that the learner has comprehended what is being said e.g., 'The bus leaves at 6:30. Do you understand?'
2. Clarification request : Effort made by the learner to urge the native speaker simplify an utterance which was not understood e.g. 'Could you repeat that please?'
3. Self-paraphrasing : The native speaker says again his or her sentence either partially or in its entirety to accommodate the EFL/ESL learner e.g., 'She got lost on her way home from school. She was walking home from school. She got lost.'

Research findings have demonstrated that conversational adjustments can aid and foster comprehension. There is evidence that modified interaction leads to a better understanding than linguistic simplification does, particularly the modification which is planned

in advance (Masatoshi and Susan, 2016). Vygotsky (1962) gave another viewpoint on the function of interaction in second language learning which is based on the theory of human mental processing. Vygotsky's theory claims that all cognitive processes including language modifications results as an outcome of social interaction between individuals. Considering Vygotskian second language acquisition theory, Long and Sato (1983) believe that collaboration and interaction promote second language learners' high-order linguistic competence when this collaboration takes place between knowledgeable speakers and advanced learners.

Vygotsky evoked the concept of the Zone of Proximal Development (ZPD) which is the stage of comprehension, in which a learner is able to show off when there is a motivation due to communication with a more advanced speaker. This can be seen in a variation of speech techniques used by advanced speakers to create helpful environment for the second language learner to understand and produce language using simplification, repetition, modeling techniques.

In short, if a questioning approach is a central aspect for promoting classroom interaction, teachers should utilize in-class questioning in a skillful manner in order to foster students' higher-order thinking. This issue leads us to in-class questioning as a subject of concern. Stevens (1912) believed that teachers spend approximately eighty percent school day on asking questions to students. This claim is supported with more recent research findings on teacher questioning behaviors and patterns which proved that this practice has not changed. Masatoshi and Susan (2016) claimed that teachers today ask between 300-400 questions each day to their students. According to Stevens (1912), teachers usually ask questions for the following reasons:

- To keep students actively involved in lessons;
- To seek answers for some questions and offer students the opportunity to

- openly express their ideas and thoughts;
- To enable other students to hear different explanations of the target ideas of the lesson by their peers;
- To pace their lessons and moderate student behavior; and
- To evaluate student learning and revise their lessons as necessary.

2.5 The Role of Display and Referential Questions in SLA

Having discussed the literature on the role of classroom questions in SLA, what follows next is the review of literature on the role of more focused display and referential questions in SLA.

To the concerns of communication in language classrooms, a further distinction has been made between "display" and "referential" questions. A "display" question is one to which the person asking the question already knows the answer whereas a "referential" question is the one whose answer is unknown to the questioner. Multiple studies have recently examined teachers' use of "referential" and "display" questions. Referential questions are believed to be more authentic than display questions whose main purpose is to test students' knowledge of language (Chang, (2010). On comparing the types of questions asked by native speakers to second-language speakers outside the classroom context to the questions asked by teachers to students in a classroom context, Long and Sato (1983) found that in the first case, the majority of questions are "referential" questions, whereas, in the second case, "display" questions are dominant. This means that teachers significantly asked more display questions than the referential questions.

Similarly, a study conducted by Doughty and Long (2005) examined the impact of referential questions on classroom discussions. Four seasoned ESL teachers and twenty-four non-native speakers were involved in the study; two of the teachers were trained to include referential questions into their classroom instruction, while two others were not trained. Longer and more syntactically complex and elaborate responses were given in groups in which

referential questions were asked to students. The authors of this study even maintained that teachers can learn to change the types of questions they ask and to include more referential questions thereby increasing the amount of communication taking place in the classroom.

The value of drawing a distinction between display and referential questions has been questioned by van Lier (1988), who pointed out that questions are made with the aim of eliciting language from the learners. According to van Lier, what matters more than drawing a distinction between referential and display questions is the fact that the main purpose of both the types is to get learners to produce language. Thus, what differentiates between the instructional type of questions and the conversational (non-instructional) type is not their referential or display nature, but their eliciting role. In interactional terms, the difference between the two elicitations may be minimal (van Lier, 1988). Nunan (2003) opines that while van Lier's (1988) criticism is a perfectly legitimate one, it is not implausible that the efforts involved in answering referential questions generates a better effort and depth of processing on the part of the learner. Referential questions normally require more thoughtful responses. This, in turn, may well be a greater incentive to acquisition than the answering of display questions.

Tsui (2002) also examined the effect of display and referential questions on meaningful interaction in the language class sessions with regard to the differences between social discourse and classroom discourse. She stated that the discussion and shaping of meaning is often absent in classroom exchanges initiated by display questions because the meaning of the questions and what forms an appropriate response are preconceived by the teacher. If the student response does not match what the teacher deems convenient, it will be either rejected or negatively evaluated. Tsui even brought about two limitations concerning the distinction of display or referential questions. Firstly, she warned that what looks like a referential question may not be intended as one by the teacher. Conversely, the teacher's referential question is sometimes taken as a display question by the student. Hence, when analyzing questions, it is important not to use the surface form of the question as the sole criterion for classification.

The intentions of the teacher, the interpretation of the student, the context in which the exchange arises, need also to be considered (Tsui, 2002). Secondly, Tsui quoted some studies to argue that it is not necessarily that referential questions promote longer and syntactically more elaborate responses. Wragg and Brown (2001) studied four ESL teachers' questions in terms of eliciting responses and found that referential and open-ended questions tend to be even less effective compared to display and closed questions in getting student responses. In the four lessons that he studied, although a large number of referential and open-ended questions were asked, only one elicited an elaborated response. Wu accounted such contrasting findings for the cultural background of the students. According to him, Chinese students in Hong Kong generally tend to shy away from answering and when they are asked to respond, they opt for producing short answers lest they are viewed as showing off by peers. Tsui concluded that although Wu's claims (as cited in Tsui, 2002) have yet to be substantiated by further research, it is certainly true that the type of question asked by the teacher is not the only factor that affects the kind of interaction generated in the classroom (Tsui, (2002). This remark is in line with Nunan's (2003) observation that reminds researchers of the possible factors, such as the field of research, the learner's background and contextual and interpersonal variables operating in teacher questioning in L2 studies.

Further studies on second language classroom discourse have shifted the focus to modifications of interaction structure. These researches have attempted to explore the nature of native speaker teachers' strategies in assisting learners understand their utterances in comparison to the use of these strategies by teachers interacting with non-native speaker learners (Bezemer and Mavers, 2011). Interaction modification strategies applied by teachers to help students comprehend their speech have been acknowledged and matched to those used by native speakers in NS-NNS conversations in several studies (Luk, 2013 and Luk and Lin, 2007). Bezemer and Mavers (2011) have identified comprehension checks, confirmation checks and clarification request as three types of questions denoting the direction of information flow in preceding utterances. They specifically work to sustain interaction and establish a common ground of understanding and clarity in sharing the same meaning by

language users. While comprehension check functions as an indicator of listener confirmation of reception of accurate message, confirmation checks and clarification requests enable the speaker to establish correct interpretation of the way the listener reacts. If measured for both teacher and students, these types should contribute to an index of interaction or negotiation in the classroom. It was emphasized that more interaction is an aid to acquisition (Bezemer and Mavers, 2011).

Long and Sato (1983) compared three types of questions used by L2 teachers in beginning-level classes with occurrence of question types in a set of data obtained from native speakers in dyadic information gap tasks with L2 learners. In both comparisons, while comprehension checks were the least frequent in the dyadic NS-NNS tasks, they were the most frequent in the classroom interactions. As Chaudron's (1988) study reports, confirmation checks, which were the most frequent in the dyads, were not so frequent in the two classroom data sets. Clarification requests were low in frequency in all conditions. Thus, it has been suggested that the amount of interaction modification strategies employed by the native speaker or the teacher determines how much negotiation work might take place between the native speakers and the non-native speakers or between the teacher and the students, which in turn influences the information that has been made comprehensible to the non-native speaker or the learner.

Tsui (2002) conducted a study on reading comprehension classes in Hong Kong secondary schools. Based on her findings, she argued that while it is certainly true that the instances in which modification devices are employed are indicative of how far the teacher is aware of the importance of making the input comprehensible, it must not be assumed that the use of these devices necessarily leads to the production of understandable input. The nature of negotiation is that it allows interaction between the speaker and the learner. It could be argued that it is challenging to ascertain the quality of the input and the amount of negotiation work without looking at the feedback provided by non-native speaker or learner and the degree of their involvement in the whole classroom interaction process.

Chaudron (1988) and Tsui (2002) rightly argue that it is of paramount significance to differentiate between modification devices which are used by students and those which are used by the teacher. The former is a far more important indicator of students' participation in the negotiation of meaning than the latter. It is also important to look at the kinds of modification strategies used because some are better indicators of learner participation in the negotiation task than others. For example, clarification requests and confirmation checks that can only be used in response to students' initiations are better indicators than comprehension checks and self-repetitions, which can be used many times without any student feedback. After examining the multi-level perspectives of research on teachers' knowledge and the types of classroom questions, it is now time to examine and discuss another aspect of this study, which is critical thinking. The following section is devoted to the review of research on critical thinking.

2.6 Critical Thinking

Many researchers concluded that there is no consistent agreement on how to define CT, on what are the required strategies and skills to teach CT, and constructing the best approaches and methods for this teaching. Mostly, even now educational curricula have not finalized common applications to teach CT. This section will reflect on these different views in order to construct a clear understanding of different aspects related to critical thinking.

2.6.1 Critical Perspective on Critical Thinking

An extensive review of literature indicates that some scholars employ CT and high order thinking seemingly. Elliot (2006) is one of the scholars who were interested in studying and using CT and higher order thinking skills in teaching context. However, other scholars have different use of these two concepts and implement them differently according to the proper context and aim (Zohar,2004).

The connection between "critical thinking," "higher order thinking," "thinking skills" and other terms such as "informal logic," "informal reasoning," "problem solving," "argumentation," "critical reflection," "reflective judgment," and "metacognition" have made them more difficult to apply differently. Scholars also have other issues that led to different understanding like (a) the degree to which CT is related to specific subject, (b) variations between novice and expert thinking in which novices can learn to reflect more like experts, and (d) if CT should be implemented as a process-based practice or a package of skills (Qing, Jing, and Yan, (2010); Keng (1996); Gambrell (2005); and Thayer-Bacon (2000). While some scholars have tried to use order on these different terminology e.g. "critical thinking," "higher order thinking," "thinking skills" and other terms such as "informal logic," "informal reasoning," "problem solving," "argumentation," "critical reflection," "reflective judgment," and "metacognition" Preiss and Sternberg(2010), none of them has managed to provide a description, classification, or even a theory could be conventional as ultimate classification (Martin, 2010; Folsom, 2009; Chenault and Orsello, 2008; and Dagli, 2008).

The main problem to this disagreement has rested in the grounding of different hypotheses separate fields that is related to this study. These two fields are philosophy and psychology. The current study aims to discuss both views in order to provide a clear picture of the nature of critical thinking in education.

Many philosophers have tried to discuss the quality and the nature of the results or conclusions of CT, such as analyzing justifications and reasoning responses. However, psychologists have focused on cognition procedures and process, the elements and practices implemented to indicate practical learning issues. In addition, psychology has been tested and examined through empirical research process, while philosophy has depended on logical justifications and reasoning to conclude solutions and recommendations. On the other hand, some educators realized the importance of both areas of psychology and philosophy to improve an accurate and reliable teaching methods and procedures of CT (Kuhn, 2005; Giancarlo, 2004;

and Marzano, Pickering, and Pollock, 2003). The current study reflects on this dilemma and will try to investigate the correlation between psychological and philosophical aspects that affect teaching and learning of ESL (English as a second language) through CT. Having looked at the critical perspectives on CT there is a need to discuss it in its more focused perspective. The following section deals with a discussion of researches related to philosophy-based theories and definitions followed by psychology-oriented theories and definitions in the following section to aid the ongoing discussion in the right direction.

2.6.2 Philosophy-Based Theories and Definitions

Since the time of Socrates, CT has been related to philosophy. The importance of CT in the recent different educational reform projects has been directly related to informal logic as a separate field within the area of philosophy. Many researchers consider informal logic as logic that is based on evaluation, interpretation, construction of argument and justification. Informal logicians consider CT as a general concept that includes and depends on the recommendations and conclusions of informal logic, not only that but it takes from other types of logic as well as from advantages of other fields (Johnson, 2008). The contribution of informal logic has been considered as a basic theoretical ground for CT.

Philosophers consider informal logic as a meeting point for examining and improving CT and philosophy-based theories, also, they have addressed various aspects and factors that are related closely to CT and positively differ in a number of ways, they also reveal common concerns (Ennis, 2008; Siegel and Biro, 2008; Paul and Elder, 2006; and Lipman 2003). Johnson's (2008) analysis reflects the importance of CT on cognitive theories and competences introduced by informal logic with neglecting the active role of "affective propensities" on the practice of those skills.

The work of Paul and Elder (2006) is considered as an important work for scholars that uses both psychological and philosophical contributions to CT. Paul and Elder's theory of CT was constructed on a philosophical ground that has been considered by CT theorists as supportive area for informal logic research. In contrast, other informal logicians have neglected clarifications of concepts, taxonomies, and specific procedures of justification analysis. While their analysis has avoided most formal terminology, it has reflected findings from other fields. Much of Paul and Elder's reasoning has revolved around every day issues or problems that are included in the content and structure knowledge different academic fields. Paul and Elder have often referred to these "ill-structured, multidisciplinary problems as multi-logical issues".

Paul (2010) has also discussed the difference between two major concepts in his theory, "strong sense" critical thinking and "weak sense" critical thinking, that indicate a basic moral concern in his theory with pervasive bias and egocentric thinking. Paul's conclusions of cognitive qualities have been introduced to begin a clear distinction between (a) "sophistic or weak-sense thinkers". This kind of thinkers tend to support their personal interests and defend inadequacies in justifications and arguments of others, rather than use those skills to their reasoning; (b) "true critical thinkers", this kind of thinkers struggle to know how to control their egocentric and ethnocentric biases. They use their cognitive skills to their justification and reasoning, and their interest to reach reality. Therefore, self-judgment has been a further research interest of Paul's theory.

Paul has claimed that CT must be defined in many various ways that should not be considered equally selected. In his different definitions of CT, it is viewed as "thinking about your thinking while you're thinking to make your thinking better" (Paul, Willsen, and Binker, 1993, p. 91). Paul's definitions indicate the metacognitive factor of CT, self-regulating, and the knowledge of how to assess thinking based on certain criteria and standards. He considers CT as a method of eliminating the domination of narrow-mindedness, unjustified statements, and illogical practices that are brought to discussions, and this is the type of CT skills that is

expected to be developed by the participants in phase II of this study. Paul's disagreement to attach himself to a particular definition of CT has shown his priority in constructing comprehensive definition of CT, a definition that reflects different perspectives and insights of other fields.

Resnick (2005) concluded that the contribution of philosophy to CT introducing a systematic thinking that can protect people against their human weakness to egocentric thinking, misleading statements, and drawing irrational ideas, this is because it is less difficult than conclusion drawn through logical thinking options. In logical sequence, the discussion of philosophy-based theories and definitions leads to the discussion of psychology-oriented theories and definitions in the following section.

2.6.3 Psychology-Based Theories and Definitions

In comparison to philosophy, psychology researchers have based their conclusions about CT on theories of intelligence, cognitive and developmental psychology (Walker, Brophy, Hodge, and Bransford, 2006; Sternberg, Roediger, and Halpern, 2007; and Halpern, 2002). Psychologists have tried to relate problem solving to CT. However, philosophers tend to deal with problem solving and CT as equivalent concepts. Sternberg et al (2007) has concluded that CT is "purposeful, reasoned, and goal directed". It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions" (p. 5). Though, Sternberg prefers using CT, other psychologists have used "thinking skills" as a general term (Miri, David, and Uri, 2007). Generally, cognitive-based theorists have analyzed and examined skills that contributed to thinking critically, often neglected characteristics of a good critical thinker and criteria for assessing thinking. In addition to general conclusions, recently some psychologists have started giving attention to learners' acts and have in order to examine models of CT (Wegerif, 2007, and Zohar, 2007).

Many psychologists believe that Bloom (1956) taxonomy of educational objectives is considered as a foundation for psychology-based models and cognitive skills studies (Kite,

Stockdale, Whitley, and Johnson, 2005). Psychologists have presented fast increasing knowledge base for different models for CT (Cheng, Cheng, A., and Tang, 2010; Hoa, 2008; and Yamashita, 2007).

Researchers of CT and thinking skills have directed their research to investigate the background knowledge and representations of knowledge in “experts and novices’ studies of different domains. This kind of investigation has added to the previous conclusions that related to how problem solving process changed improved experience and knowledge. In the past twenty years, this field of cognitive studies was limited to problem solving in structure-based fields like math and physics (Chi, Feltovich, and Glaser, 1981).

Recently, English as a second language has started to become an area of interest for this type of research by educationists and psychologists (Vickers, 2010; Clark, 2008; Kasper, 2006; Meskill and Anthony, 2006; Borg, 2006; Sasaki, 2005; and Johnson, 2003). Studies on teachers' knowledge and actual use on students CT has exposed to the knowledge of domain-specific indicators of CT and has supported to clarify instructional goals such as encouraging students to think and to practice organized dispositions and patterns of thinking. For example, Paul (2010) examined teachers' attitude and conceptions of CT and found that teachers in the study generally felt that CT must be included in the curriculum and considered as an essential part of teaching, they also pointed out training in teaching CT is needed. However, it was strongly indicated that teachers have no clear notions about CT; instead they have general understanding without having proper and specific procedures and methods to teach CT.

Psychologists have also added to our comprehension of teaching CT. For instance, Perry (1970) investigated cognitive development in male college students which was followed and then refined by studies that investigated the cognitive development of minorities and women (e.g. Goldberger and Crow, 2010; Baxter-Magolda, 2008; and Baxter-Magolda, 2004). They added useful conclusions on different developmental problems that learners might encounter

in learning to think critically, therefore, teachers' knowledge of different pedagogical issues would help EFL teachers to solve different learning challenges.

The research design that was used by Baxter-Magolda (2008) is comparable to what Perry (1970) used in his study. However, Baxter-Magolda (2008) analyzed an equal gender size; her study was longitudinal research where she followed students through their undergraduate program. In her study, Baxter-Magolda indicated percentages for the number of students who might fall into a particular category of the way in which students viewed the nature of knowledge and responded to tasks requiring CT. From the findings, 1) absolute knowing (dualism above) was common among new male students (68%) which, declined to 2% among more advanced male students; 2) transitional knowing (multiplicity/subjective knowledge) was a quality of 32% among new male students and 80% of more advanced male students; 3) independent knowing, her perception of a position close to relativism discussed above was infrequent among the target students, she concluded only 16% of more advanced students which increased to 57% after their graduation; 4) contextual knowing was low; it was 12% among participants only before graduation. The conclusions derived among psychologists indicate that instruction must provide some kind of challenge that involve students in using higher-order thinking skills which enable them to do more challengeable tasks.

Recently, limited attempts between scholars indicate the need of having common agreements between scholars who are interested in CT. However, though some scholars are from different disciplines, they refer to each other's' conclusions and findings and have attended to each other's workshops and discussions (Halpern, 2002; Goldberger and Crow, 2010). Many recommendations and findings presented by different psychologists and philosophers have presented a significant overlap Singh, Thomas, Mueller, Lims, Perkins, and Zhu, 2010; Ennis, 2008; Magolda, 2008; Johnson, 2008; and Halpern, 2002, and a number of different attempts to gather the works done by psychologists and philosophers that are related to CT have presented in the published studies and research (Leighton and Geiri, 2007; and Schuh, 2004). Another attempt to synthesize contributions was brought by Paul and Elder (2006) who have

called for combining visions of different scholars from different domains in a clear constructed theory of CT. Therefore, Paul and Elder (2006) started to introduce “stage theory” of CT that is based on the two domains of philosophy and psychology to CT.

Possibly the most widely known organized attempt of a scientific inquiry in the area of CT was initiated by the American Philosophical Association to reach an agreement by a committee of specialists in CT related to educational development of instruction and assessment in different teaching and learning stages (Paul and Elder, 2006). A group of 46 scholars gathered from different disciplines and fields contributed in the “multi-year qualitative research project” (Paul and Elder, 2006). As regards, “(52%) of the participants were philosophers, and the rest were associated with education (22%), the social sciences together with psychologists (20%), and the physical sciences (6%)”. The report resulting from this investigation is commonly known in the critical thinking literature as the Delphi Report” (Paul and Elder, 2006).

Those experts define CT as one form among different interrelated forms of higher cognitive thinking, beside other forms like decision making, creative thinking, and problem-solving (Paul and Elder, 2006). They concluded that these different forms of higher cognitive thinking are overlapped complexly and conceptually, and interrelation between them has not acceptably investigated. The concluded agreement among Delphi experts has both affective and cognitive domains. The definition of CT that was concluded is general and abroad. However, they are still divided on the concerns of whether or not CT contains a clear defined dimension as included in Paul’s interpretation.

Other scholars have dealt with higher cognitive thinking as a container that includes CT, decision making, and problem solving. In addition, there are other skills that are included in CT like argumentation, reasoning, judgment and justification about ill-structured problems. In addition, CT includes affective dispositions that must be presented with other cognitive skills. The major conclusions and agreements in Delphi report and intellectual standards recognized

by Paul and Elder (Foundation for Critical Thinking, 2006) are used as the definition of CT in this study. After being familiar with the philosophy and psychology based theories and definitions, it is now opportune time to move to the discussion of a perspective on teaching for CT.

2.6.4 Teaching for Critical Thinking

Presumably, if the college students attend their classes, participating and listening to lectures, writing their assignments and completing other learning activities, they would develop their CT skills. However, many researchers have indicated that developing students' CT skills needs more direct and well planned teaching of CT skill (Paul and Elder, 2009; Muspratti, Luke, and Leonards, 2009; Facione, 2007; and Ege and Kutieleh, 2004).

Until recently, there is no strong evidence or conclusive research findings on the most successful instructional methods for developing students' CT skills. Forsyth, Paul, Kelley, and McMillan (2009) reviewed 27 studies that examined the result of different programs and courses on CT skills among college students, and they conclude that even results do not support the application of explicit instructional or course practices to increase CT; they did support the findings that college attendance improves CT. McMillan (1987) has concerned against using these conclusions to all courses or methods, quoting incomprehensive research designs, using weak instrumentation inappropriate to the treatments being evaluated, and unjustified definition and theory of CT. Halpern (2002) has recommended that the assessment tools existed in a particular study might add to the conflict of deciding the effectiveness of different approaches for CT. she has argued that assessment measures and tools must be constructed "more sensitive" to measure comprehensive progress in CT abilities. Obviously, more research is needed to find out which learning practices provide the best increases in CT. The current study is another attempt of investigation that would provide an understanding of different variables that are related to students' critical thinking skills.

The focused research on CT, along with frequent attention in developing higher cognitive abilities and skills for learners at different levels of education and ability, has brought different methods and approaches to teaching CT skills. One has been concerned with the improvement of the courses taught and assessment procedures that involved in teaching courses (Ennis, Martin, and Sun, 2007; Savery, 2006; and David, Baumfield, Steve, Mei, and Jen, 2004). This approach has been commonly used in high schools' levels, particularly, in America such as California where the assessment and teaching of CT is a state-centered priority.

An alternative teaching approach recommends not adhering to specific plans and models while activating the enrichment of a classroom environment that supports learners' responses to CT by having a deep exposure of discussions, real questions and tasks introduced to learners, and focus on evidence and justifications to enhance written or spoken claims (Kamali and Fahim, 2011; and Beaumont, 2011). This alternative teaching approach is influenced by the concept of constructivism as a metaphor which suggests understanding of knowledge and learning through experience (Pelech & Pieper, 2010; Onuf, 2013). Furthermore, it also involves Sfard's (1998) position of participation metaphor as against acquisition metaphor that is actively involved in language acquisition. Incidentally, the acquisition metaphor compares learning to computers and containers which compels us to think of knowledge as a commodity to be accumulated in the mind as a repository, where the learner hoards the commodity (p. 5). Participation metaphor, to use Sfard's (1998) remark, defies the traditional distinction between cognition and affect and brings social factors to the forefront and thus deals with an incomparable wide range of possibly relevant aspects (p.12). As Pavlenko and Lantoff (2000) argue, participation metaphor, in opposition to acquisition metaphor, necessitates a shift in our focus from language structure to language use in context, including social context, in the process of language teaching and learning. The participation metaphor suits in constructivist learning environments and promotes students' voice, agency and inter-subjectivity. The participatory characteristics of constructivist model of education (c.f. Pelech and Pieper, 2010; Onuf, 2013) enhance learning as they focus on encouraging multiple representations of reality; avoiding oversimplification to represent the complexity and diversity

of the real world; viewing knowledge as constructed as against knowledge as given; emphasizing authentic activities and meaningful contexts; focusing on real world settings and non-linear instruction; providing stimulus for reflecting on experience; articulating context-bound characteristics of knowledge; and on acknowledging collaborative construction of knowledge through inter-personal associations and negotiations.

In addition to the above approaches just discussed and introduced, other related educational approach has incorporated plans or models to be involved in critical thinking across the curriculum (Swartz, 2009; Sternberg, Roediger, and Halpern, 2007; Keng, 1996; and Paul and Elder, 2006). Obviously, varied models and strategies are available to encourage learners to develop their CT skills, while each approach has its supporters; little empirical research has been done to decide if one approach reflects successful applications than another in developing learners' CT abilities and practices. Therefore, the current study aims to apply Paul and Elder's (2006) model of CT to investigate the result of applying teachers' knowledge of SQ questioning on students' CT. This is because Paul and Elder's (2006) model is characterized by the following five standards that reflect the nature of the current study.

1. The model raises important questions and thinking process, constructing them evidently and specifically.
2. Collects and evaluates information, processing abstract concepts to understand them efficiently.
3. Reaches well-justified conclusions and answers compared to related and relevant indicators.
4. Students think deeply within different ways of thinking, identifying and evaluating their implications and assumptions in practical consequence.
5. The model helps students to communicate commendably with others to reach solutions to difficult issues.

The strategies that are based on inquiry teaching are interactive-move specific. As a result, these strategies were to be applied by people and computer instructors as a domain-independent basis for enabling learners' Socratic inquiry (Lotter, Harwood, and Bonner, 2007). A cognitive theory for Socratic teaching consists of the Inquiry teaching strategies which are useful in the current study investigation of partially SQ. Every strategy has circumstances that would make instructors apply it and actions instructors to act the strategy.

Lotter, et al (2007) categorized them into four main procedures:

- Case Selection Procedures: where instructors choose a situation from a common area to debate that will likely reflect arguments.
- Entrapment Procedures: where instructors entrap learners into exposure of their misunderstandings.
- Hypothesis Identification Procedures where instructors attempt to scaffold learners' construction of assumptions, expectations, or procedures that would help them explain phenomena.
- Hypothesis Evaluation Procedures: where instructors involve learners to exam assumptions, expectations, or procedures.

Moor and Rudd (2002) in Figure 2.1 below specify that the basic SQ model reflects four elements

- Origin and source
- Support, reasons, evidence, and assumptions
- Conflicting views
- Implications and consequences

Every element reflects an extent to which the instructor questions learners. The instructor might question learners how they reach their conclusions. Instructors might also ask what

evidence that learners used to confirm their conclusions. They also might use questions to specify different conclusions and identify the meanings, implications, values, and consequences of a specific conclusion. Having reviewed the literature on Teachers' questions, CT skills, philosophy-based and psychology-oriented theories and definitions thus far, what we need to discuss further is the literature on the aspects of assessment of CT in the following section.

Figure 2.1: The Socratic Questioning Model, Moor and Rudd, (2002)

2.6.5 Assessment of Critical Thinking

Critical thinking assessment is a main issue in constructing programs that develop learners' CT abilities and skills. It is very difficult to develop solid teaching models of CT unless scholars and researchers reach a widely agreed definition of CT. In spite of having unclear theories of CT, different attempts have been made to improve assessment methods and techniques. There are three common methods that are widely used to assess CT: (a) commercial standardized tests that include general knowledge indicators; (b) instructor or teacher constructed assessments that try to identify indicators of CT closely related to the aims of the target course; and (c) learners' self-assessment where they assess their own thinking. Every one of these methods will be highlighted with reference to its suitability to the current study.

Commercial standardized tests that assess general CT abilities such as the Cornell Critical Thinking Tests, Watson-Glaser Critical Thinking Appraisal, and the California Critical Thinking Skills Test. Some psychologists have usually depended on multiple choice indicators that assess the main features of CT, together with interpretation, analysis, inference, recognition of assumptions, assessing credibility (West, Toplak, and Stonovich, 2008). Researchers have not claimed that those instruments test all aspects of CT. These tools have been carefully constructed and evaluated in relation to its validity and reliability, not only that but have been commonly applied as tools for assessing learners' skills to think critically (Zascavage, Masten, and Nichols, 2007). Its application as assessment measures reflects their simplicity of marking. However, though they assess how good learners reason from written texts, commercial standardized tests cannot test where learners are competent enough to produce a logical oral or written justified arguments, whether they can provide solution to open-ended problems, or can improve practices to apply CT abilities when appropriate. Some scholars have indicated that multiple-choice exams are not valid items of CT skills because learners have no choice to regulate their own queries or use their own "evaluative criteria" (Kelly, 2009).

Some researchers have supported learners-generated answers, including written genres, to assess effectively students' CT (Enger and Yager, 2009; Stein, 2008; and Norris,

Leighton, and Phillips, 2004). Many different general knowledge standardized essay exams for CT have been constructed as replacements to multiple-choice exams to test learners' abilities to activate arguments and to control the problem solving nature of CT. The Ennis-Weir CT Essay Test (Ennis and Weir, 1985) the best-recognized and most commonly used. It requires learners to read a text on daily matters having many reasoning inaccuracies and to build up their own justifications. This type of CT tests offers a number of benefits over multiple choice exams or teacher-constructed essay exams, as well as learners-generated answers, well-constructed reliability and validity. On the other hand, while this type of assessment has involved conditions and standards for marking essays, the cost and time involved in marking open-ended questions and the proficiency needed to mark them consistently has narrowed their use.

Further approaches based on learners' reasoning for their judgments and responses or produce their own justifications on this type of tests of CT are being researched as well. Tsui (2002) has argued that a learner's justification for a specific response should be considered, and she has suggested follow-up multiple-choice statements that measure learners' answers. Theresa (2007) has recommended the use of verbal feedback of thinking to measure multiple-choice answers. Oakleaf (2008) has argued for the additions of multiple-rating statements that help learners to rank, from different possible selections. He has proposed building up test statements so that a list of favorable responses can be referred to any number of self-determining test items, and individual responses can be selected a number of times or not at all. These approaches would narrow guessing of answers as an issue in exam scores. While several additions to assessments of CT are being examined by these and other scholars, this type of CT tests that contain these improvements are not yet commercially offered.

Current attempts have introduced the subject of CT nature in the form of a standardized commercially accessible test. CT assessment has been frequently interpreted as an essential part of CT or as a distinct but overlapping conception. The Ennis-Weir CT Essay Test assesses some CT practices together with reasoning ability, but the idea of assessing CT practices separately from CT skills is comparatively new. Carl (2010) has pointed out that a good test

must assess both a learner's CT abilities and whether they can apply those abilities without being directed to do so. The California CT Dispositions Inventory, constructed on the conclusions that is related to theoretical model and items identified by the Delphi Report professionals, tests for seven subsets of CT practices using a six-point Likert scale (Facione,2007). This kind of CT tests a limited effect to effectively measure changes in learners' CT abilities, however their careful development, consistent marking, and common procedures make them favorable to use by educational scholars and researchers.

Another approach to testing CT is teacher designed tests. Norris and Ennis (1989) have introduced criteria and indicators for teachers involved in improving assessment methods and tools for such aims as assessing domain-specific CT, assessing a CT course, ongoing assessments, or marking and grades system. While instructor-made exams must be used in the classrooms to measure CT, their application in educational studies measuring the value of different models or methods to teach CT has important limitations. Tools constructed for a particular experimental model or technique for CT might best control its strengths however the subsequent variety of tools and testing procedures has led to complications comparing the results of educational research projects. Maybe the most suitable method to measure learners' CT skills is to teach them to evaluate their own individual thinking.

Paul has discussed the importance of helping learners to evaluate their own individual learning, and he has claimed that to the degree that learners need advice from teachers, they have not attained a high level of critical thinking (Foundation for Critical Thinking, 2006). Angelo and Cross (2008) have also discussed the need for learner self-assessment practices. This method seems to involve a basic amount of teaching for CT and requires to be introduced widely by educational researchers. While it is very important for classroom practice, on the other hand, it needs a comprehensive understanding of CT and obligation from both teachers and the learners. In addition, this method of assessment, for several causes, does not match the required needs of demanding educational research.

Current thoughts on CT indicate that recent assessment applications should be reviewed, rejected, or changed. Researchers have concluded the need to construct reliable and valid assessments that measures the whole concept while providing consistency in marking. Till now there is no best method, and each method has its advantages and disadvantages. This study will use the Watson-Glaser Critical Thinking Appraisal Form A (WGCTA-A) to assess common CT skills due to its fixable design and consistent marking nature. Chapter three discussed further details about Watson-Glaser Critical Thinking Appraisal Form. After discussing the available literature on the assessment of CT, the researcher includes a brief description of higher order thinking strategies which is followed by a discussion of SQ and further researches conducted on it in the following section.

2.7 Higher Order Thinking Strategies

The following higher order thinking strategies are presented in this research for improving students' higher order thinking skills. These strategies are not exhaustive to CT, but there are other strategies and methods which can enhance students' CT. Socratic questioning (SQ), the inquiry learning, and cognitive questions are discussed in this study as strategies that might take an effective role in developing students' CT. The process of implementing higher order thinking skills could provide teachers with directions to practice methods of using more focused and purposeful learning environment, which develops higher level thinking.

2.7.1 Socratic Questioning

The real ancient texts that function as basic foundations of what Socrates really said and reflected a comprehensive knowledge of how Socrates applied the process of applying "counter-examples" to improve a series of questions that could increase the amount of thoughts, ideas, examples, and views to carry out the dialogue to a maximum advantage. It is clear that discussions and dialogues are the good means to maximize comprehension of the course content. On the other hand, there is no precise written manual that demonstrates how Socratic questioning should be used (Paul and Elder, 2006). Generally, there is no specific

Socratic model where someone can thoroughly try to use in teaching and learning practices.

Many scholars believe that the cause behind existence and extension of this approach is the way that is interpreted. Sahamed (2004) indicates that Socratic questioning (SQ), having occurred about 2,500 years, has logically developed in its various practices today and the method has better adjusted to suit various purposes. Clearly scholars believe that questioning is an important tool of comprehension in teaching and learning context. As a result, scholars believe that Socratic questioning (SQ) can be adjusted and applied in various ways to different stages of comprehension (Paul and Elder, 2006).

As a result, in literature we can underline different acts of applying this method, e.g. the "method of Dialectic "Socratic Method" (Paul and Elder, 2006). It also has been termed Socratic seminar and Socratic dialogue (Brickhouse and Smith, 2007). Keng (1996) considers that over different applications of SQ as unreliable and considers that as misapplication of the pedagogy since only unreal features are applied without the nature of SQ. These misapplications have directed other researchers to consider SQ just open-ended questions and answer procedures (McCoy, 2008).

There is clear evidence that SQ is not restricted to a particular conclusion of research. It is generally applied as a tool in different fields of education. A study called "The Rhodes" typology of questions was done by Wenning (2006). It is "a comprehensive treatment of content-directed question types", and is really well matched to be used in SQ centered by experiment and observation. All content-based inquiries in this typology are categorized into one of: *interpretive questions* have more than one answer that can be supported with evidence from the text, *informational questions* where the questioner seeks knowledge concerning a particular fact; *evaluation questions* are used in classroom as an assessment and evaluation tool; *explanatory questions* help students to describe the main ideas in the text; *procedural questions* relate to classroom procedures, routines and classroom management, *verificational questions* help the students to verify the required information from the text; and *heuristic*

questions help the students to process problem solving and self-discovery tasks (Wenning, 2006).

SQ has also been applied in therapy, most important "Cognitive Therapy". The aim here is to support reveal the evidence and hypothesis that identify thoughts in relation of problems. Thoughtful application of SQ allows a therapist to test a "illogical thinking" while having an open case that respects the internal logic to even the most superficially illogical thoughts (Paul and Elder, 2006; and Kenzic, Wubbels, Elbers, and Hajar, 2010).

Other researchers pointed out that education reform is required in school curriculum and change according to Socratic logic is really needed. A reform movement was led by Alder (1982) called "Padeia Proposal", believing that SQ is an essential need for learning. According to Alder (1982), learning should be applied in three phases; in the first phase learners gain the knowledge over lectures and this is what has been called "Didactic Method". In the second phase, learners are supported to construct their cognitive abilities through the application of taught activities with the teacher. The idea is when the learners are prepared with the comprehended ideas they move to the third phase which is SQ and arguments (Alder, 1982).

In spite of the variation in the forms of SQ, the underlying principles reflect that some general features demonstrated are that (Sahamid, 2004):

- It is generally led by one person, the instructor, asking the questions to another person;
- It proceeds with the application of a set of questions to develop the topic at hand;
- The questioner initiates a position of "open-mindedness" and acceptance to create response from those being asked;
- A clear comprehension of the content or subject of those being asked is needed for the content to be developed.

2.7.2 Previous Studies on Classroom Questions and CT Queries

Having discussed various aspects related to the assessment of CT followed by relevant perspective of research on it and SQ in the preceding section, the following section briefs on previous studies conducted on classroom questions and CT queries. Garret (2006) conducted an action research study on higher level questioning at Socratic cafes. The purpose of this study was to determine how teachers pose effective questions, and then practicing SQ at lunch meetings, would positively affect sixth grade students' ability to engage in more reflective discussions about literature. At the beginning of the study, all sixth grade students read a story from the Open Court Reading Series for Sixth grade entitled *The Table Where Rich People Sit* by Byrd Baylor. Next, students began the weekly meetings of the Socrates Café Philosophers' Club. In order to introduce the Socrates Café, the researcher read the book, the Philosopher's club, to the classes orally. In it, the author, Christopher Phillips, helps the listener understand that his goal is not to insure that the listener knows how to answer questions, but rather how to pose thoughtful questions. The students were invited to join the experimenter during lunch in order to discuss topics of their choice. The experimenter quickly discovered that a small group is much more conducive to discussion and held separate Café meetings for different classes. In a large group, they were not as likely to hear from the quiet student.

At the closing of the study, students read a story from the Open Court Series, called *King Midash*. In order to more carefully control the consistency of classifying the questions, all students' questions were read and categorized according of the level of Bloom's Taxonomy. Based on the data collected, sixth grade students were dramatically affected by participation in the Socrates Café Philosopher's Club. The experimenter documented a substantial increase in the number of student generated questions written at the higher levels of Bloom's Taxonomy. The percentage of questions written at the Application, Analysis, Synthesis, and Evaluation levels of Bloom's Taxonomy rose from 29.4% in the pre-test to 76.1% in the post test. While this falls short of the hypothesized 90%, the experimenter now believes that 75-80% may have been a more realistic hypothesis, based on the cognitive level of most sixth graders.

Furthermore, the experimenter predicts that practicing Socratic discourse could have a long term effect on students. Philosophical reasoning using abstract thought has the potential to promote conflict resolution. Other important results included the transfer of learning from the students who chose to participate in Socrates Café to those who did not. Some of the students who participated in our meetings of the Philosopher's Club began to lead by example during literature discussions and could adequately provide examples of "Fat questions" that provoke higher level thinking skills. While there certainly was significant growth in the students' ability to construct higher level thinking skills the number of students who wrote questions at the higher levels of Bloom's Taxonomy was 76.1% rather than the hypothesized 90%. In retrospect, perhaps the speculated 90% was not only too high to attain, but was also probably higher than is desired based on sixth grade cognitive development. The experimenter would now seek to have students write 75-80% on their questions at the higher levels of Bloom's Taxonomy.

Schwarz, Hershkewitz, and Azmon (2006) investigated the role of the teacher in turning claims to Socratic arguments. The purpose of their experimental study was to identify structures that emerge from talk and interactions between the teacher and her students. The data has been taken from a research project in which sequences of activities were designed to lead Grade 8 students to learn conceptual knowledge in probability. The unit includes: Written individual pre-tests, 5 activities designed to take 10 lessons, structured as sequences of problem situations in different social contexts (whole classroom discussion, small group collaborative problem solving, and homework), and individual written post-tests.

Experimenters followed two teachers from two different schools who volunteered to take part in this experiment. Students in both schools were high achievers. Teacher A had taught probability once before; Teacher B taught probability for the first time. The teachers received the same instructions concerning guidance, specifically concerning leading the class discourse and knowledge construction. Each teacher was free to choose social contexts for teaching the unit (individually in homework assignments, in small groups, or in teacher-led

discussion with the whole class). They chose here one of the problem situations in which the two teachers led a discussion with the whole class on the same problem. By doing so, comparison between the two teachers was possible. The entire lesson for each of the two teachers was video-recorded, transcribed and analyzed. They sought to identify how they helped in constructing knowledge. Therefore, they looked for argumentative patterns of interaction the building bricks necessary for this purpose are: (a) claims – declarations, viewpoints, etc. and (b) arguments– claims followed by explanations. They present here one episode to compare the argumentative patterns of interaction for each teacher.

They found that teachers convey different subject norms and classroom cultures: with teacher A, students feel obligated to support claims by explaining; they are used to crystallize ideas by reaching agreement and negotiating mathematical meanings; with teacher B, students are committed to tune to the teacher's questions and to adopt her explanation as theirs. Weusijana (2006) investigates the impact of Socratic asking system through a web-based program on students' interaction. There were 63 senior undergraduate students in the Northwestern University course Biomedical Engineering entitled "Biomedical Engineering Laboratory". Sixty of them consented to be part of the study. All the students in the course were Biomedical Engineering seniors who attended the same class sessions but chose to attend one of two laboratory sections. About half of the students were in the morning laboratory section with the other half in the afternoon laboratory section held on the same day. Students work in the lab together in self-selected workgroups of 2 to 4 students. The morning control section had 29 students. The afternoon experimental section had 34 students. There were four groups, High Experimental Workgroup, Low experimental Workgroup, High Control Workgroup, and Low Control Workgroup.

The researcher found that, the High Experimental Workgroup has the highest total on-task utterances, followed by the Low Control workgroup, the Low Experimental Workgroup, and the finally the High Control Workgroup had the least. Students who were in High Experimental Workgroup were abler to reason for themselves, to derive general principles from

specific cases, and to apply the general principles that have been learned to new cases.

Sahamid (2004) conducted an action research study on Socratic questioning in the teaching of short stories. The purpose of the study was to explore the role of a particular model of Socratic questioning as a classroom strategy to develop critical thinking skills through learning short stories. The participants were the learners of residential school called the Maktab Rendah asaaim. These 15-16 years old learners whom were chosen would be cognitively ready to be presented with a strategy that dealt with the abstracts. The class was fairly typical of the classes in other 30 MRSM schools in the country; it was co-educational class made up of 11 male students and 13 female students, all ethnic Malay. The selected students had to attend MRSM, based on their academic achievements in national exam in the previous year, specifically for their outstanding grades in the subjects of Mathematics and Science. These students, in other words, were streamed to do the Sciences in MRSM based on their results. The study went through three phases which spanned the duration of over five months. During this time five short stories from the Literature Component of the Form Four and Form Five Syllabus was taught to students of class 4A. The main written tasks were assigned as data and assessed according to the Grade Profile. The purpose of assigning students the three writing tasks was to check for students' developing in critical thinking.

The data collected at the end of each phase was studied and decision made in order to make changes to teaching strategy and implement it into the next phase. The analysis at the end of each phase was more reflective in nature compared to the analysis at the end of the three phases. Sahmid found that in the first two phases students' anxiety was considerably reduced; their perceptions of the English teacher and what is expected of them in English classes were slowly evolving into something more realistic. By the third phase and into the fourth month of the study students were showing signs of adapting to the culture and practice of Socratic questioning.

Al-Kindi and Al-Mekhalfi's (2017) study focuses on investigating post-basic English teachers' practices of CT skills and challenges they face while teaching English language skills in Omani EFL classroom. on the practice and challenges of implementation CT skills. the sample of the study was drawn from twelve post-basic schools. The total number of 30 grade 11 and 12 English teachers from these schools were included. The data was collected through a questionnaire and class observation form. The questionnaire elicited the challenges that teachers' faced in the implementation of the CT skill. The class observation form was used to determine the frequency of teacher behavior to invite or to enhance CT in classroom. The results of the study show that there was (a) a lack in teachers' behavior to nurture CT skills; (b) there was a need to consider a number of extra curricula activities from the teachers responsible for; and (c) a need to establish a strong relationship between school and parents to help promote CT instruction inside the school; (d) English teachers should use questioning that invite CT skills in class and they should focus on asking more questions at the analysis, synthesis, and evaluation levels in order to create opportunities for students to practice CT skills.

Tuzlukova and Prabhukanth's (2018) paper reports on a study which examined that to what extent ESP foundation program courses at Sultan Qaboos University in Oman impacted CT and problem-solving skills intended for empowering the students' and developing their personalities. The study mainly focuses on how teaching and learning approaches related to CT and problem-solving skills, models and practices adopted in the context of ESP foundation program in Oman. The data was collected from 85 ESP foundation program students from Sultan Qaboos University using an on line questionnaire based on the motivated strategies for learning. The questionnaire contained statements related to students' knowledge, application, and integration of CT skills. The main finding of the study includes 54% of the sample were able to apply foundational knowledge involved in the cognitive processes at work in the ESP context and they were able apply CT, creative, and practical thinking skills in the given tasks. They were also positive about integration of reasoning skills for making value judgement in the CT and evaluation activities.

Kumar and James's (2015) study evaluated the level of CT variables among the students in Nizwa College of Technology Oman. The data was collected from 281 diploma students from engineering, IT, and business departments using a structured questionnaire which was adapted from Watson and Glasser (2006) CT model which includes five parameters of inference, assumptions, deduction, interpretation, and arguments. The findings of this study states 25% of the students in Nizwa College of Technology have high level of inference, 21% of the students have high level of assumption and deduction, and only 12% of students have high level of interpretation and evaluation of arguments skill.

Mehta, Al-Mahrooqi, Denman, and Al-Aghbari's (2018) study investigated CT skills of first year students at a public university in Oman to determine whether the CT skills were adequately developed at the school's level. An adapted version of the Cornell Class-Reasoning Test Form X was administered to 60, students who had just graduated from high school and entered the university. The results of this quantitative study revealed that participants had failed to master four of the five principles tested, and that they also scored at the levels that were mostly comparable to grades four to six school learners in the United States. Also, the participants of this study received lower scores than a similar group of students at Arab Gulf University who took the Cornell Class-Conditioning Test Form X. These findings suggest that the way CT is integrated into the curriculum in Omani schools may need to be reviewed, with particular attention to curriculum, teaching practices, textbooks, and assessment.

Mehta and Al-Mahrooqi's (2014) qualitative case study of university students from English language and literature major explored the way in which CT could be taught in EFL context. The results suggest that CT involves a series of skills, which can be continuously worked upon by students for whom constant revision and application of these skills is a significant way of internalizing what has often been seen to be just an attitude or bent of mind. The CT skills are further enhanced when students get an opportunity to write in the areas which have been discussed in the class, particularly if they are of some relevance to the

students' own contexts. The study concludes that continuous practice, both oral and written, provide opportunities for students to develop their CT abilities as they become more successful in incorporating subtle and CT ideas into their academic writings. By implication, this study suggests that students' academic and personal achievements depend on their CT abilities as inability to read critically will result in an inability to write insightfully. Having examined the perspective of research on classroom questions and CT queries, the discussion of the role of cognitive questions and affective questions is in scope and sequence in the following section.

2.7.3 The Role of Cognitive Questions

An extensive review of related literature on teachers' classroom questions reflect that there are many various and complex taxonomies and classifications. However, in the present study, the focus will be on SQ that enhance students' CT thinking.

Teaching learners to develop their thinking skills is increasingly known as a significant aim of education. If learners are to participate productively in the labor market society, they must be prepared with lifelong abilities and thinking skills needed to gain and apply skills and knowledge in the challenging work environment (Beanmont, 2010). Also, Cate (2002) indicates that the need of cognitive learning is considered as important need for learners to improve their attitudes and abilities of effective thinking.

Teacher questions are often regarded as a method that instructors apply to encourage learners' critical thinking (CT) abilities. Long and Sato (1983) conclude that only 29% of teachers' questions in basic education required CT. Chin (1983) believes that teacher questions can inspire learners' thoughts and develop thinking abilities. However, the current educational systems have not been "developing citizens who can make reflective decisions on their own about those things that concern them and the society at large" (Beanmont, 2010, p. 1). Many teachers focus on "recall learning" which involves learners only to repeat knowledge and ideas learned before

(Chin, 2006). The literature concludes information on the kinds of teacher questions frequently applied by instructors. Shamoossi (2004) found that instructors ask questions that cover only recall level of thinking.

Using Bloom's Taxonomy (Bloom, Englehart, Furst, Will, and Krathwohl, 1956) is possibly the commonly applied paradigm in educational projects to evaluate the kinds of questions instructors apply in classroom focusing on the cognitive domain of teaching and learning. Bloom et al.'s (1956) typology is a classification assumes that the cognitive level of the question asked is identified by the answer provided by the learners. Their central concern was to provide a method to form a classification structure of cognitive objectives by classifying questions as knowledge, comprehension, application, analysis, synthesis and evaluation. The first three types are usually applied to implicate lower-cognitive learning, whereas the latter three types implicate higher-cognitive learning. Therefore, to effectively interpret, analyze, synthesize, and evaluate, a clear comprehension of the lower cognitive types, such as memorization and comprehension is important. "When case analysis and discussion are limited to the first two or three categories, useful higher order insights are lost. Likewise, attempting to bypass the lower cognitive levels and start at the higher levels may be detrimental to learning" (Wood and Anderson, 2001, p 12).

Bloom et al. (1956) described each level in his classification, and then in a later book (Hayman, 1979, p. 8) on affective domain indicate that he is not sure whether his cognitive classification is really a hierarchical order of cognitive objectives or only a normal order of cognitive procedures with no logical classification at all. Bloom et al. (1956) defined the taxonomy categorizing as it correlates to learners' achievements as "What we are categorizing is the planned behavior of learners, the behavior in which learners are to act, think, or feel as the outcome of practicing some parts of instruction" (p. 78).

In Bloom's Classification (1956) higher order cognitive application has been expected to be more likely when learners are involved in evaluating or synthesizing information than

when they are involved in normal recall of information. Scholars have expected that learners who ask higher order cognitive questions that support synthesis, analysis, and evaluation rather than lower order cognitive questions that focus on recall of information, provoke CT abilities.

Bloom et al. (1956) defined the lower domain of the cognitive questions, as: "The most important behavior reflected in knowledge is whether or not the learner can recall and either refers to or identifies positive statements in answering specific questions. While somewhat more than actual memory is needed for knowledge, the type of the question and the level of accuracy and thoroughness needed should not be too different from the way in which the knowledge was naturally learned" (p. 78). Therefore, with a low level question the answer needs straightforward memory or recall rather than more complex cognitive processes.

Higher-level cognitive process is identified by Bloom et al. (1956) as "It is possible that tasks including synthesis objectives reflect a concrete experience than those including mainly acquisition of thoughts" (p. 167). One could accept that higher level cognitive questioning results higher level thinking. Research has concluded that higher cognitive questions aim to provoke higher level cognitive skills Tan, 2007; Gayle, Preiss, and Allen, 2006; and Xun and Land, 2004 findings indicate that effective instructors use higher cognitive questions. Other findings have also concluded that thinking is considered beyond simple memorization when instructors use questions higher than the level of simple recall (Renaud and Murray, 2007; Rupp, Ferne, and Choi, 2006; Scardamalia and Bereiter, 2006; and Jariah and Talif, 2005). Akkaya and Demirel, 2012; Gonzalez, 2010; and Taylor, Alber, and Walker, 2002 found that students who were exposed to lower cognitive questions had less critical thinking skills.

Other researchers support the use of higher domain of cognitive questions in educational research projects. Studies have concluded that when instructors apply and promote higher level cognitive questions that will result significant learning achievements. A meta-analysis of 18 experiments (the same 18 used by Winne plus two additional studies) by Redfield

and Rousseau (1981) has shown that the reattempted use of higher cognitive questions while teaching produced positive achievement on test results of factual recall and application of thinking abilities. "This analysis demonstrates, regardless of type of study or degree of experimental validity, teachers' predominant use of higher cognitive questions has a positive effect on student achievement. These overall findings lend support to previous conclusions regarding the importance of teachers' questioning behavior on student achievement" (p. 244).

The relationship between the kinds of questions applied by teachers and the learners' responses to questions is also investigated by (Beck and McKeown, 2002). They concluded that "talk dealing with text", in particular, is the major procedure to help learners acquire reading skills. Having discussed "decontextualized" language and constructing ideas using bottom-up approach was found to be important to understanding text, therefore learning from it. Beck and McKeown (2002) found that, learners in primary classrooms were not encouraged to interpret words and how words and ideas are related. After investigating primary teachers during story read aloud, they indicate that most text comprehension gains focus on simple recall or literal response from text where learners could produce one or two words to answer their questions. In an attempt to develop teaching performance, they activated "Text Talk" that encourages teachers to learn to provide learners more alternatives to respond to decontextualized language in more comprehensible ways. Thus far, studies on cognitive questions are discussed which further leads the discussion research on the role of affective question in language learning.

2.7.4 The Role of Affective Questions

According to Bloom (1956) the affective domain is constituted using feelings, beliefs, values as well as interests and appreciations of the students. The types of questions that students face in this domain require that they translate feelings and experiences that left an impact on them into words, and explain the way this frame of reference dictates their actions to a large extent. Although the cognitive domain is used more than the affective domain, the latter is as important as the former in instructions. The main preoccupation in education is linked with working on

improving cognitive abilities. However, honing cognitive skills warrants that students develop self-confidence, positive self-image, and a readiness to work hard to learn.

What matters more than investigating the learning outcomes and accomplishments of students, is the learners' attitude which should be part and parcel of teacher's questioning in addition to assessing their grades and improvement in terms of their test results. The distinctive components of attitudes are the cognitive, the affective, and the behavioral (Al-Kindi and Al-Mekhlafi, 2017). By helping students gain knowledge about important information, teachers are expected to ascertain that students develop a positive attitude towards what they are learning and towards the material they are using in the learning process. This can be done by engaging students in discussions in which teachers make sure that the discussions pave the way for students to shape the right attitudes towards the subject matter. Teachers are justified in attempting to have students learn about this technique and to develop either positive or negative attitudes towards it when speaking their minds and expressing their feelings during class discussions with teachers and voice their attitudes in order to enhance their awareness (Al-Mahrooqi, 2012). Research has proven that it is incumbent upon the teacher to devise a learning environment conducive to students' active involvement. Otherwise, students will fail to become active learners because they were deprived from the chance to get involved in the learning experience and from acquiring the needed skills to effectively perform in their learning (Chang, 2010).

2.8 Summary and Conclusion

This chapter on the literature review is divided into three core sections. The first section is devoted to the discussion of the concept of teachers' knowledge, classroom research relevant to teachers' knowledge, and the role of classroom questions in SLA. The second section covers the definitions of critical thinking, and the discussion of philosophy-based and psychology-based theories and definitions of critical thinking. This section also highlights the critical pedagogical perspectives relevant to teaching and assessment of critical thinking. The third

section deals with higher-order thinking strategies of Socratic Questioning, and presents the highlights of important studies conducted on Socratic questioning. This section also discusses the role of cognitive questions and affective questions in the broad framework of Bloom's Taxonomy.

Based on the educational researches, the teachers' knowledge has been categorized as content knowledge, pedagogical knowledge, and pedagogical content knowledge. Pedagogical content knowledge helps in identifying the need for teachers to transform the content knowledge into formalized pedagogical knowledge that is accessible to students for reflection on how to accomplish and achieve it. Classroom research on the teachers' knowledge is primarily based on contemporary constructivist paradigm of teaching practices that are evolved from a critical intellectual perspective which suggests that the knowledge is constructed rather than received, and also, it is revealed through interaction rather than retrieved. The implications of the findings of classroom research on teachers' knowledge and practice suggest that SLA can be successfully effected through integrated and interactive framework. Many SLA interactional theories and researches are considerably influenced by psychology-oriented learning theories which focus on the need for modified interaction in language learning and acquisition. SLA also involves the role of display and referential questions for meaningful interaction in terms of the social discourse as well as the classroom discourse.

Though attempts to agreements have been concluded, and widely recognized definitions of CT presented, scholars have not agreed on a definition of CT consistently. There is adequate consensus, nevertheless, to practice research on procedures for developing CT abilities and applications. Many models and approaches to teaching CT have been improved, but few of them have been assessed empirically by unbiased researchers. Disagreements on the definition of CT have also weakened attempts to improve methods for evaluating CT.

Many standardized tests for CT are available; however, they usually have not been successful to account for good aspects of CT. Short open-ended answers and essay tests have

reflected many of the concerns about multiple choice tools, though these continue to be difficult to grade reliably. The amount to which learning of CT in particular areas transfers to other subjects and to daily reasoning has continued to be a cause of debate, but being obvious and reflecting alternatives for application seems to make the possibility of transfer more likely. In EFL, CT abilities are often constructed through the use of teaching activities and teachers' practice; however, the research studies and literature investigating teaching for CT in EFL is limited. More studies are required to identify which aids and strategies are required in developing learners' skills to think cognitively and critically in EFL contexts. More research is also required on explicit methods of teaching for CT so that achievement of effective learning will take place.

Higher order thinking strategies can actually be employed within any subject. The basic understanding of these strategies is to allow for different interpretations and discussing various "correct answers" as against "one right answer". It is not about finding one correct answer but discovering different interpretations for it and finding out that all or any of these may be acceptable. Many scholars described how the engagement of the learner and the teaching materials is a personal one making it possible for multiple interpretations since each learner brings different meanings into assigned teaching materials.

The role of the teacher is to motivate students to go beyond merely understanding the teaching materials and toward thoughtful interpretations. The teacher becomes a facilitator of improving the individual's capacity to evoke meaning from the lesson by leading him/her to reflect him/herself critically in this process. The teacher who conducts higher order strategies can play a vital role in helping students to make connections with the lesson by asking questions that will lead to thinking and reflecting. Thinking is only productive if it leads to thoughtful interpretations and the questions we ask determine where the thinking goes. Though Socratic Questioning is a fundamental base for this study, still we have to look at other questioning taxonomies and illustrate the importance of those questions in the light of their significance in language teaching and learning. On a separate note, social aspects of language use call for a

literature review that focus on the social and cultural dimension of language use as it might help in developing a multi-level perspective of the SQ and CT in Omani context. As Krishnaswami, Verma, and Nararajan (1992) argue, language and society are so intertwined that it is impossible to understand one without the other and there is no human society that does not depend upon, is not shaped by, and does not itself shape language (p. 13). Language as a sociocultural system is culture-preserving and culture-transmitting. Sapir-Whorf hypothesis (Whorf, 1956), based on the principle of linguistic relativity or linguistic determinism, relates our thought and conduct to language. Sapir and Whorf argue that we dissect nature along the lines laid down by our relative language. In this sense, as Scollon and Scollon (2010) remark, "human discourse is influenced by socially shared habits of thought, perception and behavior" (p. 538) because, to use Goddard and Mean's (2009) argument, when we acquire language, we acquire ways of thinking, i.e. "conceptual systems" (p. 2). Furthermore, according to Austin (1962) and Searle (1969, 1979), whenever we use language, we perform speech act(s) such as questioning, probing, answering, complaining, requesting, negotiating, and so on and so forth. In this sense the whole speech is simply an exchange of speech acts. Successful performance of speech acts depends on their felicity conditions. These felicity conditions are shaped by the socio-cultural norms of the society. The norms of solidarity-oriented society and formality-oriented society are likely to differ based on the power principle and solidarity principle of the two types of societies. Incidentally, Asian societies tend to be solidarity-oriented societies whereas European societies formality-oriented. According to Thakur's (2002) central argument Western socio-cultural considerations in communication are predominantly driven by formality-oriented texture of society whereas Asian socio-cultural considerations by solidarity-oriented social texture. Thus, our belief systems and the cognitive schemata are highly dynamic structures which are "constructed in and through discourse and can be continuously revised and updated" (Weber's 1998, p.116). In view of these socio-cultural considerations of language use House and Kasper (1981) argue that pragmatic approach would be very beneficial to minimize native cultural interference in interpreting the speaker utterances effectively or appropriately. After examining and discussing various aspects and strands of SQ and CT skills, based on the research studies, discussion papers and research projects, the next chapter in

sequence is devoted to the discussion of Research Design and Methodology of the present study.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The third chapter on the research methodology is devoted to a comprehensive discussion of the theoretical framework, and the detailed methodology evolved to the analytical concerns of the study. Descriptive explanations of the research design, details of the sample of the study, information related to the instruments used, a discussion on the validity and reliability of the instruments, the description of the procedures of data collection, the explanation of the analytical procedures to examine the data, and the discussion on the intervention stages have been included in this chapter.

The current study has adopted a “multi-method” (M. Teresa, Angel, Jose, Pedro, and Anthony, 2018) design which was conducted on 230 grade 12 female teachers, and 60 grade 12 female students in Dhofar region of Oman, who were divided into 30 students in the experimental group and the other 30 in the control group. Five data collection instruments were used in this study. These instruments were subjected to various relevant processes of ensuring validity and reliability. The required data for this study was collected using the relevant instruments at appropriate stages.

The collected data was analyzed using standard relevant statistical procedures as applicable in the two phases of the study. Phase I of the study identifies the relationship

between the teachers' knowledge of SQ and their actual use of SQ in their instructional practices in grade 12 EFL classroom through analyzing data collected through quantitative questionnaires administered on 230 grade 12 female teachers, and classroom observations of 33 grade 12 female teachers using classroom observation checklist. In Phase 2, three instruments were administered to collect data. CT test was conducted as a pre and post measure to examine students CT skills. Furthermore, four quizzes (Assessment Day, 2017) at the intervals of every two weeks, were conducted to gauge the gradual progression of students' CT skills, and the structured interviews were conducted with five students from the experimental group to triangulate students' responses with their results of the four quizzes. What follows next is a detailed discussion and explanation of different aspects of research methodology used in the study.

3.2 Research Design

This is a predominately multi-method study that was conducted in two phases where both quantitative and qualitative data (M. Teresa, Angel, Jose, Pedro, and Anthony, 2018) were employed. In phase I, the correlation research design was employed to collect descriptive quantitative data that helped to identify the relationship between teachers' knowledge of SQ and their actual use of SQ. The result of this phase helped to determine the nature of the intervention in phase II. The quantitative data in phase I helped the researcher in two ways; (a) to link the data that obtained from phase I and that obtained from the participants via quasi-experimental results in phase II to understand the application of teachers' knowledge of SQ on students' CT skills, (b) understand the reality of classroom context in Omani post-basic classes in relation to teaching methodologies and different pedagogical strategies related to classroom questions (see section 4.3.2). The conceptual model of the study in Figure (1) shows the relationship between the independent variable (SQ) and the dependent variable (CT). In addition, it illustrates the inter-relationship between the constructs of the main variables of the study.

In phase II of the current study the researcher used a quasi-experimental design. The reason behind choosing this specific design is that “quasi-experiments” include assignment, but not random assignment of participants to groups. This is because the experimenter cannot artificially create groups for the experiment (Creswell, 2012, p. 309). It is important to select the proper research design that “fits for the purpose of the study” (Gorard, 2002a, p. 354). The nature of an experimental method is that researchers “intentionally manipulate the settings prior to the intervention phase to maintain normality and homogeneity which determine an equal distribution of the sample (Creswell, 2012, p. 303)”.

In other words, correlation studies only define or measure the possible relationship between dependent and independent variables but cannot demonstrate the main cause-and-effect relationship between the two variables. However, an experimental design is a design that exposes experimental groups to a particular treatment and certain statistical procedures that could identify the cause-and-effect relationship between the two variables.

Since the current study investigated the application of SQ strategies in EFL reading classes on students’ CT skills, the researcher believes that conducting an experiment is the appropriate procedure for the current study to answer the target questions in this research. Two types of data were collected in phase I and II of the study, namely quantitative data that is related to research question one (RQ1) and research question two (RQ 2), and qualitative data that are relevant to research question three (RQ 3). Both types of data are discussed in details in chapters three and four. Incidentally, both the quantitative and qualitative data are presented, analyzed, and discussed at two levels. Both the results are first presented holistically including all the thirty participants of the experimental group in the study. After that, the two sets of data are also presented in its focused perspective which involves five students from the experimental group, who had volunteered to participate in the qualitative structured interviews. The results obtained from the holistic and focused perspectives are used to evolve a multi-level discussion of the findings.

3.3 Sample of the Study

In phase I of the current study the researcher used “purposive sampling technique”. This technique is considered as a non-probability sample where all subjects in the sample share same characteristics (Crossman, 2018). The sample size for phase I of the current study was 230 teachers, which equals the total population of the EFL teachers working in eight major post-basic grades 11 and 12 schools in Dhofar region. Before starting the implementation of the current study, the researcher had the permission to implement this study in the target schools related to the current sample (appendix 1). Phase I of the current study aimed to measure the relationship between Omani EFL teachers’ knowledge and use of SQ. All target participants were Omani female teachers who taught grade 12 students in post-basic schools. There were three reasons behind choosing only female teachers, as discussed below.

First, Omanization policy of replacing expatriate teachers with Omani teachers was highly successful with female teachers (Ministry of Education, Oman, 2008). This was due to the fact that male members of Omani community were not interested in the teaching jobs over their preference for jobs in other fields such as engineering, law, military, police, and other private sectors. Second, as the main target of the study is Omani teachers, all English teachers who taught in post-basic schools could not be considered as they were expatriates. Third, mixed gender classes were not allowed in the post-basic schools in Oman. Mixed gender classes were allowed only in cycle 1 schools which includes classes from 1 to 4. Therefore, the findings of the current study could not be generalized to all post-basic Omani teachers.

Those teachers who participated in this study taught English language reading skills in the academic year 2016/2017 and came from same educational backgrounds. Important measures were taken to ensure that all teachers who participated in the study did not have any major individual differences that might possibly affect the result of the study. All teachers who participated in the study have Bachelor Degree in teaching English as a foreign language. In addition, they had the same range of teaching experience which is more than 5 years. All of them had well informed awareness of Omani social life, and the local language. The local

language is called Al Shahri. This language is not totally Arabic. It is a Sematic-Omani language and it is considered as a first language in most of Dhofar areas.

Those teachers who taught English language reading skills in the academic year (2016/2017) answered the questionnaire meant to measure their knowledge of SQ. In addition, Classroom observations were conducted with 33volunteered respondents from the same sample of teachers who completed the questionnaire as illustrated in Table 3.1 below. This was because not all teachers were happy to have observations in their classes, only 33 teachers agreed to have observations in their classes. This was to examine the actual use of SQ inside the classroom by using classroom observation checklist. The researcher considered the 33 teachers, whose classes were observed as representative sample of Dhofar region as well as of the other regions of Oman. This is because the context and the circumstances of teaching and learning in schools across Oman are more or less the same. Thus, with a reasonable degree of variation, the results of this study are likely to provide useful information and insights for the benefit of English language teachers of post-basic schools of Salalah in particular and Oman in general regarding the knowledge and effective use of SQ and CT skills in grade 12 EFL classrooms.

Table 3.1: The Schools and the Volunteered Teachers Participated in the Interviews

No	Number of the School	Volunteered Teacher (T)
1	School 1 Post-Basic School	T9, T20, , T84, T153, & T176
2	School 2 Post-Basic School	T1, T7, T119, T33, T78, & T222
3	School 3 Post-Basic School	T27, T67, & T209
4	School 4 Post-Basic School	T7, & T11
5	School 5 Post-Basic School	T94, T25, T120, & T70
6	School 6 Post-Basic School	T5, T19, & T93
7	School 7 Post-Basic School	T8, T21, T100, & T219
8	School 8 Post-Basic School	T41, T60, &T124

In phase II of the current study which was quasi-experimental design, a random school was selected from the list of the eight schools included for the purpose of the study. The schools were listed alphabetically and the first school was selected for the purpose of the study. Again, two grade 12 post-basic classes with the same teacher were assigned to act as a control group and experimental group. The total number of students in each class was 30. Their age was between 16-18 years old, and all of them use the same course book. Group A was the control group and group B was the experimental group.

3.4 The Use of Multi-Method Data

This study is driven by a multi-method research approach. It integrates quantitative and qualitative data to research design in the form of triangulation, however, mixed method design is when data is in sequential order (Teddlie and Tashakkori, 2009). Teresa, Angel, Jose, Pedro, and Anthony (2018) conclude that both quantitative and qualitative can be accrued in different stages of the research, which can be indicated in the planning phase, research questions, research instruments, and analysis lead to research findings.

Before the researcher discusses the quantitative and qualitative data related to RQs 2 and 3, it would be in the fitness of the study to briefly talk about the instruments related to (RQ1). For the purpose of RQ1, two instruments had been used, which are (a) a quantitative questionnaire to measure teachers' knowledge of SQ; and (b) classroom observation-based checklist to account for teachers' actual use of SQ. These instruments are discussed in details in forth coming sections. What follows is the discussion of the use of multi-method approach.

The current study includes two phases. Quantitative design in phase I was considered as an important platform that determined the nature of the intervention that was implemented and used in phase II as a treatment. In other words, the findings of phase I in this study identified the relation between teachers' knowledge of SQ and actual use of SQ strategies. Thus, the findings of phase I formed the bases for intervention in phase II of this study. As a

result, the intervention helped to identify the result of applying teachers' knowledge of SQ on students' CT skills. Therefore, in phase II, three instruments were used to collect the required data. First, a T-test was conducted to measure the statistical significant differences between mean scores in pre-test and post-test relevant to the research question 2 (RQ2). Second, the scores of quizzes conducted four times during the intervention phase were recorded. CT quiz scores were tabulated holistically in one table to count the means and total average of CT skills. Furthermore, the quiz scores were also tabulated separately in terms of individual five CT skills along with the total mean score of each CT skill. Third, structured interviews were conducted on the experimental group as a qualitative data instrument in order to triangulate the effect of the intervention phase II result. The qualitative data collected in the form of interview responses were recorded, and transcribed, after that, the emerging themes from the responses were identified, coded, and duly categorized. The qualitative data and the quantitative data collected through the results of quizzes and interviews were analyzed to answer research question 3 (RQ3).

3.5 Intervention: In Situ Vs SQ Method

The classification of SQ questioning strategies developed by Paul and Elder (2006) is not ordered and organized in a rigged traditional manner. These categories (elements) in the model proposed by Paul are flexible, which allows teachers to select questions (elements) randomly based on the category of questions relevant to the target skill they want to develop. The task of successful and effective teacher is to sustain and maintain the process of inquiry. Thus, it is very important to illustrate the difference between the teaching strategies and teaching materials for both groups.

The subjects in both groups, i.e. the control and the experimental group used and studied the same course materials and reading texts published by the Ministry of Education for grade 12 as indicated in (appendix 3). The students regularly met with the teacher for eight weeks. The control group was taught normally using IN SITU method which reflects the teaching of the control group that employed existing textbooks, usual learning tasks, and other

materials. This method mainly focused on three reading strategies, surveying the reading text, skimming the reading text, and scanning the reading text. This method depends on mechanical reading and the role of students is limited to just pick up the correct answers according to the given questions for each reading text. By the end of each reading strategy the teachers will select the correct answers, because the focus was on mechanical reading and identifying the correct answers which reflects product-based learning. On the other hand, the experimental group was taught through the SQ Framework (SQF). SQF implemented the same reading strategies, which have been used through IN SITU method; however, the role of the teachers is to sustain and maintain the process of inquiry. Obviously, the focus in the experimental group was on using SQ for each reading strategy and evoke oral discussion related to the selected answers.

Students should engage in CT process and discussion (process-based learning) that reflect the five CT skills which had been introduced to them at the beginning of the intervention program. To elaborate further, the intervention included 8 reading topics (appendix 3). Each topic was taught and covered the five CT skills in one week (5 hours). SQ strategies used as teaching strategies in each reading topic which cover the five CT skills included in the questionnaire and observation checklist. The nature of SQ strategies is based on questions that evoke students' CT. This helps the teacher and the students to exchange ideas at higher cognitive level through authentic discussions in the lesson. The questions in Table 3.2 below used for teaching, are examples of SQ which reflect the core of SQ strategies.

In terms of practicality, it is not necessarily that the teacher would be able to use all the categories and questions listed in the below table in each reading lesson. The teacher will target the type of relevant questions to evoke students CT in terms of the demands of reflection the reading text pauses on the students.

**Table 3.2: Examples of Socratic Questions
Paul and Elder (2006)**

No.	Category	Examples
1	Questions for clarification	Why do you say that? What do you mean by...? How does this relate to our discussion?
2	Questions that probe assumptions	What could we assume instead? How can you verify or disapprove that assumption? On what basis do we think this way?
3	Questions that probe reasons and evidence	What would be an example? What is....analogous to? What do you think causes to happen...? Why?
4	Questions about viewpoints and perspectives	What would be an alternative? What is another way to look at it? Why is the best?
5	Questions that probe implications and consequences	What generalizations can you make? What are the consequences of that assumption? What are you implying?

Adapted from:

https://www.kpu.ca/sites/default/files/Learning%20Centres/Think_Critical_LA.pdf

The 8 reading texts along with lesson plans, students' tasks, activities, and related assessment procedures to be used in the intervention phase of 8 weeks were constructed and designed according to ADDIE's model in (Peterson, 2003). The proposed teaching materials were reviewed for its content validity through two major sources. First, these documents were

presented to and discussed with five EFL language supervisors and 13 EFL senior teachers in eight workshop sessions that were conducted by the Ministry of Education in Dhofar region.

Detailed feedback was collected from the participants through face to face discussion and by recording the discussion in the form of diary notes. Second, the intervention program documents were sent to three experienced reviewers. The first reviewer was the Head of Education Department in Dhofar University, the second reviewer was the chief supervisor of the Supervision Department in the Ministry of Education, and, the third reviewer was an associate professor in Languages and Translation Department in Dhofar University. The feedback collected from the two sources was utilized for further improvement.

The following suggestions from the feedback were implemented to enhance reliability and validity.

1. Some reviewers suggested to increase the intervention period from five weeks to eight weeks, in order to get a reasonably comprehensive data to meet the demands of the study.
2. In response to the researcher's decision to use five units of reading texts from the prescribed textbooks, some reviewers suggested eight units according to the eight weeks' duration of the program.
3. It was suggested that, in addition to using the only WGCT-FA test, four quizzes should be conducted at the rate of one quiz after the completion of every two weeks in order to measure students' progress of achievement on a continuous basis.
4. As suggested by some reviewers, teacher training was crucial in this study.

The teacher who taught the two groups and participated in this study was exposed to a comprehensive training plan. First, the teacher was given three induction sessions in the form of face to face meetings as introduction to SQ strategies. Second, the same teacher was also involved in the eight content validity workshops with the Ministry of education EFL supervisors and teachers on intervention. These workshop sessions enabled the teacher to develop clear

understanding and comfortable level of confidence in exploiting SQ strategies in teaching practices with the experimental group.

Table 3.3: Summary of the 8 Weeks Intervention

No.	Lesson Presented	Duration	Assessment
1	Induction week	Week 1	Oral discussion and illustrations CT skills. Introducing the textbook.
2	Readings 1 & 2	Week 2	Quiz 1: reading using 5 critical thinking skills
		Week 3	
3	Readings 3 & 4	Week 4	Quiz 2: reading using 5 critical thinking skills
		Week 5	
4	Readings 5 & 6	Week 6	Quiz 3: reading using 5 critical thinking skills
		Week 7	
5	Readings 7 & 8	Week 8	Quiz 4: reading using 5 critical thinking skills
		Week 9	

3.6 The Proposed Lesson Plan

The following prototype general lesson plan guide, as shown in Table 3.4, reflects the specific steps for each reading lesson. Each reading lesson contains four tasks. Every week the teacher is expected to plan and use five CT skills. However, the lesson plans have some variations according to the demands of individual reading text.

Table 3.4: The General Daily Lesson Plan Guide

Teachers' Actions	Taught CT Skills	Students' Actions
(Task 1): Pre-reading to read: Teacher asks some general questions that evoke students' background knowledge about the topic (SQ).	Inference	Students bring their knowledge, experience, and opinions about the topic. This will help students to relate the topic to their own context.
(Task 2): While-reading: Teacher asks some specific questions related to specific details about the topic (SQ).	Recognition of Assumptions	Students read the article and match the main ideas to the paragraphs. Students exchange with the teacher specific details and ideas about the topic according to the types of questions that teacher asks(SQ).
(Task 3): While-reading: Reading for the Main idea. Teacher directs students to read the beginning of each paragraph to understand the main idea. Teacher explains the rubric and the written questions to students (SQ).	Deduction and Interpretation	Students read the first sentence or two of each paragraph to understand the main idea. Students answer questions about the main idea (SQ).
(Task 4): Discussion: Teacher introduces two questions for discussion (SQ).	Evaluation of Arguments	Students work with partners or individually to answer and discuss the questions (SQ)

The general lesson plan guide (developed by the researcher)

3.7 Research Instruments

The description of multi-method data, the current study had two phases in which five different instruments were used. In phase I, two instruments were employed. First, a questionnaire was used to identify teachers' knowledge of SQ. Second, observation checklist aimed to measure teachers' use of SQ. Both the quantitative instruments were adapted from (Paul and Elder,

2006). In phase II, Watson-Glaser Critical Thinking Appraisal Test (2006) (WGCTA-AS) was used to measure the effect of SQ strategies on students' CT skills. Also, quizzes and interviews were used to measure the CT strategies developed by students' as a result of intervention implementation in phase II.

3.7.1 Phase I: The Pilot Study, Instruments, and Procedural Aspects

This section of the study deals with a detailed description and explanation of the process of piloting the questionnaire instrument that was used in this phase of the study. The questionnaire is developed based on Paul and Elder's (2006) five CT domains as explained in detailed in section 3.7.8. This section also discusses the expected procedures to check and control the possible threat factors which might affect the reliability and the validity of the instrument adversely. This section also includes the discussion of further procedural measures in order to strengthen different aspects of reliability and validity of the questionnaire.

There were three reasons behind piloting the questionnaire of the current study. First, as stated by Seliger and Shohamy (1998), the aim of piloting the questionnaire is to check the quality of it in order to improve the quality of the items and make it more comprehensive to collect the data. On a similar note, other researchers believed that piloting is important in detecting complications or problems so that the questionnaire variables can be developed and improved before implementing the main study (Fraenkel and Wallen, 2000). Second, piloting questionnaire can familiarize the researcher with the practical applications during administration of the data collection process. Third, a pilot assessment is associated with the procedures of validity and reliability (Seliger and Shohamy, 1998). Piloting the questionnaire of the current study enabled the researcher to find possible gaps and problems, thus, the researcher modified the needed improvements in the questionnaire to meet the required level of reliability and validity of the questionnaire (Creswell, 2012).

The participants in the pilot assessment were 130 EFL teachers. Those teachers taught English to Omani students in Dhofar region "post-basic schools" (Grade 12 students). The

students and teachers who participated in the pilot study shared similar circumstances and backgrounds of those who participated in the main study. The sample of teachers was selected randomly among the post-basic schools in the region. The researcher would like to clarify at this point that the selected sample of 130 EFL Omani teachers was not used as a sample of the data collection of the main study. Those were two different groups of EFL teachers.

The questionnaire in this study was adapted to find out teachers' knowledge of SQ (appendix 4). It consisted of twenty-nine items that are classified into five related dimensions. The items were developed from the related Paul and Elder's (2006) checklist. The dimensions included were the same ones which were used in both the questionnaire and the checklist. However, items in certain dimensions of the questionnaire were simplified in order to help the target teachers (participants) to understand the nature of the questionnaire. Also, examples are given for each item (question) to help the target teachers (participants) to understand and clarify the purpose of each item.

The dimensions are as follows: knowledge of clarification questions, knowledge of questions that probe purpose, knowledge of questions that probe students' expectations, knowledge of questions that probe information, reasons, and evidence, and questions that probe implications and consequences. Obviously, as Paul and Elder (2006) remark, there was no classified taxonomy of SQ, however, each dimension represents an area in which the teacher can question students. For example, the teacher can ask students how could they arrive at their particular point of view. They can also inquire into what evidence they gathered to support their conclusions. The teacher can use questions to identify other points of view and determines the implications and consequences of a particular conclusion based on what is known about the topic and the other points of view. Many topics taught in English language classes lend themselves well to teaching using SQ.

The twenty-nine items in the questionnaire were planned on a five-point Likert scale to get the respondents' level of knowledge of SQ. The answer scale was as follows: 5= Strongly Know, 4 = Know, 3 = Not Sure (neither Know nor Don't Know), 2 = Don't Know, and 1 = Strongly Don't Know (Cambridge University Press, 2011, p. 43). The researcher used the weighted mean to answer research question 1. This is because the target sample related to research question 1 is large and has a wide range of opinions. "When there is a range of opinions or views on a particular matter, we tend to adopt the view that we consider to carry the greatest 'weight' (Cambridge University Press, 2011, p. 43).

Table 3.5: Weighted Mean Values

Very Low	Low	Medium	High	Very High
1 – 1.79	1.80 – 2.59	2.60 – 3.39	3.40 – 4.19	4.20 – 5

The weight normally reflects the degree of authority based on expertise and experience that supports the view" (Cambridge University Press, 2011, p. 43). Table 3.5 illustrates the weighted mean values that correspond with the five-point Likert scale used in the current study. The researcher distributed questionnaires and collected the data through the Department of Teachers Training and Professional Development, Ministry of Education. The Department of Teachers Training and Professional Development had sent and collected the questionnaires to target participants in different schools they worked in. The Department of Teachers Training and Professional Development send and collect the completed questionnaires from the target participants within three weeks' time. The researcher's e-mail was attached to the questionnaires if target teachers needed any further clarifications.

3.7.2 Screening of the Questionnaire

For this study, the researcher applied screening methods to identify respondents who could not put in enough effort when responding to questionnaires. Usually, respondents to questionnaire have different attention and focus when responding to questionnaires items. There were three

techniques researchers might apply to find out respondents who fail to apply adequate attention and effort in order to increase the objectivity of analysis and improve the credibility of study findings. Researchers classified screening techniques into three types Justin, DeSimone, Harms, and Alice (2015) direct, archival, and statistical. Direct screening techniques are those techniques that need to include extra items into a questionnaire, which help researchers to assess participants on the basis of certain responses made while completing the questionnaire. Archival screening techniques do not involve any change in the questionnaire, but emphasizes particular responses and behavior while respondents are completing the questionnaire. Finally, statistical screening techniques need no changes in the questionnaire and depend on statistical procedures to spot unusual responses.

In the current study the researcher used the first screening technique which is a direct technique. Direct screening technique includes items (30 and 31) into a questionnaire before administration. In direct screening participants are likely to be aware of the function of these included items. The knowledge of inserted items could motivate participants to avoid dealing with the items in negative mode. There are three different procedures in direct technique, Self-report indices of data quality, Instructed items, and Bogus items (Justin, DeSimone, Harms, and Alice, 2015).

In this study the researcher applied the Self-report indices of data quality and instructed items procedures for both levels of pre and post implementation of the questionnaire. In terms of self-report indices of data quality, one item was added at the end of the questionnaire which was number 31 to make sure that the respondents were aware enough and focused while they were completing the questionnaire. Also, in terms of instructed item, two items were added to the questionnaire which is items number 30 and 31. This was due to the reason that it targeted respondents who answered the questionnaire items with low effort and it raised participants' awareness towards answering the questionnaire item because self-report indices were inserted before administration of the questionnaires. Therefore, the target indices were not included and analyzed in the statistical procedures of the current study.

3.7.3 Missing Data

Tabachnick and Fidell (1996) suggest that the situation of missing data can be repaired. However, it depends on whether data is “missing randomly” or “if there is a reason/pattern” that explains why the data scores are missing. This can be identified by looking at the data set. The researcher in the current study found two missing values in the questionnaire which were for items 1 and 17. Each missing value was replaced by the average of the same item scores of the total sample.

3.7.4 Normality Test

Normality test was conducted to check normality distribution of the sample. Skewness and Kurtosis values were calculated for the pilot study. Skewness value was -0.26, and the Kurtosis value was -0.37. These results indicate normal distribution of the sample in the pilot study. In addition, Skewness value (-0.26) was divided by the standard error value of the Skewness ($SE = 0.21$) which resulted in (-1.22) and the value of normal distribution ranged between (1.96 to -1.96). Also, Kurtosis value (-0.37) was divided by the standard error value of Kurtosis ($SE = 0.42$) which resulted in (-0.87) and the value of normal distribution ranged between (1.96 to -1.96), which is considered a reasonable distribution data of the pilot study (Tabachnick and Fidell, 1996). Figure 3.1 shows the distribution of the data set which is close to curve shape.

Figure

3.1: Normality Distribution of the Sample

In addition, in the pre- study stage the researcher used Kolmogorov-Smirnov test and Shapiro Wilk test in Table 3.6 to assess the normality distribution of data in the independent variable (teachers' knowledge of SQ-questionnaire). The result indicated that the values of Kolmogorov-Smirnov and Shapiro Wilk test were not significant which reflected a normal distribution of variables. Again, the values were within the range of what is considered a reasonable approximation to the normal curve.

Table 3.6: The Distribution of Data in the Independent Variable

Independent Variable	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	Df.	Sig.	Statistic	Df.	Sig.
Questionnaire total	.072	130	.179	.983	130	.116

3.7.5 Outliers Assessment

Outliers are the responses or the scores that affect the result and interpretation of the study. In addition, the statistical outcomes of analysis will be affected by outliers. To avoid the negative impact of outliers, the target data must be tested and checked (Tabachnick and Fidell, 1996). It is clear that, Figure 3.2 shows the outlier in the data set before detection. As a result, Figure 3.3 illustrates the distribution of data in a curve shape before deduction. However, Figure 3.4 shows the outliers in the data set after detection. As a result, Figure 3.5 illustrates the distribution of data in a curve shape after deduction.

To illustrate further, there was only one extreme value that reflects case number (95) in Figure 3.2 with an extreme value (87.00) as distributed in Figure 3.3. However, the extreme value was detected and removed from the data set as indicated in Figures 3.4 and 3.5.

Figure 3.2: Outliers before Detection I

Figure 3.3: Distribution of the Outliers before Detection II

Figure 3.4: Outliers After Detection

Figure 3.5 Distribution of the Outliers after Detection

3.7.6 Non-Responses Bias Test

Non-response bias occurs when there is a clear difference among those who answered the questionnaire and those who did not respond to the questionnaire. This may happen for a variety of reasons, including: The aim of non-response bias test is to achieve a representative sample by minimizing non-response bias as much as possible. In the current study non-response bias was (0). Therefore, there was no non-response bias which indicates that the sample used in this study was representative of the population.

3.7.7 Reliability and Validity of the Questionnaire

Reliability reflects confirmation if the questionnaire is gathering data in a consistent way (Seliger and Shohamy, 1989). Pedhazur and Schmelkin (1991) stated that Cronbach's alpha reflects a suitable statistical procedure that measures internal consistency by assessing the degree of

reliability of items in the questionnaire. Also, it was further explained by the two authors that Cronbach's alpha will enable researchers to assess relationships between the items of the questionnaire.

In order to confirm the reliability of the questionnaire for the current study concerning the Omani teachers' knowledge of Socratic questioning, the researcher used Cronbach's alpha to ensure the required reliability coefficient of each item of the questionnaire in the pilot study. At the beginning of the pilot study and before any changes were made to the questionnaire Cronbach's alpha was 0.77 which was considered appropriate coefficient. The representative benchmark for measuring this coefficient is between 0.60 and 0.70 (Hair, et al., 1995).

Brown (1988) and Seliger and Shohamy (1989) stated that validity indicates that instrument or questionnaire should measure what it is intended to be measured. Therefore, validity is considered as important as reliability in assessing a new instrument. Content validity is another important indicator of instrument validation. This can be achieved by distributing the questionnaire and the observation checklist to reviewers who are familiar with the content knowledge of subject matter (Fraenkel and Wallen, 2000). Following this guideline, the questionnaire and the observation were shown to a group of 5 referees, further information about the referees included in (appendix 2). These referees were asked to comment on both instrument and to indicate any suggestions for improvement. The aim of this was to ensure the maximum validity for this instrument. The questionnaire and the observation were subjected to some modifications, following the feedback obtained from the referees and the instruments were ready for pilot assessments. Some of the important suggestions for modifications include deleting certain items that were out of context; e.g. questions that are related to "display" and "close" questions. This type of questions does not enhance students' CT skills, however, it can be answered only by "yes" or "No". Another comment by referees is to provide explanation for the strategies related to each domain, e.g. referees suggested to add examples of questions related to each item in the questionnaire and observation checklist (see appendix 6). All recommendations by referees were met and implemented.

3.7.8 Construct Validity of the Questionnaire

Construct validity is considered the most significant statistical indicator during validity assessment. Researchers used construct validity when they want to assess whether the data gathering procedures are appropriate and representative of the variables that are intended to be measured (Seliger and Shohamy, 1989). In the current study the researcher conducted Exploratory Factor Analysis to ensure maximum validity to the questionnaire. When the researcher received back the completed questionnaires, he coded the measured data and documented the participants' responses in a soft copy file that uploaded to SPSS 11.0 version.

The researcher conducted a Factor Analysis using the Principal Component method. The purpose of this analysis was to ensure a higher level of construct validity of the questionnaire. The Kaiser-Meyer-Olkin (KMO) value was (0.59) which was greater than 0.5 that indicates the sample size is appropriate. Also, Bartlett's Test of Sphericity value was (0.001) which is significant at the level (0.01) which means that at least two questions were correlated. Following the application of the Varimax Rotation method to the questionnaire items (29 items); The five factors were renamed based on the results that emerged from the Factor Analysis. Guilford standard (3.0) was used to test the saturation function. In case of saturation of one phrase on two factors with two distant values, the greatest saturation was considered while factors with three item plus saturation of a minimum value of 3.0 were retained (appendix 5). This ensures a better level of purity for factors.

The questionnaire items were, therefore, invariably maintained and measured only five factors that interpreted that were (42.52) from the overall variance of the sample. The Tables numbered 3.7 to 3.11 below provide domain wise descriptions of the factors.

Factor 1

Table 3.7: The Saturation Values for Factor One Knowledge of Probing and Illustrating Questions

No	Item	Value
1	Questions that track relevant information. e.g. What do you mean by ____?	0.71
8	Questions that help students to be aware of the topic. e.g. How do the purposes of these two groups vary--?	0.52
9	Questions that clarify the target information. e.g. Was this purpose clear?	0.60
11	Questions that make students answer relevant to the question being addressed at any given point. e.g. Why do you say that?	0.38
14	Questions that question understanding when appropriate or significant. e.g. You seem to be assuming <u>X</u> . How do you justify taking that for granted?	0.59
16	Questions that help the students to focus on the main idea. e.g. All of your reasoning depends on the idea that _____. Why have you based your reasoning on <u>X</u> instead of <u>Y</u> ?	0.33
17	Questions that question conclusions. e.g. How did you reach to conclusion that it applies to this case?	0.53
22	Questions that encourage participants to think aloud. e.g. Why do you say that?	0.54

25	Questions that ask participants to make overviews of the topic . e.g. What overviews can you make about what you learnt?	0.45
The eigenvalue is (4.32) and the rate of variance is (14.91%)		

Table 3.7 shows the saturated items in the first factor which was knowledge of clarification questions. It is clear from the above table that the underlying eigenvalue of this factor amounted to (4.32) and interpreter factorial variance ratio is (14.91%) and the items saturation values ranged between (0.33) and (0.71): this factor saturated with 9 items, and all saturation values are significant and positive. The items of this factor reflect the nature of teacher questions that can help students to understand the topic that has been presented. Based on this result, the first factor can be renamed as knowledge of questions probing and illustrating topics.

Factor 2

**Table 3.8: The Saturation Values for Factor Two
Knowledge of Analytical Questions**

No.	Item	Value
2	Questions that reflect main ideas or concepts. e.g. How does <u>X</u> relate to <u>Y</u> ?	0.62
5	Questions that help students to think of other options to think about the problem. e.g. How does this relate to our problem/discussion/issue?	0.54
13	Questions that question interpretations when appropriate or significant. e.g. You seem to be assuming _____. Do I understand you correctly?	0.40

18	Questions that distinguish subjective questions from factual questions, from those requiring reasoned judgments within conflicting view points. e.g. What are your reasons justifications for saying that?	0.67
26	Questions that ask participants to reflect on main ideas of the assumption. e.g. What are the main ideas of that assumption?	0.48
29	Questions that tie in with facts that were covered earlier. e.g. How does the X information connect with what we learned before?	0.74
The eigenvalue is (2.46) and the rate of variance is (8.49 %)		

Table 3.8 shows the saturated items in the second factor which was knowledge of questions that probe purpose. It is clear from the above table that the underlying eigenvalue of this factor amounted to (2.46) and interpreter factorial variance ratio is (8.49 %) and the items saturation values ranged between (0.40) and (0.74). This factor has been saturated with 6 items, and all saturation values are significant and positive. The items of this factor reflect the nature of teacher questions that can help students to analyze and have a deep cognition about the topic presented. Based on this result, the second factor can be renamed as knowledge of analytical questions.

Factor 3

**Table 3.9: The Saturation Values for Factor Three
Knowledge of Recognition of Assumptions**

No.	Item	Value
4	Questions that ask for more explanation of context when necessary. e.g. Could you put that another way, using the background information?	0.80
20	Questions that call for more precision when necessary. e.g. what the most important points in the text?	0.58

24	Questions that carry the discussion smoothly with the different contributions being effectively blended into an intelligible whole. e.g. What made you draw this conclusion?	0.41
28	Questions that focus on cause and effect from participants' point of view. e.g. How does X affect Y in this text?	0.76
The eigenvalue is (2.14) and the rate of variance is (7.37 %)		

Table 3.9 shows the saturated items in the third factor which was knowledge of questions that probe expectations. It is clear from the above table that the underlying eigenvalue of this factor amounted to (2.14), and interpreter factorial variance ratio was (7.37%), and the items saturation values ranged between (0.41) to (0.80). This factor has been saturated with 4 items, and all saturation values are significant and positive. The items of this factor reflect the nature of teacher questions that can help students to recognize assumptions related to the main idea. Based on this result, the third factor can be renamed as knowledge of questions that promote recognition of assumptions.

Factor 4

Table 3.10 : The Saturation Values for Factor Four Knowledge of Deduction Questions

No.	Item	Value
3	Questions that help students to focus on the point of view. e.g. What is your understanding about X?	0.78
6	Questions that help students to be aware of the difficulties in question on the floor. e.g. Let me see if I understood you correctly; did you mean X or Y?	0.39
7	Questions that deduct the goal of discussion. e.g. What is the main idea of the text?	0.31

10	Questions that help students to focus on the question being asked. e.g. What is the purpose of this question?	0.45
19	Questions that call for more details when necessary. e.g. Can you provide more explanations?	0.49
The eigenvalue is (1.74) and the rate of variance is (5.98 %)		

Table 3.10 shows the saturated items in the fourth factor which was knowledge of questions that probe information, reasons, and evidence. It is clear from the above table that the underlying eigenvalue of this factor amounted to (1.74) and interpreter factorial variance ratio was (5.98%), and the items saturation values ranged between (0.31) and (0.78). This factor has been saturated with 4 items, and all saturation values are significant and positive. The items of this factor reflect the nature of teacher questions that can help students to deduct information and knowledge related to the topic. Based on this result, the fourth factor can be renamed as knowledge of questions that deduct information.

Factor 5

**Table 3.11: The Saturation Values for Factor Five
Knowledge of Evaluative Questions**

No.	Item	Value
12	Questions that check students' inferences when appropriate or significant. e.g. What is the real meaning of this or that?	0.48
15	Questions that note questionable assumptions. e.g. are you sure this is the right conclusion?	0.42
21	Questions that evaluate participants' deep issues. e.g. What made you take that position?	0.52
23	Questions that periodically summarize when the lesson is in accomplishing stage. e.g. What is the main theme in the lesson?	0.61
27	Questions that help participants to reflect on the implications. e.g. How are the ideas expressed in text relevant to society?	0.52
The eigenvalue is (1.67) and the rate of variance is (5.76 %)		

Table 3.11 shows the saturated items in the fifth factor which was knowledge of evaluative questions. It is clear from the above table that the underlying eigenvalue of this factor amounted to (1.67), and interpreter factorial variance ratio was (5.76%), and the items saturation values ranged between (0.42) and (0.61). This factor has been saturated with 5 items, and all saturation values are significant and positive. The items of this factor reflect the nature of teacher questions that can help teachers to assess and evaluate students' information and knowledge related to the topic. Based on this result, the fifth factor can be renamed as knowledge of evaluative questions.

3.7.9 Questionnaire Modifications for the Main Study

After piloting and assessing process to questionnaire, the researcher produced the final form of the questionnaire that was used in the main study. (appendix 6). The final form of the questionnaire combined some modifications to the items under each factor by modifying and improving the items. Table 3.11 shows the items and factors of the modified questionnaire that was implemented in the main study.

Table 3.12: The Final Structure of the Questionnaire for the Main Study

No.	Factor	Items
1	Knowledge of Probing and Illustrating Questions	1,8,9,11,14,16,17,22,25
2	Knowledge of Analytical Questions	2,5,13,18,26,29
3	Knowledge of Recognition of Assumptions	4,20,24,28
4	Knowledge of Deduction Questions	3,6,7,10,19
5	Knowledge of Evaluative Questions	12,15,21,23,27

A final validation to the questionnaire conducted to make sure that the revised items and factors were practically measured what was planned to be measured. The researcher mixed up the items and the five factors of the questionnaire and put them in random order. Two colleagues were requested to validate the reviewed questionnaire by asking them to distribute the items according to each factor.

Factors were identified and after that the items were matched to each factor. Therefore, the researcher was able to validate the questionnaire and confirm questionnaire validation.

3.7.10 Classroom Observation Checklist

Foster (1996) explained that observation records the behavior of people that is watched in situations to collect data about the phenomenon of interest. Classroom observation is

implemented for different reasons. Many teachers claim knowledge, however, they do not apply it in real classroom settings. Therefore, classroom observation becomes an important method of research to collect information and evidence of teachers' claims (Foster, 1996). In addition, Foster believed that observations provide specific information that cannot be provided by other data collection tools, this kind of data is considered more likely precise and detailed which will not have relied on what participants claim about their classroom acts through interviews or questionnaires.

The current study investigated the result of applying the Omani teachers' knowledge of Socratic questioning on students' critical thinking in post-basic schools. One investigated domain in this study that was covered in phase 1 was the relationship between Omani teachers' knowledge of Socratic questioning and the actual use of this knowledge in classroom context. Therefore, teachers' knowledge was measured through a questionnaire and teachers' applications measured through classroom observation checklist.

Consequently, it was important to implement classroom observations in this study. Besides, the reviewed literature indicated a mismatch between theoretical and practical aspects of SQ in the teaching context (Moosa, Suyansah, and WardatulAkmam, 2015).

The observation checklist was used to collect data on teachers' actual use of SQ inside classroom. This instrument required a systematic observation process of teachers' use of questions in actual teaching situation. It was adapted from (Paul and Elder, 2006) checklist. The checklist was adapted on the basis of the related literature and the nature of this study (appendix 7). The items in the checklist were classified into five major factors which were similar to the factors and items in the questionnaire, and consisted of twenty-nine items. The checklist followed frequencies and tallies of SQ. The purpose of using this response mode is to facilitate the subsequent tabulation and analysis of data. Each teacher was observed one time and the lessons were audio-taped, which gave quick access to clarify ambiguous questions inside the classroom. Each lesson lasted 35 to 40 minutes.

3.7.11 Reliability and Validity of the Classroom Observation Checklist

Many researchers considered reliability is a very important aspect in classroom observations. Seliger and Shohamy (1989) and Allwright and Bailey (1991) pointed out, reliability provides consistent and stable data which reflect meaningful interpretation and result. Therefore, in this study, the researcher implemented two procedures in data gathering and analysis to maximize the level of reliability in the observation checklist. The two procedures were inter-rater reliability and data cross-checking procedure. Before conducting classroom observations, the researcher and another English senior lecturer had three trial observations followed by detailed discussions, in order to clear the ambiguous items in the checklist and familiarize the other senior lecturer with the observation context. The researcher proceeded to pilot the observation checklist. The researcher and the senior lecturer visited 10 teachers in their classrooms together. The aim behind this was to establish the inter-rater reliability of the observation checklist. The two observers sat through the same lesson, at the same time and utilized the same observation checklist. Cohen's Kappa statistical procedure was used to test the degree of agreement between the two raters. The inter-rater reliability coefficient for the observers was ($r = .70$), which is considered good level of agreement between the observers (Fleiss, 1981).

Additionally, the researcher cross-checked the observation data with the yielded questionnaire data, which were about Omani EFL teachers' knowledge and actual use of SQ questioning in classroom applications. The researcher aimed behind this methodological triangulation in order to find support and a desired positive matching between the two data collection methods (Seliger and Shohamy, 1989).

Validity reflects the judgements about what was observed, measured, or explained (Mason, 1996). As a result, Creswell (2012) stated that the observational tool and result drawn from it should be expressive, appropriate, and significant. Also, it has to be practical and meaningful to further studies. Beretta (1986) listed three principles that observation checklist should encounter. The First principle is; observation checklist should provide a precise

documentation of what was delivered. The second principle is, that observed data should be relevant to the target event. The third principle is, the observed data should be complete and reflect the whole target event.

In order to meet the above three principles, the researcher, concentrated on what was taught which was related to teachers' use and application of SQ in classroom in relation data construction or generation method, data analysis, and data interpretation. By data construction and generation, the researcher thought that the perceived data sources and data methods should possibly indicate how well they can ensure this (Mason, 1996). Definitely, this involved a logic interpretation whether the method reflects the perceived data to the research question posed. Therefore, the target observations are based on how teachers' questions in classroom was understood by Omani EFL teachers in classroom teaching context. This consideration added a clear and valid understanding of using classroom observation checklist.

Furthermore, validity includes questioning how effective the data analysis is and the explanation of the underlying principles of justified data (Mason, 1996). Therefore, the researcher justified the basis of the observation checklist that had been used (appendix 8). In the current study on Paul and Elder (2006) and interpretations were made to enhance the validity.

3.7.12 Preliminary Analysis of Correlation

Before applying correlation analysis, it was very important to implement some statistical procedures that will allow the researcher to check for violation assumptions related to the study. Assumptions like outlier detection, normality, linearity and, homoscedasticity are considered serious assumptions that must be checked before implementing the study. Creswell (2012) believes that scatterplot is an appropriate procedure that can be used to check the violation of the assumptions of linearity and homoscedasticity, also, it will clarify the nature of relationship between the variables. Therefore, these assumptions have to be tested before implementing

the correlation analysis.

3.7.13 Outliers Detection

As the researcher discussed previously the importance of outliers detection in the pilot stage of the current study. Pallant (2007) explained that outliers can seriously misrepresent the result obtained. Therefore, Pallant (2007) recommends that extreme outliers must be removed from the data set or recode them down to not extreme value. Regarding the questionnaire for the main study, the researcher used the statistical analysis (SPSS) to check extreme values and outliers. Figure 3.6 and Figure 3.7 indicate no extreme values in the data set which reflected a normal values distribution in both variables.

Figure 3.6: Extreme Values and Outliers in the Questionnaire

Figure 3.7: Extreme Values and Outliers in the Observation

3.7.14 Normality Test for the Main Study

Normality test was conducted to check normality distribution of the sample. The Skewness and Kurtosis values calculated for the main study from the questionnaire. Skewness value was (-0.05) which is close to (0), and the Kurtosis value was (-0.37). These results indicate a normal distribution of the sample in the main study (Tabachnick and Fidell, 1996). In addition, Skewness value (-0.05) was divided by the standard error value of the Skewness ($SE = 0.16$) which resulted in (-0.32), and the value of normal distribution ranged between (1.96) and (-1.96).

Also, Kurtosis value (-0.37) was divided by the standard error value of Kurtosis ($SE = 0.32$) which resulted in (-1.17), and the value of normal distribution ranged between (1.96) and (-1.96), which is considered reasonable distribution of data in the main study. In addition, Figure 3.8 below shows the sample distribution of the independent variable $N=230$. Also, the Skewness and Kurtosis values from the observation were calculated for the main study. Skewness value was (-0.06), and the Kurtosis value was (0.52). These results indicate a normal

distribution of the sample in the main study (Tabachnick and Fidell, 1996).

In addition, Skewness value (-0.06) was divided by the standard error value of the Skewness ($SE = 0.41$) which resulted in (-0.15), and the value of normal distribution ranged between (1.96) and (-1.96). Also, Kurtosis value (-0.52) was divided by the standard error value of Kurtosis ($SE = 0.80$), which resulted in (-0.65), and the value of normal distribution ranged between (1.96) and (-1.96), which is considered a reasonable distribution data of the main study. Figure 3.9 shows sample distribution of the dependent variable.

Figure 3.8: Sample Distribution of the Independent Variable

Figure 3.9: Sample Distribution of the Dependent Variable

In addition, the researcher in table 3.12 below used Kolmogorov-Smirnov test and Shapiro Wilk test to assess the normality distribution of the independent variable (teachers' knowledge of SQ from questionnaire), and the dependent variable (teachers' use of SQ from observation). The result indicated that the values of Kolmogorov-Smirnov test and Shapiro Wilk test were not significant which reflected a normal distribution of variables.

Table 3.13: The Test Results of Normality

Variables	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	D.f.	Sig.	Statistic	D.f.	Sig.
Total Questionnaire	0.06	230	0.06	0.99	230	0.15
Total Observation	0.11	33	0.20*	0.96	33	0.28

3.7.15 Linearity and Homoscedasticity

The two variables in Figure 3.10, the independent variable (teachers' knowledge of SQ) and the dependent variable (teachers' use of SQ) reflected that the assumption of linearity is reasonable, as teachers' knowledge of SQ increases, teachers' use of SQ increases. Also, Figure 3.10 shows that all the points were arranged in a narrow cigar shape which indicates a strong correlation between the two variables. This suggests, there was no violation of the assumption of linearity and homoscedasticity.

Figure 3.10: Linearity and Homoscedasticity of Independent and Dependent Variables

Thus far, the researcher described, discussed, and explained the research design and approach of the study, sample of the study, intervention program, research instruments and the relevant aspects of the reliability and validity of the test and instruments. This leads to the next stage

of the actual implementation of the study relevant to phase I which will further lead to the phase II of the undertaken study. The researcher has already discussed in details the different stages of phase I study in proceeding chapters.

3.8 Phase II: Data Collection Instruments and Procedural Aspects

This section is devoted to the discussion of the three instruments that have been used in phase II to collect the required data for the study. The section also explains as to how collected data would be used for the purposes of controlled analysis, and for developing a multilevel discussion of pedagogical issues relevant to the concerns of the study.

3.8.1 The Watson-Glaser CT Appraisal Form-A Test

The Watson-Glaser CT Appraisal Form-A (WGCTA-A) intended to test essential aspects of CT and has been used in different educational situations (Watson and Glaser, 2006). Gadzella (2005) indicated that the WGCTA-A is a practical tool to assess CT for learners. In the guidebook for the WGCTA-A, Watson and Glaser (2006) regarded CT as knowledge, attitude, and ability. Related to the aspects integrated into their practices, Watson and Glaser (2006) assumed that CT should contain the following: the ability to identify the indicators of conflicts and recognize the common requirements for confirmation to support what is stated as facts; ability to develop awareness of the nature of practical inferences, perceptions, and overviews in which the accuracy or weight of various types of confirmation are rationally concluded; and also, the ability to use them appropriately. The present study reflects various features of this explanation: (a) CT is not the conclusion of how to solve a problem but applying the skills to identify a problem; (b) CT must reflect basic support for what is concluded to a fact; (c) CT must contain the possibility to hypothesize before actual ideas are involved to solve a problem; and (d) CT must have the skill to assess and take appropriate decisions based on evidences.

The Watson-Glaser CT test contains 16-situations and 80-question items which are divided into five sub-sets. Each sub-set attempts to assess different element of CT. Watson and Glaser (2006) have linked the question types in the five sub-sets with the following performative

functions of those questions: (a) Inference; (b) Recognition of Assumptions, (c) Deduction, (d) Interpretation, and (e) Evaluation of Arguments. Each sub-set has a reading text or a situation that contains statements, arguments, interpretations, and problems, and information related to these met in everyday practices (Watson and Glaser, 2006). Reading texts are categorized as either "neutral" or "controversial". A "neutral" text provides a context in which a learner is expected to have no judgment about the topic, while, a "controversial" text expects the learner to use scenarios that provoke real "emotional feelings". The respondents' total score in this test is calculated based on the total number of questions answered appropriately from the comprehensive list and not from the individual sub-set. The reliability of this CT test depends on assessing the individual sub-sets and calculating the scores separately for each sub-set (Watson and Glaser, 2006).

This test includes 80 multiple-choice items ranging from 2 to 5 in its five sub-sets. Students were given five scenarios to evaluate the performative function involved in each scenario. The five scenarios offer scores for five different subtests ranging from 0 to 80. The five sub-sets focused on the following: (1) Inference: respondents are supposed to conclude inference among different levels of truth or falsity from given data. (2) Recognition of Assumptions: respondents are expected to identify unstated assumptions or sup-positions implied in given statements or assertions. (3) Deductions: respondents are required to determine if certain conclusions logically follow from the given statements and the data. (4) Interpretation: respondents need to evaluate the evidence and come to a decision if conclusions or overviews are based on the given information. (5) Evaluation of Arguments: respondents are subjected to different between discussions that are relevant and well-built, and those that are irrelevant and not properly justified in a particular situation (Watson and Glaser, 2006). This test covers a comprehensive range of CT: therefore, it is a valid benchmark for different professional and academic indices of success. This test is intended to collect quantitative data as a pre and post-test relevant to RQ2.

3.8.2 Pedagogical Strengths of Watson-Glaser CT Test

The Cronbach's alpha coefficient for WGCTA-A CT test was $r = .81$, as reported by (Watson and Glaser, 2006). Recently, reliability investigation results were very close to former scores, ranging from (0.77) to (0.86). The confirmation for strong validity of the WGCTA-A has been related mainly to the use of anticipative validity studies (Impara and Plake, 1998).

In spite of some valid concerns expressed by some scholars about the applications of the Watson-Glaser CT Appraisal test, it has still gained high approvals for its application in assessing CT abilities. Generally, criticisms on inventory have been positive in its use. Impara and Plake (1998) have considered the WGCTA-A to be one of the best tools available to test CT abilities. Furthermore, Impara and Plake (1998) claim that the WGCTA-A Test reflects a practical and successful application in educational and evaluation research courses and programs, and it is adequately reliable and therefore should be considered to be valid. It also, considered to be a useful and practical instrument for assessing CT abilities among teachers in instructional setting. Led by these motivating reasons, the researcher has preferred to choose the Watson-Glaser CT Appraisal, Form-A, for his study.

3.8.3 Interviews and Quizzes

Interviews and quizzes were considered important useful instruments in phase II of the study. Both interviews and quizzes were conducted only with the experimental group after the intervention period and post-test. Interviews aimed to collect qualitative data, whereas quizzes intended to collect quantitative data relevant to RQ3.

There were 30 students in the experimental group. By the end of the course and after the post-test was conducted, the instructor of the group selected those students who volunteered and agreed to participate in the interview. Only five students (20%) of the sample size agreed to participate in the interview stage (student1, student 7, student 19, student 20, and student 21). The researcher met these students, took their consent to participate in the

interview and explained to them about the interview process and all of them agreed to participate in the interviews. Interviews between the researcher and the participants were conducted individually after the eighth week. The time was given for each interview was about 10 minutes. All the interviews were completely and carefully recorded in the form of dairy notes.

The interview questions were developed by the researcher in a reflective framework taking focus and direction from the five thematic considerations of SQ in Paul and Elder's (2006) model. The questions used in the interview reflected most of the CT skills and SQ elements. Interviewees were asked 6 questions in order to get a deeper insight concerning how could they comprehend the concepts of CT, the level of difficulties they faced, and, how the course content and the language skills they were exposed to during the intervention phase helped them finding the appropriate answers to various CT questions.

Before starting the actual intervention phase, the students were trained in 1-week workshop. The teacher explained to the students the procedural aspects and introduced the tasks, activities, and assessment procedures related to CT strategies and SQ in which they will be involved during the intervention phase. By the end of the workshop the students were clear about their roles and responsibilities during the forthcoming eight-week long intervention.

Students were tested on 4 quizzes during the intervention period of eight weeks. Students were made to take a quiz at the end of every second week. Each quiz contained the five CT skills (questions) based on the five thematic components of SQ as suggested in Paul and Elder's (2006) model and based on the framework of WGCTA-A Test. These questions, developed by the researcher were open-ended and were based on SQ strategies in a reflective framework. Needless to say, the participants were tested on the questions based on the five CT-SQ components of Paul and Elder's (2006) model. Table 3.14 below describes the schedule of quiz sessions. On the other hand, the response mechanism on the part of interviewees was also reflective and analytical in nature. The researcher's intention in this, was not only to get

the responses to interview questions, but also to elicit an evidence of the reflective thinking process in which the participants had to engage in during the process of using enquiry oriented-CT skills with a questioning mind.

Table 3.14: Quiz Plan During the Intervention Stage

Time	Quiz Number	Reading Topic
A week before the intervention program	Students induction week, they are familiarized with SQ teaching strategies and course materials.	
week 2	Quiz 1	Reading topics 1 & 2
week 4	Quiz 2	Reading topics 3&4
week 6	Quiz 3	Reading topics 5 & 6
week 8	Quiz 4	Reading topics 7 & 8

To address the issues of validity and reliability of the two instruments discussed in this section, the researcher would like to make two comments. (i); as far as the validity and reliability of quizzes is concerned, it was ascertained as a part of the documents related to the intervention phase, as already. (ii); the validity and reliability of the structured interview can be considered as “met” or “achieved” because the findings of this data are close to the findings of the post-test data collected through the instruments of Watson and Glaser (2006) CT test, and the results of the quizzes that assessed the progress line of the participants during the intervention.

3.8.4 Statistical Analysis Procedures

The current study followed quantitative approach to research design. In phase I of the study, a questionnaire and an observation checklist were used to collect the required data related to teachers’ application of SQ strategies. The questionnaire was used to find out teachers’ knowledge of SQ whereas the observation checklist was used to determine the actual use of SQ in a real classroom situation.

To analyze the required data and find the relationship between the target variable in phase I, Pearson's correlation coefficient was used to find out the correlation between teachers' knowledge of SQ and actual use of SQ related research question 1 (RQ 1). In phase II, experimental design was used in which pre-test and post-test were conducted to find out the result of the intervention on the experimental target group which related to research question 2 (RQ 2). Also, in this phase multi-method measures were used as means to collect sufficient data in order to facilitate the process of triangulation of data related to research question 3 (RQ 3). To further elaborate, T-test, frequencies, and interviews analysis, were used to measure the various results of SQ as a result of the intervention on developing students' CT skills.

Table 3.15: Summary of Data Analysis Procedures

Research Questions	Data Collection Instruments	Sample Of Participants	Data Analysis Procedure
1. What is the relationship between teachers' knowledge of SQ and their actual use of SQ?	Quantitative Data Questionnaire and Observation Checklist	Teachers	Relationship: Pearson Coefficient
2. Are there any statistically significant differences between mean scores of students who are taught through SQ and mean scores obtained by students who are taught CT skills in a normal setting?	Quantitative Data WGCTA CT Test Pre and Post-Tests	Students	T-test
3. What CT strategies the students were able	Quantitative & Largely qualitative Data	Students	Total Scores

to develop and apply at the end of the intervention phase?	Quizzes and Interviews		and Coding
--	---------------------------	--	------------

3.8.5 Ethical Considerations of Research

Ethical principles and procedures direct research studies. Therefore, researchers are driven by the ethical regulations that make them committed to the roles and principles of research ethics. These differ according to nature and background of research studies. However, they govern academic integrity and honesty. In addition, they confirm that research has nothing violates privacy (Creswell, 2012).

The current study has three ethical considerations which were thoughtfully considered during the implementation process of this study. Firstly, permission was approved before the study began from the Ministry of Education to conduct the study in the public schools (appendix 1). Secondly, the participants' agreement to participate in the study was mentioned and explained while distributing the questionnaires to concerned teachers. Regarding the students who participated in the interviews, they volunteered and agreed to have the interviews without any pressure on them. Thirdly, since the respondents were all Muslim Oman females, the researcher is fully aware of Omani social customs and cautions.

3.9 Summary and Conclusion

This chapter was mainly devoted to the description, discussion, and explanation of the research design evolved in this study, research instruments to be used, and required methodological considerations. This study adopted quantitative-qualitative data, and based its procedures on five research instruments, namely self-developed teachers' knowledge questionnaire, classroom observation checklist, WGCTA-A CT test, four classroom-based quizzes, and self-developed structured interview. Due ethical procedures were followed to develop and validate

the data collection instruments. The study was planned to be carried out in major two phases. In the first phase, the questionnaire and interview were piloted to ensure the required validity and reliability of the instruments.

After that, the two instruments were subjected to the full study and the data was collected from the total sample of the study and the correlation coefficient analysis was conducted. Further, in the second phase of this study, the other three instruments namely, CT test, classroom quizzes, and the structured interviews were used to collect the required data from the participants of this study. The data collected through these instruments were properly organized to be ready for the analysis. What follows next is the analysis and findings of the data collected for this study.

CHAPTER 4

ANALYSIS AND FINDINGS

4.1 Introduction

The present study is primarily focused on a pedagogical investigation of a very crucial and vital aspect of classroom instructional practice namely SQ in the immediate context of Omani post-basic schools located in the Dhofar region. This study was planned and conducted in two phases. This chapter presents the statistical analysis and the findings of the data collected for this study in two phases. Phase I of the study is devoted to the analysis and findings of RQ 1. And, Phase II deals with the analysis and findings that are related to RQ2 and RQ 3 of the study.

This chapter presents the analyses and findings related to three major objectives of the research which are realized through the investigation of the three focused research questions. Research question 1 (RQ1) is intended to measure the relationship between teachers' knowledge of SQ and their use of SQ in their classroom performances in phase I of the study. Similarly, research question 2 (RQ2) is aimed at examining the statistically significant differences between students' score in the experimental group and the control group. And, research question 3 (RQ3) is used to investigate the CT skills which were developed by the students during the intervention phase. This preliminary discussion of the research questions linked with research objectives leads to the expected analyses and findings of data and results in the following sections.

4.2 Phase I: Analysis and Findings

The main purpose of Phase 1 of the study was to collect required data, and to investigate the relationship between *teachers' knowledge of SQ* and *teachers' actual use of SQ* in a real classroom context. In this chapter, phase I data is analyzed to examine the corresponding relationship between these two variables "teachers' knowledge of SQ" and "teachers' actual use of SQ" which relates to RQ1: What is the relationship between teachers' knowledge of SQ and their actual use of SQ? The data were gathered through two instruments. First, the questionnaires were administered on 230 (n=230) teachers, and total number of completed questionnaires was also 230 teachers (n=230). Based on the RQ1 data, the following table shows the weighted mean score of teachers' knowledge of different domains of SQ (section 3.6.1. weighted mean values scale).

Table 4.1: The Teachers' Responses to their Knowledge in Different Domains of SQ Questioning (N=230)

Teachers' Knowledge of SQ	Mean	St. Dev.	Weighted Mean	Level of Response
Inference	15.87	3.33	2.65	Medium
Recognition of Assumptions	9.64	2.09	1.93	Low
Deduction	9.60	1.98	1.92	Low
Interpretation	21.11	3.54	2.63	Medium
Evaluation of Arguments	10.64	1.70	2.13	Low
Teachers Knowledge of SQ	66.86	6.12	2.30	Low

Table 4.1 shows that the level of teachers' knowledge of SQ questioning was low on the total score of the questionnaire with a weighted mean score of 2.30 and standard deviation of 6.12. In terms of individual domains, the level of response is found to be low in recognition of assumptions category with the weighted mean score of 1.93 and standard deviation of 2.9; in the deduction category the weighted mean score was 1.92 and standard deviation of 1.98; and in the evaluation of arguments domain the weighted mean score was 2.13 with standard deviation of 1.70. On the other hand, the level of response was medium in the domain of inference with the weighted mean score of 2.65 and standard deviation of 3.33; and the same was in the interpretation domain with the weighted mean score of 2.63 and standard deviation of 3.54.

The holistic result of teachers' total knowledge of SQ with a weighted mean score of 2.30 is low and standard deviation of 6.12. The narrow range of standard deviation in the five domains suggests that the sample of the study is approximately a homogenous group in terms of low level of teachers' knowledge of SQ, which reflects that the majority of participants are at a lower level on the scale of expected SQ knowledge. After examining the questionnaire data of 230 teachers, there is a need to look at the responses to the questionnaire of those 33 teachers separately who were observed by the researcher using classroom observation checklist for the purpose of correlation a little later in the forthcoming discussion.

Table 4.2: The Teachers' Responses to their Knowledge in Different Domains of SQ Questioning (N=33)

Teachers' Knowledge of SQ	Mean	St. Dev.	Weighted Mean	Level of Response
Inference	15.82	2.99	2.64	Medium
Recognition of Assumptions	9.84	2.21	1.97	Low
Deduction	10.36	1.71	2.07	Low

Interpretation	22.69	4.16	2.83	Medium
Evaluation of Arguments	10.88	1.78	2.18	Low
Teachers' Knowledge of SQ	69.61	6.83	2.40	Low

Table 4.2 shows that the level of teachers' knowledge of SQ questioning was low on the total score of the questionnaire with a weighted mean of 2.40 and standard deviation of 6.83. In terms of individual domains, the level of response is found to be low in recognition of assumptions category with the weighted mean score of 1.97 and standard deviation of 2.21; in the deduction category the weighted mean score was 2.07 and standard deviation of 1.71; and in the evaluation of arguments domain the weighted mean score was 2.18 with standard deviation of 1.78. On the other hand, the level of response was medium in the domain of inference with the weighted mean score of 2.64 and standard deviation of 2.99; and the same was in the interpretation domain with the weighted mean score of 2.83 and standard deviation of 4.16.

The holistic result of teachers' total knowledge of SQ with a weighted mean score of 2.40 is low and standard deviation of 6.83. The narrow range of standard deviation in the five domains suggests that the sample of the study is approximately a homogenous group in terms of low level of teachers' knowledge of SQ, which clearly indicates that all the participants have more or less the same level of lack of desired knowledge of SQ questioning.

The next table comprises of data generated from classroom observations to measure the teachers' use of SQ. This data was collected from 33 teachers who, upon researcher's request, volunteered to undergo classroom observations. The required data was collected using the observation checklist (appendix 8). This sample size of 33 teachers is a reasonably

representative sample size over the total sample of the study for classroom observation.

Table 4.3: The Levels of Teachers' Use of SQ Questioning in Different Domains of SQ Questioning (N=33)

Teachers' use of SQ	Mean	St. Dev.	Weighted Mean	Level of Response
Inference	12.03	2.63	2.01	low
Recognition of Assumptions	10.55	2.32	2.11	low
Deduction	9.52	2.0	1.90	low
Interpretation	13.27	2.43	1.66	very low
Evaluation of Arguments	10.42	1.92	2.08	low
Teachers' use of SQ	55.79	6.92	1.92	low

Table 4.3 shows that the level of teachers' use of SQ questioning was low on the total score on the observation checklist with a weighted mean of 1.92 and standard deviation of 6.92. In terms of individual domains, the level of response is found to be low in the inference domain with a weighted mean score of 2.01 and standard deviation of 2.63; in recognition of assumptions category with the weighted mean score of 2.11 and standard deviation of 2.32; in the deduction category the weighted mean score was 1.90 and standard deviation of 2.0; and in the evaluation of arguments domain the weighted mean score was 2.08 with standard deviation of 1.92. However, the level of response in the interpretation domain was very low with weighted mean score of 1.66 and standard deviation of 2.43. The overall result of teachers' total use of SQ with a weighted mean score of 1.92 is low and standard deviation of 6.92. The narrow range of standard deviation in the five domains suggests that the sample of the study is approximately a homogenous group in terms of the level of using teachers' knowledge of SQ.

The same teachers' questionnaires and observation checklists were coded to maintain exact correlation between knowledge and use for each specific teacher. The correlation results

were analyzed for the 33 teachers whose classrooms sessions were observed by the teacher.

Demographic data has no relevant concern for the main focus of this study: hence, it was not targeted. Therefore, the questionnaires and the observation checklists played an important role to cover the pedagogical data related to the aims and objectives of the study and the related research question. What is needed and follows next is a discussion of findings related to RQ1 which focuses on the relationship between teachers' knowledge of SQ and its use in their instructional practices.

4.2.1 Main Findings of RQ1

To answer research question 1 which investigates the relationship between teachers' knowledge of SQ and teachers' actual use of SQ, Pearson's correlation coefficient was subjected to the 33 teachers who volunteered to have classroom observation. The relationship between *teachers' knowledge of SQ* which acts as an independent variable (measured by the questionnaire), and *teachers' actual use of SQ* which acts as a dependent variable (measured by the observation) was investigated using Pearson product-moment correlation coefficient. Preliminary analyses were conducted to ensure that there were no instance(s) of violation of the assumptions of normality, linearity, and homoscedasticity as discussed in chapter three.

As evident in Tables 4.2 and 4.3 above, weighted mean score of teachers' total knowledge of SQ was 2.40 and the corresponding weighted mean of teachers' total use of SQ was 1.92- both fall in the range of low level of response. The breakup of the total knowledge of SQ and the total use of SQ from the five CT domains suggests that there is no significant deviation in the weighted mean scores of the five CT domains.

In the inference domain, the weighted mean score of the teachers' knowledge of SQ was at the medium level (i.e., 2.64 as illustrated in Table 4.2), but the corresponding weighted mean score for the teachers' actual use of their knowledge in the same domain was at the low level (i.e., 2.01 as shown in Table 4.3). the finding indicates that the teachers were less likely

to apply their knowledge of SQ in the inference domain in the actual classroom. Similarly, the medium level of response of teachers' knowledge of SQ in interpretation domain with the weighted mean score of 2.83 became very low level of response with the weighted mean score of 1.66 in the teachers' use of SQ. On the Other hand, teachers' knowledge of SQ in recognition of assumptions domain with a weighted mean score of 1.97; in deduction domain with the weighted mean score of 2.07; and in evaluation of arguments domain with a weighted mean score of 2.18 remained low in teachers use of SQ with a weighted mean score of 2.11 in recognition of assumptions domain; in deduction domain with weighted mean score of 1.90; and in evaluation of arguments domain with a weighted mean score of 2.08. It is obvious from the data that there is a close correspondence between the teachers' knowledge and use of SQ in most of CT domains except interpretation domain which emerged as a challenging domain in terms of its use in the class.

The above statistics, when subjected to Pearson's coefficient statistical procedure suggested that there was a strong and positive correlation between the two variables of teachers' knowledge and teachers' use of SQ. Table 4.4 below shows a strong, positive correlation between the two variables, $r = 0.79$, $n = 33$, $p < 0.01$. This means that teachers who had low scores in the knowledge of SQ variable are likely to achieve lower scores in the corresponding variable of actual use of SQ in the classroom context. This is evident in the result of the study where teachers with low scores in the knowledge variable have scored low in the use variable

Table 4.4 : The Correlations between the Two Variables of Questionnaire & Observation

		Total	Total
Total Questionnaire	Pearson Correlation	1	0.79**
	Sig. (2-tailed)		.000
	N	33	33
Total Observation	Pearson Correlation	0.79**	1
	Sig. (2-tailed)	.000	
	N	33	33

** . Correlation is significant at the 0.01 level (2-tailed).

4.3 Phase II: Analysis and Findings

This section presents the findings of phase II in the current study. Phase II was proposed to investigate the research RQ 2 and RQ 3 Before presenting the actual findings of phase II of the study, it would be in the fitness of the discussion to describe the various tests that were involved in phase II prior to the intervention stage to maintain normality distribution of the sample, and homogeneity and equivalency test to ensure that the two groups were equal in relation to some external variables prior to the beginning of the experimental stage.

4.3.1 Normality and Homogeneity Measures

The description on the sample of the study, as presented in the section 3.3 on methodology in chapter three, holds relevance for the discussion of sample in this section. Hence, the discussion moves on to the description and purposes of tests used in phase II of the study. The first test used was the normality test. The second test was homogeneity and equivalency which test students' English language achievement score (ELAS), their total grade point average (TGPA), critical thinking skills score (CTSS), and intelligence quotient score (IQS) prior to the intervention stage. Third test was homogeneity and equivalency test which measures the participants' level of CT prior to the intervention stage in two modes. First, general holistic level

of CT as highlighted in Table 4.5, and second at the level of five individual domain-specific categories as displayed in Table 4.6.

i) Normality Test

Normality test was conducted on the control group and experimental group to maintain normality distribution of the sample. Skewness coefficient for the control group ranged from (0.0) to (0.35), and for the experimental group from (0.0) to (0.58). Both results are close to (0) which indicate normal distribution of the sample in the two groups. Also, Kurtosis coefficient for the control group ranged from (0.46) to (-1.33) and (-0.39) to (-1.10) for the experimental group which is close to (1) for both the groups. This indicates normal distribution of the sample in the two groups. The researcher also used Kolmogorov-Smirnov test to determine if the two samples related to the control group and experimental group had the similar normal distribution. The Z value was (0.65), and the significance value was (0.78) which means that there was no statistical significance at the level of 0.05.

ii) Homogeneity and Equivalency Test

In order to measure the homogeneity and the equivalency of the control group and the experimental group, before implementing the intervention of the variables, the participants' English language achievement score (ELAS), their total grade point average (TGPA), critical thinking skills score (CTSS), and intelligence quotient score (IQS) obtained in the tests conducted prior to the intervention phase were subjected to t-test in order to calculate the differences of the afore-mentioned pre-variables between the two groups. Table 4.5 below shows the differences of the significance between the participants in the two groups related to external variables.

Table 4.5: Significance of Differences between the Two Groups in the External Variables

CT Variables	Control Group N=30		Experimental Group N=30		t-value	Sig.
	Mean	Std. Div.	Mean	Std. Div.		
ELAS	79.30	9.78	79.37	9.81	- 0.03	0.98
TGPA	717.80	67.34	714.73	62.81	0.18	0.86
CTSS	30.40	02.37	29.77	1.92	1.14	0.26
IQS	45.63	05.73	44.73	5.30	0.63	0.53

Table 4.5 shows that there were no statistically significant differences between the mean scores of the experimental group and the control group at the level of (0.05) in the pre-test of the ELAS (0.98), TGPA (0.86), CTSS (0.26), and IQS (0.53). This shows that the two groups were equal in relation to afore-mentioned test scores prior to the beginning of the experiment.

iii) Homogeneity and Equivalency of CT Skills

The researcher conducted t-test in the pre-test stage which reflects the differences between the mean scores of the control group and the experimental group in the five domains (Inference, Recognition of Assumptions, Deduction, Interpretation, and Evaluation of Arguments) of Watson and Glasser (2006) CT Test. This was to make sure that the dependent variable which is students' CT skills in both the control group and the experimental group was equal prior to the implementation stage. Table 4.6 below explains the equality of variances between the two groups.

Table 4.6: Significance of Differences in CT Test between the Two Groups in the Pre-Test

CT Domains	Control Group N=30		Experimental Group N=30		t-value	Sig.
	Mean	Std. Div.	Mean	Std. Div.		
Inference	6.10	1.16	6.20	1.03	- 0.35	0.73
Recognition of Assumptions	6.27	1.11	5.97	0.85	1.17	0.25
Deduction	5.83	0.87	5.63	0.89	0.88	0.38
Interpretation	6.17	0.75	6.17	0.59	0.00	1.00
Evaluation of Arguments	6.03	0.89	5.80	0.92	0.99	0.32
Total of CTSS	30.40	2.37	29.77	1.92	1.14	0.26

The above Table represents the statistics related to the results of pre CT test taken by the control group and experimental group. The results of t-test in terms of t-value and significance value are also highlighted in this table. As evident in the above table, the total mean of the control group in the five CT variables is 30.40 with a standard deviation of 2.37, and the total mean of experimental group in the five CT variables is 29.77 with a standard deviation of 1.92. This suggests that both the control and the experimental group were more or less at the same level of CT abilities. If we compare the individual means of CT variables in the two groups, it is obvious that there are no statistically significant differences in the pre-test mean scores of students' CTS between the control group and the experimental group in the five domains and the total score of CTSS-Watson-Glasser's CT test.

The t-value of the difference between the two groups of inference domain was ($t=-0.35$) which is not significant ($\text{Sig.}=0.73$). The t-value of the difference between the two groups of recognition of assumptions domain was ($t=1.17$) which is considered as not significant at ($\text{Sig.}=0.25$). The t-value of the difference between the two groups of deduction domain was ($t=0.88$) which is counted as not significant at ($\text{Sig.}=0.38$). The t-value of the difference

between the two groups of interpretation domain was ($t=0.00$) which is not worthy of considering it as significant at ($\text{Sig.}=1.00$). The t -value of the difference between the two groups of evaluation of arguments domain was ($t=0.99$) which is not to be considered as significant at ($\text{Sig.}=0.32$). As a result, both the control group and the experimental group were at equal level in all the five domains of Watson and Glasser's (2006) CT Test.

4.3.2 Phase II: Analysis and Findings of RQ 2

This section is meant to discuss phase II results relevant to the RQ2, as shown in Table 4.7 below. The RQ2 is; Are there any statistically significant differences between mean scores of those who are taught through SQ and those obtained by students who are taught CT skills in a normal setting? The data in Table 4.7 highlights the differences between mean scores of those who were taught CT through SQ and those who were taught CT skills in a normal setting. The t -test was conducted to investigate the differences between the two groups in relation to the five CT domains as specified in Watson-Glasser's (2006) CT test. It has been reported that so far as the domain of "interpretation" was concerned, there was a large difference between the mean scores obtained by the control group (i.e., 5.80) and that of the experimental group (i.e., 7.90), with a p -value of less than 0.05. However, it was previously stated that "interpretation" was the domain in which there was an obvious difference in the teachers' knowledge of SQ and their actual use of SQ. Looking at the data obtained from the teachers in phase I and that obtained from the students via the quasi-experimental study, it appears that despite the lack of teachers' readiness to apply the SQ knowledge (relating to "interpretation" in particular) in the actual classroom, the students' scores could still be significantly improved if the students were taught the CT skills via SQ (in the experimental group).

Table 4.7: Significance of Differences in CT Test between the Two Groups in the Post-Test

CT Domains	Control Group N=30		Experimental Group N=30		t-value	Sig.
	Mean	Std. Div.	Mean	Std. Div.		
Inference	6.53	1.07	7.07	0.94	-2.04	0.046
Recognition of Assumptions	6.20	0.89	6.93	1.08	-2.87	0.006
Deduction	6.03	1.30	7.83	1.05	-5.89	0.000
Interpretation	5.800	1.03	7.90	1.12	-7.54	0.000
Evaluation of Arguments	6.07	0.87	6.57	0.97	-2.10	0.040
Total of CTSS	30.63	2.87	36.37	2.03	-8.94	0.000

The above table highlights the statistics related to the results of post CT test taken by the control group and experimental group. A t-test was used to determine the statistical significance related to the null hypothesis. The null hypothesis was, there were no statistical significant differences at the level of (0.05) between mean scores of those who are in the first group, who were taught through SQ and those in the other group, who were taught CT skills in a normal setting. As evident in the above table, the total mean of the control group in the five CT domains is 30.63 with a standard deviation of 2.87, and the total mean of experimental group in the five CT domains is 36.37 with a standard deviation of 2.03. This suggests that there is a noticeable improvement of the total score of both the total mean and the standard deviation in the experimental group. If we compare the individual means of CT domains in the two groups, there is a noticeable improvement in the post-test mean scores of students' CTs in the experimental group as against the control in the five CT. Furthermore, the result of t-test

comparisons shows statistically significant differences between mean scores of those who were taught CT through SQ and those who were taught CT skills in a normal setting.

There are significant differences at the level of (0.05) between mean scores of those who are in the first group, who were taught through SQ and those in the other group, who were taught CT skills in a normal setting in the individual five domains of CT and the CT total score; hence, the null hypothesis was rejected. This result indicates that teaching CT skills through SQ strategies is beneficial and more productive. Furthermore, t-test results show significant differences between the two groups in relation to the five CT domains.

i) The Main Findings of RQ2

The total mean score of the experimental group at the pre-test stage was 29.77 with a standard deviation of 1.92. However, the total mean score in the post test was raised to 36.37 with a standard deviation of 2.03 as a positive effect of the intervention treatment on the experimental group. In terms of standard deviation, there is no significant deviation noticed between the pre-test stage and the posttest stage scores. This suggests that the participants maintained more or less the same level of homogeneity in their efforts to improve. Further examination of the individual CT domains in the pre-test and the posttest of the experimental group suggest that there was a considerable degree of improvement.

To illustrate, the pre-test mean score of 6.20 in the inference domain became 7.07 in the posttest. The pre-test mean score of recognition of assumptions which was 5.97 became 6.93 in the posttest. Similarly, the pre-test mean score of 5.63 in the deduction domain was raised to 7.83 in the posttest, and the pre-test mean score of 6.17 in the interpretation domain became 7.90 in the posttest. Lastly, the pre-test mean score of evaluation of arguments domain was 5.80 which was improved to the mean score of 6.57 in the posttest.

It is evident from the result that those students who were made to undergo intervention improve upon their CT skills and earned higher scores than the ones who were taught CT in a normal setting. The mean scores of the five CT domains show that students who were taught CT using SQ were able to develop and use the target skills in the five domains better than those who were taught in a normal setting. This result clearly demonstrates that the students who were taught CT skills through SQ developed abilities to use CT skills in the five domains at enhanced levels of competence. This also supports the premise that pedagogical training programs are integral to better skills development.

4.3.3 Phase II: Analysis and Findings of RQ3

The aim of RQ3 is to identify SQ strategies that helped the students to develop their CT skills during their experimental EFL reading classes which were meant to enhance their CT skills in the intervention phase.

To answer RQ 3, the researcher deployed multi-method research design involving quantitative as well as qualitative data. Four quizzes were used to collect the quantitative data that is related to RQ 3. Each quiz was conducted after completing 2 thematic reading units. Eight units were taught during the course of intervention. All the quizzes were developed based on Watson and Glasser's (2006) five domains of CT Test (inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments). The purpose of these quizzes was to measure the developmental progress of CT skills that students made through the intervention phase related to quantitative data.

The second instrument was the structured interview. Five students, out of thirty in the experimental group agreed and volunteered to be interviewed. In this structured interview, the students were asked to inform about what did they actually do to answer the questions in the quizzes and to justify their responses. Quiz 1 was selected as the starting point, and quiz 4 as the finishing point to investigate students' progression and improvement in CT skills.

After the general description of data collection instruments relevant to RQ 3, it is now time to present the findings of both quantitative and qualitative data followed by a juxtaposed analysis of data.

i) Quantitative Data Analysis

As discussed in section 4.5, the quantitative data was collected through the results of the four quizzes conducted at the end of every two weeks. The quiz results were duly tabulated, as highlighted in Table 4.8. This table shows the individual scores of the five CT skill. This information is tabulated in the ascending order of quizzes.

**Table 4.8: The Mean Scores of CT Skills in the Four Quizzes
N=30**

No.	Critical Thinking Skills (3 marks each)	Mean Score			
		Quiz 1	Quiz 2	Quiz3	Quiz4
1	Inference	1.80	2.40	2.47	2.93
2	Recognition of Assumptions	1.93	2.43	2.03	2.90
3	Deduction	1.63	1.67	2.53	2.83
4	Interpretation	1.87	2.00	2.37	2.83
5	Evaluation of Arguments	1.43	1.57	1.70	2.37
6	Total Mean score	8.67	10.07	11.10	13.87

Table 4.8 shows the progress scale of thirty students' development in each CT skill exposed to during the experimental phase. The tabulated data clearly indicates that all thirty students at the initial intervention phase of quiz 1 were at a lower level with total mean score of 8.67. However, in the final phase of quiz 4, after further intervention input of six weeks, the

results changed significantly with a total mean score of 13.87. If we compare the total mean score of quiz 1 and quiz 4, it is obvious that the scaffolding of CT skills has grown significantly by the difference of 5.20. After every phase of intervention from quiz 1 till quiz 4, every student gained a higher score in every domain of CT skills. The overall result in the ascending order of quizzes reflects a steady upward scaffolding and growth in CT skills. On a separate note, natural order of individual differences between the results of the participants in each CT skill are also clearly visible which suggests that the mixed abilities. Of students prevailed during the intervention phase. The above discussion of the total mean score of the four quizzes leads to the discussion of individual achievements of the five students in the five CT skills.

**Table 4.9: Students' Results of the Four Quizzes in Each Skill
N=5**

Student	Quiz Number	Critical Thinking Skills				
		Inference 3 marks	Recognition of Assumptions 3 marks	Deduction 3 marks	Interpretation 3 marks	Evaluation of Arguments 3 marks
Student 1	Quiz 1	1	1	1	1	1
	Quiz 2	1	2	1	1	1
	Quiz 3	2	2	2	2	2
	Quiz 4	3	3	2	2	3
Student 7	Quiz 1	2	2	1	0	1
	Quiz 2	2	1	1	1	2
	Quiz 3	3	3	2	1	1
	Quiz 4	3	3	2	2	2
Student 19	Quiz 1	2	2	1	2	2
	Quiz 2	3	3	2	2	2
	Quiz 3	3	2	3	2	2
	Quiz 4	3	3	3	3	2
Student 20	Quiz 1	2	2	2	2	2
	Quiz 2	2	2	1	1	1
	Quiz 3	3	2	2	3	2
	Quiz 4	3	3	3	3	3
Student 27	Quiz 1	2	2	2	2	1
	Quiz 2	3	3	2	3	2
	Quiz 3	2	2	3	2	1
	Quiz 4	3	3	3	3	2

Table 4.9 illustrates the line of progression the five individual students took in CT skills as a result of different stages of the intervention phase. The data in the above table suggests that all the five students in the initial phase of quiz 1 obtained lower scores in all the five skills. The total mean score of quiz 1 results for student 1 and student 7 was in a low range of 0.33 to 0.40 whereas the total mean score of student 19, 20, and 27 was in a slightly better range of 0.60 to 0.67. The total mean score of quiz 2 for students 1, 7, and 20 was still in a low range of 0.40 to 0.47 whereas the total mean score of student 19 and 27 was in a slightly better range of 0.80 to 0.87. The total mean score of quiz 3 results for student 1, 7, and 27 was in a low range at 0.67 whereas the total mean score of student 19 and 20 was comparatively slightly better at 0.80. Though the total mean score of quiz 4 for student 7, 1, and 27 further improved, it was in a low range of 0.80 to 0.93 whereas student 19 and 20 were comparatively at a higher range with the total mean score of 1.

The overall perspective of the total mean score of 4 quizzes suggests that the total mean score of students 1, 7, and 20 was at a lower range of 2.27 to 2.93 whereas the total mean scores of student 27 and student 19 were at a slightly higher range of 3.07 to 3.20. Based on this discussion of holistic as well as individual total mean scores of the five students in the results quizzes, it can be concluded that there is a reciprocal and reflective relationship between the two types of data. Furthermore, the data result of the five students similarly corresponds with the data results of the thirty participants discussed earlier in Table 4.8. This result is expected in a bigger group of participants with their learning strengths and weaknesses.

The above discussion of the results of thirty participants needs to be narrowed down in order to lead the discussion of quiz results of four quizzes relevant to the five students who volunteered and were interviewed accordingly. The following five tables show the results of quizzes taken by these students in the five CT skills adapted from Paul and Elder's, (2006) model.

Table 4.10: Students' Quiz Scores in Inference Skill (N=5)

Students	Quiz 1	Quiz 2	Quiz 3	Quiz 4
Student 1	1	1	2	3
Student 7	2	2	3	3
Student 19	2	3	3	3
Student 20	2	2	3	3
Student 27	2	3	2	3
Mean score	1.8	2.2	2.6	3

Table 4.10 aims to present students' scores obtained in Quiz 1 and Quiz 4 in inference skill during the intervention stage. The scores in the quizzes show that there is a gradual improvement in the inference skills of all students from quiz 1 to quiz 4. The scores obtained in the initial stage (quiz 1) indicate students' average achievement as they are in their initial stage at this point; student 1 scored (1 mark), and student 7, student 19, student 20, and student 27, scored (2 marks) each. However, after 8 weeks of intervention, students' scores obtained in quiz 4 show a significant improvement in their inference skills. All of them scored 3 out of 3 in the inference question. Hence, this good result in the inference skill could be based on the argument that frequent practice of inference-based tasks led to this positive achievement by the end of this intervention stage.

Table 4.11: Students' Quiz Scores in Recognition of Assumptions Skill (N=5)

Students	Quiz 1	Quiz 2	Quiz 3	Quiz 4
Student 1	1	2	2	3
Student 7	2	1	3	3
Student 19	2	3	2	3
Student 20	2	2	2	3
Student 27	2	3	2	3
Mean score	1.8	2.2	2.2	3

Students' scores highlighted in the above table that obtained from Quiz 1 through Quiz 4 in the recognition of assumptions skill shown in the above table, reflect a gradual improvement. Like the students' achievement in the inference skill, the scores obtained in the initial stage (quiz 1) indicate students' average achievement since this is their initial stage; student 1 (1 mark), student 7 (2 marks), students 19, 20, and 27 scored (2 marks) each. However, after 8 weeks of intervention all the students were able to improve upon their recognition of assumptions skill to score 3 out of 3 marks. This achievement could be considered as an indicator of students' success as a favourable result of training during the intervention phase.

Table 4.12: Students' Quiz Scores in Deduction Skill (N=5)

Students	Quiz 1	Quiz 2	Quiz 3	Quiz 4
Student 1	1	1	2	2
Student 7	1	1	2	2
Student 19	1	2	3	3
Student 20	2	1	2	3
Student 27	2	2	3	3
Mean score	1.4	1.4	2.4	2.6

Students' scores in the deduction skill from quiz 1 to Quiz 4 presented in Table 4.12 show a gradual improvement. The scores obtained in the initial stage of the intervention (quiz 1) turned out to be below average with (1 mark) each for the students 1, 7, and 19, and (2 marks) each for the students 20, and 27. However, at the final phase of the intervention (Quiz 4), students' scores show a significant improvement in the deduction skill. Students 1, and 7 scored (2 marks each) whereas students 19, 20, and 27 obtained 3 out of 3 marks each. The low results obtained in the deduction skill in comparison to the inference and recognition of assumptions skill indicate towards a higher level of difficulty involved in dealing with somewhat more challenging level of deduction skill as evident in the performance of students 1 and 7.

Table 4.13: Students' Quiz Scores in Interpretation Skill (N=5)

Students	Quiz 1	Quiz 2	Quiz 3	Quiz 4
Student 1	1	1	2	2
Student 7	0	1	1	2
Student 19	2	2	2	3
Student 20	2	1	3	3
Student 27	2	3	2	3
Mean score	1.4	1.6	2	2.6

The students' scores in the interpretation skill highlighted in the above table suggest an average progression and achievement. The entry level of students 1 and 7 was very low with the scores of 1 and 0 whereas the scores of students 19, 20, and 27 were average with the scores of 2 out of 3 marks each. At the end of the intervention phase in quiz 4 students 1 and 7 scored 2 out of 3 each whereas students 19, 20, and 27 obtained 3 out of 3 marks each, which is considered as a significant improvement in interpretation skill. Like the deduction skill, the interpretation skill also seems to be equally challenging in view of the low scores. The gradual improvements in the interpretation skill of the five students may be considered to testify that frequent practice of the interpretation oriented-tasks during the intervention phase led to the gradual enhancement of scores from the quiz 1 to 4.

Table 4.14: Students' Quiz Scores in Evaluation of Arguments Skill (N=5)

Students	Quiz 1	Quiz 2	Quiz 3	Quiz 4
Student 1	1	1	2	3
Student 7	1	2	1	2
Student 19	2	2	2	2
Student 20	2	1	3	3
Student 27	2	1	2	3
Mean score	1.6	1.4	2	2.6

The scores in the four quizzes presented in the above table show a trend of a gradual improvement in all the students' evaluation of arguments skill from quiz 1 till quiz 4. The scores obtained in the initial stage (quiz 1) reflect students' below average achievement with the score of (1 mark)each for students 1 and 7, and average achievement with (2 marks) each for students 19, 20, and 27. After 8 weeks of intervention, students' scores obtained in quiz 4 show a significant improvement in students' evaluation of arguments skill of students 1, 20, and 27 with the scores 3 out of 3 whereas, in the case of students 7 and student 19 it was and average improvement with 2 out of 3 marks each. The majority of the results in terms of scores obtained by students in Table 4.14 suggest that evaluation of argument skill emerged as more challenging skill for students as compared with the other four CT skill. However, the overall improvement in this skill is due to the effect of the eight weeks' intervention.

Having discussed the quantitative data thus far, it is now the time to move to the discussion of the qualitative data results which will provide a testing ground for the triangulation of both types of results.

ii) Qualitative Data Analysis

The qualitative data for this study was elicited through the structured interview mode. The interview questions focused only on the initial quiz and the final quiz, i.e. quiz 1 and quiz 4 as indicators of starting point as mentioned in the preceding section. Quiz 1 was conducted immediately after the completion of the first two units/themes of the intervention in the experimental phase. This was the phase where the students had started gaining CT skills. And, quiz 4 was conducted finally after the completion of eight units/themes. Thus, quiz 1 can be considered to indicate the initial level of competence in students' CT skills, and quiz 4 results to indicate the final level of competence of students in their CT skills gained as a result of the implementation of the total experimental intervention program.

The interview questions were analytical in nature (appendix 10) as the students were asked to explain and justify their responses in quiz 1 and quiz 4, on the basis of relevant

strategies that they had exploited to answer the questions for each CT skill in the two quizzes. This analytical interview helped the researcher to get comparatively more detailed and deeper insights into their thought process involved while answering the questions in the two quizzes. The interview results also helped the researcher in triangulating these results with the results of the quantitative data to testify the credibility of the two types of instruments and the results. A summary of students' responses to interview questions for each CT skill related to quiz 1 and quiz 4 is presented in the following tables.

Table 4.15 below comprises students' responses to the five CT skills tested in quiz 1. This summary is not exhaustive. It only includes the most relevant and representative comments of the participants who were directly linked with the main focus and the purpose of the study (Clarke and Braun, 2013). The redundant and irrelevant parts have been ignored. It is to be noted that the participants' responses included in the following table are not full responses as they were made by them. The reason for doing so is due to the fact that these responses included lot of repetitive, redundant, and sometimes irrelevant information, which might affect the readability adversely. Therefore, only the thematically relevant phrases and clauses have been reproduced here. Detailed students' responses are included in (appendix 10).

Table 4.15: Students' CT Responses to the Interview Questions in Quiz-1 (N=5)

Critical Thinking Skills – Quiz 1					
Students	Inference	Recognition of Assumptions	Deduction	Interpretation	Evaluation of Arguments
	Quiz 1	Quiz 1	Quiz 1	Quiz 1	Quiz 1
1	was difficult, just select any answers	too much time, frustrating	too much pressure at the first quiz, reading text too long, lost reading the text	answers were very close to each other; focused on answers not questions	very hard, reading, guessing
7	many times I read the text, was not clear.	can't, reading the text only, try to guess answer	question was not clear, I read the text, difficult	answers looks all correct	read conclusions but not understand
19	some words were difficult....and time was passing very quickly	answers were same	reading but too much information	only focused on the statement, forgot to connect with the text	can't match between the argument and the statement
20	(pause) didn't understand the text, question was difficult	couldn't control the time, statement wasn't clear	couldn't focus on major information in the statement	was tensioned, don't know how to start, just guessing	couldn't read the arguments with appropriate understanding
27	first quiz and second one were difficult,	was difficult to make sense of	couldn't use deduction skill, many	I think you need more than reading,	couldn't link different arguments to

	I read the text but couldn't understand the conclusions	assumptions	information in the text	was not really focused	the main statement
--	---	-------------	-------------------------	------------------------	--------------------

After presenting the students' responses to their achievement in the five CT skills in quiz 1, what follows next in sequence is the Table 4.16 which shows the responses of the students relevant to their achievement in the five domains of CT skills in quiz 4. The students' responses included in Table 4.16 are also those which were of direct relevance to the main focus and direction of the study, and detailed students' responses are included in (appendix 11).

Table 4.16: Students' CT Responses to the Interview Questions in Quiz-4 (N=5)

Critical Thinking Skills – Quiz 4					
Students	Inference	Recognition of Assumptions	Deduction	Interpretation	Evaluation of Arguments
	Quiz 4	Quiz 4	Quiz 4	Quiz 4	Quiz 4
1	read the question first, understand the question	was clear, assumptions should be identified, cause we had this many times, sometimes no assumptions in the questions, easy to select the correct answer	able to identify extra information, understand the question	Um, reading the text, many times, underlying key words, link to given answers	find connections between the statement and the text, check the truth of the statement, statement could be true but the answer.....conclusion could be wrong
7	identify key words, read the text and make notes, underline expected	planning, reordering events, read and link to answers, some words help to	I understand the question, focus, identify the answer,link the	Um, read statement very carefully, read the text many times, read,	truth, sometimes conclusions are not convincing, which one close by comparing conclusions

	wrong answers	identify the assumptions	conclusion to the statement	compare answers with statement	
19	planning my reading, identify difficult words, read again, ask the teacher. Compare answers with the statement	read the statement many times, take notes, I think before choosing the answer	focus on the question, underline details, find specific information in the statement	I focused on the text, also, then I linked the answers to statement	tried to find if the argument is strongly connected to the statement
20	quickly read the text, underline key words, read the text again, teacher helped to explain difficult words	read the statement many times and compare it with the arguments, find connections between them	identify key information in the statement, compare it to given answers	try to find the relation between the statement and conclusions, also, check the importance of information.	read the arguments with good understanding, asking the teachers for difficult words, and make connections with the statement.
27	I identified the purpose	the concept was clear; we	identify key information in	reading the text but with	identify wrong statements,

	of the question, read the text, then I apply the information	had practiced it many times. I analyzed the given assumptions and relate them the text	the statement, compare it to given answers	expectations, also, I had to remember different ideas in the text	check answers and link them with the text
--	--	--	--	---	---

A general observation of the responses made by the five participants to the interview questions related to quiz 1 clearly suggests that the students lacked a relevant knowledge and ability to use the five CT skills. Whereas, the responses of the same students to the interview questions related to quiz 4 strongly suggest that they were able to answer the CT questions in quiz 4 linked with their higher achievement. The problems encountered by the students in quiz 1 had improved significantly. In response to the interview questions to justify the scores achieved in quiz 4, the participants claimed to have the ability to reflect upon the CT strategies and techniques they had used to answer questions in quiz 4.

Close examination of the students' responses in Table 4.15 related to quiz 1 and Table 10.16 related to quiz 4 suggest that the responses in the initial quiz 1 were weak, tentative, and negative. In contrast, the responses in Table 4.16 related to the final quiz 4 were strong, confident, and positive. It is evident that all the five students have made more or less similar comments both in Tables 4.15 and 4.16 related to quiz 1 and quiz 4 respectively. For example, in relation to inference in quiz 1 student 7 commented, "*many times I read the text, was not clear*"; in response to recognition of assumptions she said, "*can't, reading the text only, try to guess answer*"; responding to deduction skill, *she said, "question was not clear, I read the text, difficult"*; commenting on interpretation skill, she said, "*answers look all correct*"; and responding to the evaluation of argument skill, she said, "*read conclusions but not understand*". However, the responses of the same student 7 in quiz 4 were transformed relatively into more

confident and positive responses of achievement. For example, in responding to inference skill, the student claimed she was able to *"identify key words, read the text and make notes, underline expected wrong answers"*; in relation to recognition of assumptions skill, the student said that she was ready to, *"planning, reordering events, read and link to answers, some words help to identify the assumptions"*; ; referring to deduction skill, she commented that, *"I understand the question, focus, identify the answer,.....link the conclusion to the statement"*; in response to interpretation skill, the student said, *"Um, read statement very carefully, read the text many times, read, compare answers with statement"*; refereeing to evaluation of argument skill, she commented that *"truth, sometimes conclusions are not convincing, which one cloze by comparing conclusions"*.

Another set of responses to interview questions worth mentioning here is related to student number 20, who claimed common ground with student number 7. In response to inference skill in quiz 1, this student said: *"pause" did not understand the text, question was difficult"*; in relation to the recognition of assumptions skill, she confessed that she, *"couldn't control the time, statement wasn't clear"*; reflecting upon the deduction skill the student accepted that *"couldn't focus on major information in the statement"*; in response to interpretation skill question, the student explained that she *"was tensioned, don't know how to start, just guessing"*; in answer to the fifth category of evaluation of argument skill, the student accepted that she *" couldn't read the arguments with appropriate understanding"*.

Like the student number 7, the ability of students 20 to respond to the questions related to the five CT skills improved significantly as evident in her answers to interview question related to quiz 4. Responding to inference skill question the student commented that she *"quickly read the text, underline key words, read the text again, teacher helped to explain difficult words"*; referring to recognition of assumption question, the student said that she *"read the statement many times and compare it with the arguments, find connections between them"*; her response to deduction skill question was that she could, *"identify key information in the statement, compare it to given answers"*. Her answer to interpretation skill question was

that she was able to *"try to find the relation between the statement and conclusions, also, check the importance of information"*; and the student's response to the evaluation of argument skill question was that she was able to *"read the arguments with good understanding, asking the teachers for difficult words and make connections with the statement"*.

Having presented the results of the qualitative data holistically related to the two tables above, it is now time to present the detailed analytical discussion of students' responses to the interview questions relevant to the five CT skills separately that were asked at the end of quiz 1 and quiz 4. These results constitute the main findings relevant to RQ 3.

These responses are the main source of data that answer research question 3. There are many factors related to the challenges involved in applications of CT skills in Omani post-basic schools. These factors have been studied and investigated in this study in order to understand and examine problems associated with the applications of CT skills. For the analytical purposes, themes were extracted Clarke and Braun (2013) from students' responses to the interview question concerning quiz1 and quiz 4 are related to RQ 3, including. The themes extracted include; (1) Inference; (2) Recognition of Assumptions; (3) Deduction; (4) Interpretation; and (5) Evaluation of Arguments. The interviews conducted in phase 2 of this study have sufficiently questioned those challenging factors in view of the relevant studies, findings of phase I, and findings derived from the CT skills test conducted in phase II of this study.

The main concern of this section of study was to conduct a semi-structured interview (Clarke and Braun, 2013). The interview questions integrated research question 3, CT skills test scores relevant to RQ 3, and quiz 1 and quiz 4 results. This gave focus and direction to the interview process in order to get the input required to answer RQ 3. The audio-recorded interviews of the five students' in both quiz 1 and quiz 4 were carefully transcribed and checked and analyzed.

The results of phase I of this study indicated that the Omani teachers lacked the knowledge and application of SQ and CT skills. This result prepared the ground for further investigation of these challenges in phase II and find suitable strategies to address these challenges. The construction of the themes has progressively framed during the course of this long study. The related literature and studies have not provided enough scoop to draw strong conclusions and recommendations related to CT skills in the teaching and learning contexts of Oman. Furthermore, the related literature and studies included in the current study have concluded findings and recommendations focusing mainly on macro-challenges (see chapter five) of CT skills. However, it helped the researcher of this study to develop the framework and the structure of this study which includes practical models and methodologies (Paul and Elder, 2006) of CT processes relevant to teaching and learning context.

To analyze the qualitative data related to RQ3, the researcher used Thematic Analysis Method (Clarke and Braun, 2013). The goal of a thematic analysis (TAM) is to identify themes, i.e. patterns in the data that are important or interesting, and use these themes to address the research or say something about an issue (Clarke and Braun, 2013). According to Clarke and Braun (2013), the TAM has six-phase process which is a very useful framework for conducting this kind of analysis, (1) become familiar with the data, (2) generate initial codes, (3) search for themes, (4) review themes, (5) define themes, and (6) write-up your report. Following the method propounded by Clarke and Braun (2013), all students' responses were dully transcribed, coded, and themed by the researcher. Subsequently, coded responses were classified into the five thematic domains related to RQ 3, namely (1) Inference; (2) Recognition of Assumptions; (3) Deduction; (4) Interpretation; and (5) Evaluation of Arguments.

Finally, hermeneutic interpretation stage (Creswell, 2012) was implemented to confirm the target meaning of each theme and related transcribed statements. The results provided in the following section of the study are generated from the five CT themes mentioned above. These themes form the qualitative data related to RQ 3, which are analyzed and discussed independently in the five different paragraphs in the following section. What follow next are

the five tables covering the five domains of CT skills involving students' responses to the interview questions followed by analytical discussion.

Table 4.17: Students' Responses to Inference Skill in Quiz 1 and 4

Students' Responses in Quiz 1	Students' Responses in Quiz 4
S1: was difficult, just select any answers	S1: read the question first, understand the question
S7: many times I read the text, was not clear. S7: can't, reading the text only, try to guess answer. S7: question was not clear, I read the text, difficult	S7: identify key words, read the text and make notes, underline expected wrong answers S7: planning, reordering events, read and link to answers, some words help to identify the assumptions S7: I understand the question, focus, identify the answer,link the conclusion to the statement
S19: some words were difficult	S19: planning my reading, identify difficult words, read again, ask the teacher. Compare answers with the statement
S20: didn't understand the text, question was difficult. S20: couldn't read the arguments with appropriate understanding	S20: quickly read the text, underline key words, read the text again, teacher helped to explain difficult words S20: read the arguments with good understanding, asking the teachers for difficult words, and make connections with the statement
S27: first quiz and second one were difficult, I read the text but couldn't understand the conclusions	S27: I identified the purpose of the question, read the text, then I apply the information

With regard to inference skill in quiz 1, the responses of the five students who were interviewed confirmed that all of them faced difficulty in drawing inferences from reading the texts. Some students also faced the issue of clarity which led them to guess the meaning. The students' difficulty in inferring and the issue related to the clarity of the text were evident in the students' responses, e.g. the response of student 7 was that *question was not clear, I read the text, difficult* and the response of student 20 was *didn't understand the text, question was difficult*. The level of difficulty and the lack of clarity in using SQ and CT skills at the initial phase of the intervention could be logically linked to their lower level of familiarity to process the reading text using inference skill. On the other hand, students' responses in quiz 4 testify that all the students were able to develop a higher level of inference skill to process the reading texts. This achievement of the five students indicates that they were able to overcome the issue of difficulty and clarity in processing reading text through inference skill at the final phase of the intervention in quiz 4. This result can be supported with the responses of student 7 and student 20. Student 7 categorically expressed *that I understand the question, focus, identify the answer,link the conclusion to the statement* . Similarly, student 20 claimed that he *read the arguments with good understanding, asking the teachers for difficult words, and make connections with the statement*. This significant development of the inference skills of all the five students should be linked with the considerable amount of exposure, training, and practice given in the use of SQ and CT skills during the eight weeks of the intervention phase.

Table 4.18: Students' Responses to Recognition of Assumptions Skill in Quiz 1 and 4

Students' Responses in Quiz 1	Students' Responses in Quiz 4
<p>S1: too much time, frustrating</p> <p>S19: and time was passing very quickly</p> <p>S19: reading but too much information</p> <p>S20: couldn't control the time, statement wasn't clear</p>	<p>S1: was clear, assumptions should be identified, cause we had this many times,</p> <p>S19: planning my reading, identify difficult words, read again, ask the teacher. Compare answers with the statement</p> <p>S19: focus on the question, underline details, find specific information in the statement</p> <p>S20: read the statement many times and compare it with the arguments, find connections between them</p>

Students' responses related to the recognition of assumptions skill in quiz 1 project too main limitations. First constraint was the *limitation of time* to read the text, and recognize the assumptions in the text. The second constraint in completing the CT tasks was related to *anxiety of time needed* in order to be able to read the text, and complete the tasks successfully. The two constrains expressed by the students can be clearly seen in the responses of student 1 and student 19. Student 1 said that it was a matter of *too much time, frustrating*. Student 19 expressed that the *time was passing very quickly; reading but too much information*. The time constraint and the time anxiety involved in quiz 1 resulted in students' inability to successfully read the text, and identify the elements of assumptions involved in the reading text.

On the other hand, students' responses in quiz 4 demonstrate the developed CT abilities and SQ skills. They were able to recognize assumptions in the text through controlled and analytical reading using SQ strategies such as planning, identifying, and comparing. In the final phase of the intervention in quiz 4, the five students were able to mitigate the challenges faced

in quiz 1 to process the reading text for recognition of assumptions.

The developed CT skills and ability to confidently deploy the recognition of assumptions skill in the reading text can be transparently witnessed in the students' responses. Student 1 claimed that the task *was clear, assumptions should be identified, cause we had this many times*. Similarly, student 19 expressed her confidence in *planning my reading, identify difficult words, read again, ask the teacher, compare answers with the statement, focus on the question, underline details, find specific information in the statement*. The considerable improvement in students' skill to recognize assumptions through critical reading of the texts is the result of planned training and practice in using SQ and CT skills provided during the intervention phase.

Table 4.19: Students' Responses to Deduction Skill in Quiz 1 and 4

Students' Responses in Quiz 1	Students' Responses in Quiz 4
<p>S19: only focused on the statement, forgot to connect with the text</p> <p>S19: can't match between the argument and the statement</p> <p>S27: couldn't link different arguments to the main statement</p>	<p>S19: I focused on the text, also, then I linked the answers to statement</p> <p>S19: tried to find if the argument is strongly connected to the statement</p> <p>S27: identify wrong statements, check answers and link them with the text</p>

In the context of the coded students' responses related to deduction skill in quiz 1, only one problem emerged. All the students faced similar difficulty of not being able to link and match different arguments with the main statement in the reading texts. This difficulty can be

overtly seen in the students' responses. For example, student 19 stated that she *can't match between the argument and the statement*. In the similar vein, student 27 confessed that she *couldn't link different arguments to the main statement*. The students were not able to process the reading texts using deduction skills at the initial phase of the intervention simply because the systematic use of the deduction skill had not yet become a part of their schemata. In contrast, the students' coded responses in relation to the use of deduction skills in quiz 4 categorically reflect that the students' had developed deduction strategies to analyze the reading texts. The clear and confident claims made by all the students in their responses given at the final stage of the intervention strongly suggest that they were able to develop and skillfully use the deduction strategies while reading the given texts. Obviously, the deduction skills had become a part of their functional schemata now. Correspondingly speaking, the student 19 claimed that she *tried to find if the argument is strongly connected to the statement* while student 27 was able to *identify wrong statements, check answers and link them with the text*. Thus, it is consistently observable that there is a significant development in the students' deduction skills for processing the text critically which is due to frequently planned teaching and practice given to the students' during the intervention phase from quiz 1 onwards up till quiz 4 in using SQ and CT skills of deduction.

Table 4.20: Students' Responses to Interpretation Skill in Quiz 1 and 4

Students' Responses in Quiz 1	Students' Responses in Quiz 4
<p>S1: too much pressure at the first quiz, reading text too long, lost reading the text</p> <p>S20: was tensioned, don't know how to start, just guessing</p>	<p>S1: able to identify extra information, understand the question</p> <p>S20: try to find the relation between the statement and conclusions, also, check the importance of information</p>

In relation to the use of interpretation skills while reading the texts at the initial stage of the intervention in quiz 1, the students' responses suggest that they were under tension and had felt frustrated to use interpretation skill for meaning making. To illustrate, student 1 faced *too much pressure at the first quiz, reading text too long, lost reading the text*. Similarly, student 20 expressed that he was *tensioned, don't know how to start, just guessing*. The pressure and tension that students faced in applying interpretation skills to reading texts in quiz 1 were evidently due to lack of CT skills of interpreting the core ideas in the texts. However, the situation reversed when these students reached to the final phase of intervention of the study in quiz 4. The students' responses related to the use of interpretation skills while reading the texts in quiz 4 are clearly activity- driven as against their responses in quiz 1, which were activity-opposed. For illustration, student 1 claimed positively that she was *able to identify extra information, understand the question*. Likewise, student 20 stressed that she *was able to try to find the relation between the statement and conclusions, also, check the importance of information*. This transformation in the use of interpretation skills at the end of the intervention phase is led by the consistent developmental approach of providing with and practicing SQ skills and the CT strategies of interpretation during the eight weeks of the intervention program.

Table 4.21: Students' Responses to Evaluation of Arguments in Quiz 1 and 4

Students' Responses in Quiz 1	Students' Responses in Quiz 4
<p>S1: answers were very close to each other; focused on answers not questions</p> <p>S7: answers look all correct</p> <p>S19: answers were same</p>	<p>S1: Um, reading the text, many times, underlying key words, link to given answers</p> <p>S7: Um, read statement very carefully, read the text many times, read, compare answers with statement</p>

	S19: read the statement many times, take notes, I think before choosing the answer
--	--

The perspective related to evaluation of arguments that emerges from the students' responses to the interview questions in quiz 1 is indicative of two main issues, which are lack of clarity and confusion. This resulted in students' failure to evaluating arguments logically. To support this finding, the response of student 1 was that *answers were very close to each other; focused on answers not questions*. Similarly, student 7 expressed her feeling of confusion in saying that *answers look all correct*. The issues of lack of clarity and confusion expressed in students' responses to interview questions in the initial phase of the intervention reveal that the students had insufficient SQ strategies and CT skills to evaluate arguments in the reading texts. In a comparative perspective, the students' responses in the final phase of the intervention at the end of quiz 4 indicate significant development of and improvement in their ability to evaluate arguments in the given reading texts. This achievement can be clearly seen through the claims made by the students in their positive responses to the interview questions in quiz 4. To exemplify, student 1 claimed to *Um, reading the text, many times, underlying key words, link to given answers*. Similarly, student 7 also confirmed that *Um, read statement very carefully, read the text many times, read, compare answers with statement*. This significant growth and development in the students' ability to evaluate arguments using SQ strategies and CT skills is due to cumulative results of involving the students in planned tasks and activities based on evaluation of arguments during the whole intervention phase that lasted for eight weeks, i.e. quiz 1 to quiz 4. Thus far, the findings and discussion of both quantitative and qualitative data has been presented. What follows next is the recapitulation of the main findings relevant to RQ 3 based on the quantitative as well as qualitative results.

iii) Main Finding of RQ3

The overall perspective emerging from the results of the representative sample of the five students in the initial phase of quiz 1 reflects that they were at a lower level of SQ strategies and CT skills. This inadequate level of competence in SQ strategies and CT skills proved to be a barrier for the students to apply the inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments while reading academic texts. However, the planned intervention phase of eight weeks enabled these students to reach to a reasonably higher and adequate level of competence in applying SQ strategies and CT skills for decoding the ideas at the surface level and the ideas between and behind the lines in the reading texts.

The focused CT skills-wise perspective of students' performance from quiz 1 to quiz 4 indicates that there was a gradual, upward progression, and scaffolding of SQ strategies and CT skills while the reading texts. The mean score of inference skill in quiz 1 at the initial phase of the intervention was 1.8, which got considerably improved to the mean score of 3 at the final phase of the intervention in quiz 4. Similarly, the mean score of quiz 1 related to recognition of assumptions skill was 1.8. This mean score became 3 in quiz 4. However, the mean score of deduction skill was at a slightly lower level at 1.4, which was improved to 2.6 in quiz 4 as a result of the interventional progression and scaffolding. Similarly, the mean score of interpretation skill in quiz 1 was 1.4, which got elevated to 2.6 in quiz 4. Like the third and the fourth CT skills, the mean score for evaluation of argument skill was 1.6, which was improved to 2.6 at the final phase of the intervention in quiz 4.

A close examination of the pattern of progression and scaffolding from quiz 1 to quiz 4 suggests that the students' mean score in the inference and recognition of assumptions was 1.8 at the initial level which improved to 3 in the final level. As, these two CT skills are at the lower level in the taxonomy of Paul and Elder's (2006) model and; therefore, they are easier to develop and apply. For this reason, the mean scores for these two CT skills are higher than the other three skills which are at the higher level in the taxonomy. The mean scores of students in the remaining three CT skills of deduction, interpretation, and evaluation of arguments was 1.4 in quiz 1 and 2.6 in quiz 4. As these three CT skills are at the higher level of taxonomy the

mean scores are relatively lower than the scores of the first two CT skills, which corresponds with the level of difficulty involved in these higher level of CT skills. The students in this study scored higher in the lower level of CT skills and lower in the higher level of CT skills from quiz 1 to quiz 4, which lends validity to Paul and Elder's taxonomy. Similarly, the inherent validity of Paul and Elders' model of CT skills in turn lends authenticity to the data and the results of this study. Thus, the data and results used in this study proved to be a credible testing ground for the model. This discussion of quantitative data takes the discussion further to the examination of the qualitative data results.

A close examination of the responses made by the five participants in the structured interview sessions related to the initial phase of intervention in quiz 1 suggests that (a) the students' responses on their performance and achievement in quiz 1 were general, tentative, and weak; (b) all the five participants were at an inadequate level of competence in dealing with the five CT skills, as represented in Paul and Elder's (2006) model of CT and they were not equipped adequately with the strategies to deal with inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments in the reading texts; (c) the element of complaint and frustration is evident in all the responses of the five participants emanating from their inability to exploit the required CT strategies while reading the texts; and (d) the level of intensity of expressing the difficulty and frustration is clearly seen from the lowest level of inference strategy through the highest level of interpretation strategy viewed in a taxonomical framework. Having examined the responses in quiz 1, the researcher intends to examine the responses made by the same participants in quiz 4 in order to develop a comparative perspective, which is as follows.

A careful examination of the responses given by the five participants in the final phase of the intervention of quiz 4, in comparison to the results of quiz 1, indicates that (a) student's responses at the end of the intervention phase in quiz 4 turned out to be specific, confident, and strong; (b) the five participants, without any exception, seemed to be at a reasonably adequate level of competence in using the five CT skills in the reading texts as they found

themselves to be reasonably equipped with the required strategies to deal with inference, recognition of assumption, deduction, interpretation, and evaluation of arguments; (c) the element of complaint and frustration which was evident in quiz 1 results seemed to get neutralized as they had developed abilities to evolve and exploit the required CT strategies while reading the texts; and (d) the level of intensity of expressing the difficulty and frustration in quiz 1 results seemed to have transformed into positive energy and confidence in quiz 4 from the lower level of inference skill through the higher order skill of interpretation.

4.4 Summary and Conclusion

The main concern of phase I of the study was to run the collected data through the two modes of analyses- the statistical analysis and the descriptive analysis. In the first phase of the study, the questionnaire data and the classroom observations data were examined and analyzed using the statistical procedure that involved Pearson correlation coefficients. In the second phase, the data collected through the CT test was subjected to the statistical analysis, in which the t-test was conducted on the data related to the control group and the experimental group both at pre-test stage and the post-test stage to find out the statistical difference between the two groups. After that, the quantitative data collected through the four quizzes and the qualitative data collected through the interviews were subjected to the statistical analysis and thematic analyses and interpretation of coded responses. As a result, critical interpretations and alternative interpretations that emerged from the deeper insights gained from the qualitative responses, were made. Having presented a) the preliminary and introductory input of concern to the present study in chapter 1; b) the survey and review of literature related to the theme and research aims and objectives of the current study in chapter 2; c) the theoretical framework and methodology used to conduct the present study in chapter 3, and ; d) the presentation and analyses of data and the results and findings related to phase I and phase II in chapter 4, it is now the right juncture to discuss the findings, draw conclusions, and propose suitable pedagogical recommendations.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

5.1 Introduction

This last chapter is devoted to drawing pedagogical conclusions from the preceding analytical and findings chapters. Here, the major findings related to the three research objectives have been discussed. The first research objective of this study was to examine the relationship between teachers' knowledge of SQ and their actual use which has been examined through the analyses of the data collected for RQ1. The second research objective of investigating the effect of SQ as a teaching strategy on students' CT was measured through the analysis of data gathered for RQ2. Similarly, the third objective was to find out the CT strategies that were developed, both at macro and as micro levels, by students as a naturally corresponding result of using SQ as a teaching strategy was investigated through the analyses and interpretation of data collected for RQ3.

What follows after this are the concluding remarks and pedagogically important interpretations of the results presented in the preceding chapter. Furthermore, this chapter also explores and highlights the implications and recommendations of the findings, delimitations of the study, and emerging scope for further research.

5.2 The Relationship between Teachers' Knowledge and Actual Use of SQ

The first finding of this study, based on the analysis of the data related to RQ 1, is that a strong, positive correlation was found between the teachers' knowledge of SQ and their actual use of SQ with low levels of teachers' knowledge of SQ, corresponding with lower levels of teachers'

actual use of SQ, as reflected in the result: $r = .79$, $n = 33$, $p < 0.01$. This result suggests that teachers who scored low marks in the knowledge of SQ variable scored lower marks in the corresponding variable of the actual use of SQ in the classroom context. This is indicative of a phenomenon of teachers with low scores in the knowledge variable getting lower scores in the use variable. This disproportionate result related to the first category of probing and illustrating questions, indicates that most of the participants who had a lower knowledge level of probing and illustrating questions, could not use this type of SQ questions when it came to the application stage in the actual classroom context. Given the importance of probing and illustration questions in clarifying the ambiguous ideas and topics and maintaining comprehensive and interactive patterns in the ongoing classroom discourse this seemed to be a question of concern for an explanatory interpretation. Reference to other studies suggests that this result is consistent with other studies which report that many teachers have limited use of probing and illustrating questions, because, this type of questions are always connected with classroom interaction which requires social, cultural and functional language competency to negotiate meaning and clarify the target interaction (Beanmont, 2010; Cates, 2002; and Long and Sato, 1983); which requires a certain level of exposure, experience and resourcefulness.

The researcher assumes that socio-cultural and functional language competency may be a practical constraint responsible for the gap which exists in the knowledge and use of the probing and illustrating questions of the concerned teachers in the study. Wood and Anderson (2001) reiterate that illustrating and clarifying questions with higher level discussions are very important tools that result in a solid and clear comprehension. Another explanation for this disproportionate result could be linked with the teaching approaches used while teaching which could have been top-down approaches led by teacher-fronted instructional practice, which would have controlled the teaching and learning situations leaving insufficient time to negotiate meanings and exchange clarification questions either with the teacher or among the peers themselves.

Another reason for not using probing and illustrating questions while teaching could be that teachers and the students might have felt freer to one another to maintain a non-interactive and non-negotiated patterns. This may be because using probing and illustrating requests and questions in the classroom could be a possible challenge for non-native speakers, teachers with their first degree in which most of the teaching input are based on language learning and not on language acquisition.

Omani socio-cultural perspectives which are mainly led by solidarity-oriented culture as against the formality-oriented culture of native speakers could be another reason for not being able to use probing and illustrating questions (Thakur and Al Mahrooqi, 2015). The frequent use of the speech act of probing is not encouraged from the younger speakers to elder addressees. Since the teachers are non-native speakers and they are also elder to the students, the classroom interaction might be affected by the socio-cultural dynamics and norms of their first language. In solidarity-oriented cultures, such as Oman and other Arab Gulf countries, to use Thakur's (2010) argument, speech acts of questioning and probing from younger participants to elder participants are considered as face threatening acts and thereby impolite in terms of not respecting the age and relative social power of the addressee (pp. 265-66). Due to this socio-cultural constraint, students seem to have been hesitant in using probing questions. This phenomenon calls for alternative interpretation for its implications on pedagogy and research.

The result related to the second category of analytical questions shows that the participants did not use SQ to analyze reasoning and evidence. This result suggests that the target teachers were not competent enough in the use of analytical SQ. This may be due to the reason of having insufficient or inadequate knowledge, training, and preparedness in using analytical SQ. This finding corresponds with other similar research findings for example, Moosa, Suyansah, and Wardatul Akmam's (2015) study, which indicates that there is a weak negative relationship between teachers' knowledge of classroom questions and its actual use in the classroom (p. 293). Furthermore, the teachers in Paul's study (2010) strongly indicated that

they had no clear understanding of the notions that reflect CT; instead they had general knowledge that could not help them to apply appropriate and specific methods to teach CT effectively.

It was also noted in this study that teachers considered the higher-order thinking skills as crucial in the domain of teaching and learning, and therefore, strongly believed that it must be included in the curriculum and also in the related training programs. This finding of the current study and the results of Paul's (2010) study gain further support in many studies such as (Baxter-Magolda, 2008; Paul, 2010; Halpern, 2002; Goldberger and Crow, 2010; and Paul and Elder, 2006). All these studies have strongly recommended that teachers must be exposed to comprehensive training programs and attend each other's workshops and discussion sessions to share and exchange related knowledge and procedures that might be of help for them to develop their competence in teaching CT effectively.

Another contributing factor to this result can be linked with lack of a systematic theoretical framework to guide the effective teaching and assessment of CT. This reason is strongly referred to in "Delphi Report" of Paul and Elder (2006), the underlying assumption of which is that the collective vision of different scholars regarding CT will help to produce a clear methodology that could be used in different domains of teaching and learning of CT in general and analytical questioning in particular.

The result related to the third category of recognition of assumptions indicates that the majority of the teachers in the study were not able to use assumption probing questions. The reason for this could be linked with the findings of McCoy (2008), which suggests that the mismatch between the teachers' questions and students' CT was due to insufficient understanding and inappropriate application of SQ. Another perspective of this result can be traced in Chang's study (2010) in which it was found that the teachers who participated in his study believed that assumption probing questions were time-consuming, difficult, and unable

to push their students to a higher level of thinking. There is a further support to the finding related to the assumption probing questions of this study in studies conducted by Akkaya and Demirel (2012); Gonzalez (2010); and Taylor, Alber, and Walker's (2002), who found that there was no proper correspondence and alignment between teacher questioning and students CT levels.

Another reason for this result related to assumption probing questions could be related to in-service EFL/ESL teachers' frustration emanating from inadequate pre-service pedagogical preparedness in crucial aspects of instructional practice such as SQ. This may have developed a negative attitude and inferiority complex in the concerned teachers which is more likely to have affected their planning and preparation. It is a common knowledge that such micro pedagogical aspects are not a part of mainstream curriculum of EFL/ESL teacher training degree programs. This factor could have been one of the reasons for the mismatch between the teachers' knowledge and its use of assumption probing questions to surface as one of the results in this study in the Omani context.

The next result is related the fourth and fifth category of "deduction" and "evaluation" questions, which suggests that the participants of this study had limited knowledge of questions that are used to evaluate and deduct students' answers and responses. Also, they lacked in the practical experience which allows them to employ this kind of questions effectively. By implication, this result is consistent with Garret's (2006) finding, which suggests that when teachers are able to use deduction questions that evaluate students' answers, the degree of students' achievement is relatively higher and students are able to understand the target tasks.

Paul and Elder (2006) found that deductive and evaluative questions focusing on students answer at different stages of the lesson would help them to develop their thinking skills. In view of this, to use Schwarz, Hershkewitz, and Azmon's (2006) recommendation, the training and experience are very important and deserve more focus and attention in helping

EFL/ESL teachers to use deductive questions that evaluate students' answers effectively and purposefully.

Another possible reason behind this kind of unfavorable result should be that the curricular instruction does not have a strong and integrated focus on posing deductive and evaluative questions that give voice and agency to students in responding to such questions. Put another way, no effective and constructive pedagogical gains can be made in EFL/ESL classrooms without exploiting such higher-order response questions.

One more reason for this result could be the teacher's over-dependence on the use of memory-based questions both in teaching, testing and assessment of EFL/ESL skills. Consequently, the questions that ignite the cognitive process and demand critical responses get marginalized and hence neglected. Such an approach is bound to result in the deficit of CT. This brings the discussion of phase I of the study to its completion and the following section will deal with the discussion of data related to Phase II of this study.

5.3 The Effect of SQ as a Teaching Strategy on Students' CT

The second finding of this study, based on the analysis of the data related to RQ 2, shows significance differences between the two groups in relation to the five CT domains. It was clear that those students who were taught through SQ gained higher scores than the ones who were taught CT in a normal setting.

To illustrate, this finding reveals that the participants had gone through a process of internalization of the knowledge of SQ in the five domains of CT and a gradual scaffolding of this knowledge as a result of planned and explicit instructional input given to the participants of the experimental group. This result is consistent with Garret's (2006) action research study on SQ which demonstrates a significant growth in the students' ability to construct higher-level thinking in terms of using higher level of cognitive questions cognitive level of questions as a result of explicit instructional input and practice. This argument gets desired support from Paul

and Elder's (2006) remark that SQ can be adjusted and applied through various techniques at different stages of comprehension.

The researcher's argument is consistent with the experimental work of Wenning (2006) in which he proposed a typology, which is similar to SQ that can be considered as a "comprehensive treatment kit" to improve CT skills. Though this result of the study is not directly consistent with Forsyth, Paul, Kelley, and McMillan's (2009) study which reviewed 27 studies that examined the result of different programs and courses on CT skills among college students, and they drew a cumulative conclusion which does not support the view that the application of explicit instructional practices lead to a significant increase in CT. And, at the same time, they did support the findings that college attendance improves CT skills. The college attendance linked with the improvement in CT skills reflects the autonomous learning abilities of college students as against the school level students who, in contrast, depend on guided learning and therefore, explicit instructional practice to develop CT skills at school levels remains relevant which finds desired support in Thakur and Al Mahrooqi's (2015) argument that unlike ordinary thinking which is an inborn human ability, CT needs to be taught through implicit and/or explicit instruction. Thus, as an alternative interpretation of Forsyth, Paul, Kelley, and McMillan (2009) study, the researcher intends to claim that the abilities to use CT skills do develop through explicit instructional practices involving SQ strategies at school levels, where the students have not yet become independent autonomous learners unlike the college students referred to in the afore-mentioned study.

The results of Garret's (2006) action research on higher level questioning at SQ clearly demonstrates a significant growth (76%) in the students' ability to construct higher-level thinking in terms of using questions at the upper level of Bloom's taxonomy as a result of explicit instructional input and practice. There is a strong connection of this finding of Garret's study with the findings of the current study as reflected in the results of the four quizzes conducted in phase II. These results show that (a) There was an upward steady growth in terms of the participants' performance scores from quiz 1 to quiz 4, and (b) There was also a

process of internalization of the knowledge of SQ in the five domains of CT and a gradual scaffolding of this knowledge as a result of explicit instructional input given to the participants of the experimental group.

The cumulative results of the four quizzes conducted on the classroom community of 30 Omani students in the experimental group of this study indicate that the continuously given instructional input in SQ led to improved students' output in CT skills. This pedagogical gain is a result of the interactive instructional practice used in the classroom, which enhanced students' level of learning and interaction. In this situation the CT is being used as a means to transform learning and the community, which strengthens the argument of Benesch (1993); Atkinson (1997); and Fox (1994) that social practice is one of the indispensable components of CT in developing information-based community. Having discussed the results of RQ2, the study moves on to the discussion of RQ3.

5.4 The CT Skills Developed from SQ Strategies as a Teaching Method

The third finding of this study, based on the analysis of both quantitative and qualitative data related to RQ 3, shows that there were considerable improvements in the students' knowledge, skills, and competences of SQ strategies and CT skills. The following discussion on the main findings of SQ and CT skills is developed and presented from the holistic context and perspective of the thirty students who participated, as a representative sample, in quiz 1 and quiz 4 related to both quantitative and qualitative results.

First, the functional strategy of students in reading the text and applying them has obviously developed from the level of unplanned reading to the planned implementation of reading strategies whereby, to use Elder and Paul's (2010) terminology, the students who participated in this study have evolved from 'unreflective thinker' to 'practicing reflective thinker'. Second, as a result of this enhanced ability, the confidence and spirit of the students in applying the reading strategies got elevated.

Third, the students were able to demonstrate well-informed awareness and strategy-oriented approach to use SQ and CT skills in reading the texts critically. These positive pedagogical improvements were made by the students by exploiting the CT skills and SQ strategies which were advanced through explicit training and frequent experiences of using CT during phase II. These observations find direct support from Miri, David, and Uri's (2007) data-driven longitudinal case study which found that if higher order thinking strategies are purposely and persistently practiced in the classroom, there is a good chance for a consequent development of CT capabilities (p. 1). Also, this enhancement of CT skills through the eight weeks of students' engagement in the intervention program conducted on the Paul and Elder's (2006) CT skills model based on SQ finds research backing from Pavlenko and Lantoff's (2000) participation metaphor of social constructivist learning.

The analytical discussion of results clearly suggests that at the initial stage of the intervention which is quiz 1, participants' reading was an unplanned activity which was seemingly based on merely a blind guess work devoid of any evolved concrete strategy to process textual information critically, which is in line with Zohar and Dori's (2003) argument that higher order thinking skills and low achieving students are mutually exclusive. Therefore, the responses of the participants reflected negativity and frustration in their performance in the initial phase of intervention results. However, as the intervention process progressed and reached to its final phase, the performance of the students in the class and responses to the interview questions were improved considerably.

The qualitative responses and also the quantitative results of students 1 and 7 in quiz 1 clearly indicate that they were weak and both of them were not able to process the reading texts and deal with higher order demands of the five CT skills. This is very clearly reflected in their qualitative responses in quiz 1, which includes the comments that the reading text and the tasks were: not clear; difficult; time consuming and frustrating. They had to depend on guess work; they were lost in the reading and felt pressured; answers were confusing; and they read the text and tasks without understanding. This difficulty and inability of students 1

and 7 is clearly reflected in their cumulative quantitative total mean scores in quiz 1 which was ranging between 0.33 and 0.40; as a result, these two students have heavily failed in quiz 1, which is indicative of Liaw's (2007) finding that input must be comprehensible to the learner and be offered in such a way as to allow multiple opportunities to understand and use the language. However, this situation improved considerably when these students reached to the last phase of the intervention in quiz 4 as overtly reflected in both the qualitative as well as quantitative results in quiz 4. The two students were able to read and process the reading texts and answer the questions related to the five CT skills as a result of their growth and development of cognitive abilities through continuous comprehensible instructional input and related practice sessions given during the eight weeks of the intervention. This lends support to Liaw's (2007) finding that input must be comprehensible to the learner and be offered in such a way as to allow multiple opportunities to understand and use the language. Further, the qualitative responses in quiz 4 testify this situation of improvement as reflected in their comments that they were able to read and understand, focus on key words, make notes and sift wrong answers. This positive result is in line with Krashen's (1985) input hypothesis and is consistent with the findings of Liaw (2007) that if comprehensible input is provided and the anxiety level is low, then acquisition of learning skills will take place.

Both of them had the ability to read and locate answers, plan, reorder events, and find signal words for assumptions. They were also able to sift extra information, understand, relate, and conclude. In addition, they had the ability to use the strategy of repeated reading and focusing on key words, scaffolding and comparing. Furthermore, these two students had also developed the ability to find connections between the task and the text, assess conclusions, and judge the quality of argument. This finding corresponds with Bialystock's (1990) convection that higher level of proficiency leads to significantly fewer L1-based strategies of L2 learners and their varied use of learning and thinking strategies. This heightened awareness and strategy-oriented abilities that developed in the two students in the domain of the five CT skills are strongly supported by considerably improved performance in quiz 4 of quantitative results, in which their total mean score is ranging between 0.80 to 0.87. with a phenomenal raise of

50% and 38% respectively as compared to their scores in quiz 1. Having compared the qualitative and quantitative data of the students 1 and 7 the researcher, similarly, needs to juxtapose the two sets of data for the students 19 and 20 for further confirmation of effectiveness and validity of the intervention phase. This will be dealt with in the following paragraphs.

The qualitative data results of students 19 and 20 in quiz 1 also reflect their weakness and inability to process the reading texts and deal with the challenging demands of the five CT skills, which simultaneously gets confirmed by the poor results of the quantitative data. This is evident in the tone and tenor of the responses made by the participants of this study in quiz 1. Those comments are: difficult vocabulary and shortage of time; confusing answers; lot of information to process; connection lost and partial focus on the text and missed the link; difficult to match argument and statement; difficult text and difficult question; unclear statement and failed time management; lack of focus on key information; starting difficulty and guess work cause tension; gap between reading and comprehension. This difficulty and inability of students 19 and 20 corresponds with their cumulative total mean score obtained in quiz 1 which was ranging between 0.60 to 0.67. This is due to the incomprehensible input at the initial phase of the intervention stage. This situation supports the claim made in Liaw's study (2007) that the input for the learners should be comprehensible and multiple opportunities of language use should be created and practiced. This would result in lowering the learners' anxiety level and creating a conducive environment for developing and using CT skills.

The qualitative response data of the students 19 and 20 in quiz 4 is realized in numerous encouraging, positive, and confident statements influenced by their own successful performance in the five CT skills. Their responses include; identify key words and infer meaning from the context; repeated reading, note taking, assessing, judging, and finding the answer; sifting the details from the main idea; focus on the text and generate the answer; find and judge a strong connection; repeated reading, compare arguments and link; identify main ideas and compare relevance; link relevant ideas and assess key information; comprehend argument,

identify key words, and link relevant information. This finding is parallel to and compatible with the results of Sahamid's (2016) action research conducted on Malaysian English language upper-secondary school level students using Paul, Willsen, and Binker, (1993) model of SQ.

The present study claimed to find positive application of repeated practice of SQ on students' responses and the consequent development of CT skills. These positive responses are clearly developed as a result of the overall outcome of improvement in their abilities and confidence in handling the reading texts and addressing the five CT skills after the completion of the intervention phase. These responses of success are very closely associated with the scores that students 19 and 20 achieved in quiz 4. Matching with the high spirit of the qualitative responses, the students 19 and 20 scored very high marks in the domains of five CT skills with a total mean score of 1.00. This finding is strongly supported by the findings of Fahim and Bagheri's (2012) study carried out in Iranian Language Institutes in the framework of Bloom et al's (1956) and Paul's CT taxonomies (2006) relevant to SQ. One of their main findings is that students who were taught through active reading using SQ strategies of questioning were able to raise deep questions and actualized CT in classrooms. The quantitative scores established the fact that students 19 and 20 had already developed a functional ability to apply CT skills on the reading tasks and they were able to use them successfully. At the same time, the qualitative responses of these students are clear indicators of their well-informed awareness of the knowledge of relevant strategies required in the five domains of CT skills.

In the preceding discussion so far, the main findings of the quantitative and qualitative data have been discussed in a holistic analytical perspective. Now the discussion turns towards putting forth the main findings of the quantitative as well as qualitative data derived from quiz 1 and quiz 4 in more focused skill-wise format of CT skills.

The quantitative results relevant to inference and recognition of assumptions skills begin with a gradual improvement with an average achievement of the five students who had volunteered to participate in the qualitative data collection through the structured interviews.

The mean scores of these students in quiz 1 for the two CT skills was 1.8 which eventually ended up with a significant achievement with the improved mean scores of 3 in both the skills at the end of the intervention phase in quiz 4. In view of the qualitative results, the average achievement of students with the lower score in quiz 1 corresponds with the issue of difficulty and issue of clarity faced by the students in the inference skill at the initial phase of intervention and similarly the lower score related to recognition of assumption skill in quiz 1 relates with the limitation of time and the issue of frustration at the initial phase of the intervention. Conversely, higher scores of students achieved in quiz 4 reflect their ability and confidence in demonstrating the developed CT abilities and skills which can be seen in the confident responses voiced by the students at the end of the final intervention phase where they were able to deploy developed CT strategies of focusing, identifying the answer and linking the conclusions to the statements in the inference skill and using strategies of planning, identifying and comparing in the recognition of assumptions skill. These results of both quantitative and qualitative data related to inference and recognition of assumptions, starting from the average achievement with difficulty and gradually moving with improvement and scaffolding towards a significant level of application of CT skills, at the end of the intervention phase of quiz 4 testifies to the findings of a number of studies such as Paul and Elder (2009); Muspratti, Luke, and Leonards (2009); Facione (2007); and Egege and Kutieleh (2004), which claim that developing students' CT skills needs more direct and well-planned instructional intervention of CT skills. Furthermore, the students' success in developing from the phase of difficulty in quiz 1 to the confident level of an evolved ability to apply CT strategies at the final phase of intervention in quiz 4 could become possible in a natural course of events because of a planned improvement made in the way instructional and assessment practices were conducted in the framework of using SQ questioning strategies. This could not have been most probably, possible through any other non-Socratic intervention framework. This conforms to the findings of a number of studies such as Ennis, Martin, and Sun (2007); Savery (2006); and David, Baunfield, Steve, Mei, and Jen (2004) which focused on the planned improvement in the course and assessment procedures related to CT skills.

The quantitative results relevant to the CT skills of deduction, interpretation, and evaluation of arguments demonstrate a gradual improvement beginning with an average achievement of the five students in quiz 1, which improved significantly in quiz 4. This improvement was based on explicit instructional practice adopted during the eight weeks of the intervention. The mean score of the students in quiz 1 for the three CT skills were 1.4, 1.4, and 1.6 respectively, which was improved to 2.6 in all the three CT skills. The lower scores in quiz 1 related to the three skills of deduction, interpretation, and evaluation of arguments reflect and justify students' average achievement which is also reflected in their qualitative responses. The qualitative responses in the three CT skills mentioned before suggest that the students were not able to link and match the arguments to the main statements at this stage. On the other hand, the students' higher quantitative scores in quiz 4 reflect that their ability to use deduction, interpretation, and evaluation had become a part of their functional schemata. They developed the ability and confidence in successfully applying the three CT skills in the reading texts. The evidence that the five students had developed a good level of CT strategies to process the reading texts critically can be seen in their well-informed and confident qualitative responses to the interview questions at the end of the intervention phase. The students claimed that they were able to identify wrong statements, check answers, and link answers with the relevant statements in the texts. In addition, they had the capacity to identify extra information, understand the question, find the relation between the statement and conclusions, and also check the importance of the information in the reading texts. Moreover, the students could also read the text many times, focus on underlying keywords, link to given answers, read the statement very carefully, and compare answers with the statement.

This significant growth and development in the students' ability of using SQ strategies and CT skills was due to a cumulative effect of the rigorous implementation of the intervention plan based on the SQ framework. This framework confirms the findings of Kamali and Fahim (2011); and also of Beaumont (2011) that the teaching approach which activates the enrichment of a classroom environment that supports learners' responses to CT by having a deep exposure of discussions, real questions and tasks introduced to learners, and focus on

evidence and justifications to enhance written or spoken claims. In the eight weeks long intervention phase of this study that was conducted in the SQ framework of deep thinking also focused on exposing and subjecting students to deep discussion, real questions, and the CT tasks that demanded evidence of justification to qualify spoken or written claims made by the participants in a social constructivist framework.

Towards the end of this chapter, it is now time to sum up the analytical discussion of the developed SQ and CT strategies thus far based on the triangulated results juxtaposed with the findings of relevant studies. The following transformations in the students' knowledge, skills, and competence in SQ and CT skills have emerged as obvious.

First, the functional strategy of students in reading the texts critically got transformed from the stage of unplanned application to the planned implementation of reading and CT strategies as a result of adopting interactive participatory approach of social constructivism (c.f. Pelech and Pieper, 2010; Onuf, 2013) followed in the eight weeks' intervention program in correspondence with Crookes and Lechner's (1998) argument in their study on students' voice and agency that emphasizes on accomplishing the joint goals of achieving simultaneous development of English communicative abilities and the ability to apply them to develop critical awareness of the world and the ability to act on it to improve matters (p. 320).

Second, as a result of the development of this strategic ability, the confidence and spirit of the students in examining the reading texts critically got elevated. Thus, the students were able to demonstrate well-informed awareness and strategy-oriented approach to process the reading texts analytically and critically. These favorable pedagogical gains in CT were made by the students by exploiting the CT strategies developed through explicitly taught and frequently practiced SQ and CT skills during the experimental intervention phase. This phenomenon is supported by Riffel's (2014) conviction that SQ method with its constant questioning evolves CT strategies (p.25). The developmental performance and consistent improvement of the five

students in all the CT domains gets support from Miri, David, and Uris's (2007) study which strongly suggests that fostering inquiry-oriented thinking and encouraging open-ended discussions lead to consequent development of CT thinking capabilities.

Third, results of growth and development of the participants of this study in the CT skills claim common grounds with the analyses, arguments, and results of the study conducted by Sahmid (2004) who found that in the first two phases students' anxiety was considerably reduced and their perceptions of the English teacher and what is expected of them in English classes were slowly evolved into something more realistic. By the third week and into the fourth stage of Sahmid's study, students were showing signs of adapting to the culture and practice of Socratic questioning. Similar incidents were observed in the participants' progress of this study during the intervention period of quiz 1 to quiz 4, which is further supported by the findings of the studies conducted in the similar contexts by Al Darwish (2012); Robitaille and Maldonado (2015); and Hashemian, Mirzaei, and Abaszadeh (2015).

Fourth, characteristically speaking, the SQ and CT strategies developed by the participants of this study are flexible, spontaneous, diverse, based on individual responses, and authentic. Attainment of such strategies through this experimental study lends, and also gets, strong research support to Thakur's (2016) remark that in line with the socially-aligned view of competence much needed spontaneity, flexibility, and diversity accrues only through a process-centered pedagogy of voice, agency and response, which was also involved in the strategy-oriented and well informed SQ pedagogy approach during the intervention phase of this study.

Fifth, the pattern of progression and scaffolding of SQ and CT skills that emerges from the quantitative and qualitative findings in this study clearly indicates that inference and recognition of assumptions skills belong to the lower level in the taxonomy of Paul and Elder's model (2006). However, the deduction, interpretation, and evaluation of arguments are found to be at the higher level of Paul and Elder's taxonomy. This is justified in the pattern of quantitative results of the students in quiz 1 and quiz 4. The results demonstrate that the

students scored higher in the CT skills that belong to lower level and they scored lower in the CT skills that belong to the higher level in Paul and Elder's (2006) taxonomy depending on the element of difficulty involved. This finding supports the validity of Paul and Elder's model of CT skills and at the same time this study also gains authenticity of the collected data and analytical results derived from it. Thus, authentic data and analysis carried on an established model/framework became a testing ground for the validity and reliability of the model/framework itself. This may be counted as a significant contribution of this study, among others.

The analytical discussion of both the quantitative and qualitative results thus far also establishes that the present study has not only investigated the SQ strategies and CT skills at the macro level but also at the micro level; the area which was not covered by other studies. This study commenced its intervention program with the five broad CT skills included in Paul and Elder's (2006) CT model. In addition, this study adopted a rigorous process-based approach in its intervention program, in which the students' efforts and performance resulted in the successful development and use of micro level CT skills such as reading critically, locating and linking answers to the CT questions.

However, the reasonably comprehensive literature review of SQ and CT skills covered in this study suggests that the studies conducted in Omani contexts as well as in other contexts are limited to the investigation of CT skills only at the macro level following the product-based approach where the main focus was on the five CT skills listed in Paul and Elder's (2006) model. The results and multilevel perspectives developed and discussed in this study focusing on the process-based CT skills clearly advances the knowledge and scholarship of evolved micro level CT skills. This new attempt made in this study is an additional contribution of the researcher to research and the existing literature in the domain of SQ and CT, which may be considered a new contribution of the current study and an addition to the others.

5.5 Delimitations of the Study and Recommendations for Further Research

The delimitations of the study indicate towards the gaps in the present study which offer a fertile ground and scope for designing and undertaking independent studies. Participants in the current study who were exposed to Paul and Elder's (2006) model for CT to read and analyze EFL/ESL reading texts enhanced their skills to think critically and developed effective reading strategies. Whether these findings will be transferred to other learning situations, and to what extent, is open to question and further investigation. One possible line of research could be to conduct a follow-up research on participants who were involved in the current study to find out if participants who are taught through SQ and Paul and for research is Elder's (2006) CT model are able to sustain and retain the CT strategies and abilities evolved for a longer period of time or not. And also, if the participants are more likely to use them in other classroom situations when compared to students in the control group or not could be another variable for investigation.

Since the current study is the first empirical study in Oman that not only applied Paul and Elder's (2006) CT model but also broadened its scope by focusing on the process-based micro-level CT skills, its replication is clearly required for further confirmations of results, although the results of the current study reflect significant gains from integrating Paul and Elder's (2006) CT model into the curriculum.

These empirical studies should be conducted at various grade levels and focusing on different subject matters. The relative results and findings of Paul and Elder's (2006) CT model also need to be assessed against other models. It is possible that direct and explicit teaching approach of this model had as much effect as the actual gains of the model itself. Further research needs to be conducted to find out if other CT models, which might have more implications for use at other grade levels or in other Subject matters, might be equally effective or more effective.

Paul and Elder's (2006) CT model could also be tested using other research methodologies or research instruments. Instead of using the Watson-Glaser Critical Thinking Appraisal Form-A to test for changes in students' ability to reason on everyday subjects, Ennis-Weir Critical Thinking Essay Test could be used. On a theoretical level, researchers need to continue their determinations to reach a broad and unified definition of CT. The correlation between CT in everyday inferences and reasoning, and CT in different areas of expertise also requires further research.

5.6 Implications of the Study

Based on the findings of the study and the insights generated from the entire process of this study the researcher proposes multifarious implications on: (a) instructional practice that involves the teachers and students; (b) policy and decision makers; and (c) syllabus designers, testing and evaluation. The findings of the study could be useful in informing the language instructors, language material developers, language teacher trainers as well as curriculum designers about the importance and role of SQ in the development of CT skills and strategies. The study proposes an integration of SQ techniques of deep thinking and Paul and Elders model of CT skills in the instructional pedagogy of the English language curriculum of grade 12 of Omani post-basic schools. It is also recommended that well-framed professional development training programs on how to develop the knowledge of SQ and CT skills and how to use it effectively in teaching English at grade 12 of post-basic schools be organized on a regular basis.

The current study was conducted in an educational context with variables that are typically associated with Omani post-basic schools, including teachers with different teaching experience and students who came from different learning backgrounds. Findings of this study indicate a significant, and practical results showing that integrating teaching with Paul and Elder's (2006) model in teaching language and CT can develop both students' general abilities to think critically and students' abilities to think within subject area and offers an effective motivation to reflect on possible consequences of integrating Paul and Elder's (2006) model in

different educational curricula. Indeed, the findings of this study concerning the usefulness of Paul and Elder's (2006) model for critical thinking in improving students' abilities to think critically hold important implications for policy makers, educators, and individuals.

From the viewpoints of educators, future employers, preparing students to think critically is among the major responsibilities of the educational system. Critical thinking skills such as analyzing real matters and situations, creating solutions, transferring understanding to new contexts and making connections, and developing justifications for decision-making, are essential for success in real life. Labor market demands CT skills and abilities from the individuals in the society. A democratic society should afford and ensure sufficient training and fair learning opportunities for all individuals in the society to be trained for CT. If business leaders really want their employees to have CT skills and abilities and if policy makers need individuals to be able to think critically, they must direct educational institutions to integrate explicit instruction in CT across different levels of learning in all academic fields.

For educators, having both the awareness of learning to think critically and appropriate methods of teaching and learning through which this can be achieved are indeed important. There is a limited evidence to indicate that students will improve in their CT skills and abilities simply because they attend classes, even if the teacher uses CT in planning his/her courses. However, there is a strong evidence, including in the current study, to confirm that if we want our students to use and apply CT skills and strategies, we must explicitly teach and train them to do so.

In this study, training given in CT was both explicit and intense. Similarly, to develop as critical thinkers, learners must be exposed to the components of the CT model thoroughly and explicitly, and they should have frequent practices in applying the model. Paul and Elder's model (2006) needs to be comprehensively integrated into course content and learning outcomes, not just presented or introduced in isolation during the semester. Implicit and indirect modeling of CT through few lessons with CT activities is not likely to be effective for many learners.

The most important implication of the current study lies in recognizing and identifying the need for direct and explicit training and teaching for CT. The results of this study also suggest that sufficient training and awareness are essential for teachers to successfully infuse Paul and Elder's (2006) model into course syllabus and learning outcomes.

Paul and Elder's model (2006) is theoretically flexible and rich, however, it needs deep understanding and practical experience in order to adapt its principles and strategies to course materials and course content. It is an approach that needs, for most practitioners, a willingness to reflect deeply on course materials, course content, course syllabus, target students and, assessment procedures to achieve effective application of the model and intended learning outcomes. The application of Paul and Elder's (2006) model of CT will be more effective and result-oriented if the process-based micro level CT skills are targeted along with the macro level skills during the course of instructional input and practice sessions. In this context, it is strongly recommended for the English Language Teachers in EFL contexts to create constructivist learning environments and exploit them using participatory mechanisms in the classrooms, as suggested by Pelech and Pieper (2010) and Onuf (2013), which are strongly believed to be ideally poised and positioned to promote students' independent thinking, voice, and agency.

Two of the novel implications of the present study that have emerged as significant are that (a) In order to significantly add to the existing literature on SQ and CT skills with a new dimension is to carry out a comprehensive research on CT skills at the micro level to balance the existing gap and disparity between the available literature on the macro versus micro levels of CT skills, which will open doors for further researches, and (b) In view of the dynamic roles that socio-cultural dimensions play in solidarity-oriented cultures as against formality-oriented cultural norms in performing the speech acts of questioning and probing, and similarly other speech acts, one of the important implications of this phenomenon on research would be to conduct studies based on the hypothesis built around the felicity conditions of speech act that make the speech act either felicitous or infelicitous. This might result in more focused new

findings which may lead to useful pedagogical implications.

5.7 Summary and Conclusion

This study employed multi-method data collection, which focused on Omani female teachers' applied knowledge of SQ to enhance students' CT skills in Dhofar region. The study was conducted and completed in two phases using (a) The educational philosophy of 'Socratic Questioning' led by the premise of 'Deep Thinking' (Keng, 1996; Kenzik, Wubbels & Hajer (2010)); (b) the 'Reader Response Theory' premised on the process of 'Engagement with the Text' (Rosenblatt, 1978; Rosenblatt, 1995); (c) the 'Constructivist Approach to Language Competence' built on the central argument that we make sense of our experiences by 'Constructing our own Mental Models' to interpret new experiences in reading the texts (Piaget, 1966, 1972, 1974; Vygotsky 1962; Burden 1997; Seifert, 1997; Reagan, 1999; Osborn 2000; Pelech and Pieper, 2010; Onuf, 2013); and (d) Paul and Elder's (2006) 'CT Skills Model' as a framework to study the knowledge and practice of Socratic Questioning in the ESL/EFL contexts.

In Phase I of the study, the researcher examined the relationship between teachers' knowledge and their actual use of SQ. In phase II, an intervention was implemented to help teachers' enhance students' CT. This intervention was mainly based on Paul and Elders' (2006) CT model, which was also guided by several relevant theoretical orientations and perspectives, in addition to the educational philosophies highlighted in the preceding paragraph, including Philosophical Theories of Ennis, 2008; Siegel and Biro; 2008; Paul and Elder, 2006; and Lipman, 2003), and Psychological Theories of Walker, Brophy, Hodge and Bransford, 2006; Sternberg, Roediger and Halpern, 2007; and Halpern, 2002). In phase I, the researcher administered a quantitative questionnaire and conducted classroom observations via checklist to elicit the required data from 230 participants. In addition, the researcher applied CT test, conducted eight quizzes, and took five interviews in phase II to elicit data from 30 participants of the experimental group.

The main objective of this data-driven study was to investigate the gap(s) between teachers' knowledge and their actual use of SQ in order to get research-based insights into the problem under examination to help develop and enhance students' CT skills. The researcher used both quantitative and qualitative data aimed at studying participants in their daily teaching/learning settings.

The finding related to RQ1 suggests a strong positive relationship between teachers' knowledge and their actual use of SQ which indeed helped to demonstrate the real state of affairs and the issues of concern involved in teaching CT skills. This positive relationship led to the investigation of RQ 2 in the next phase of this study by determining the appropriate intervention which enabled teachers to enhance students' CT skills. This was achieved by integrating and aligning the textbooks and teaching materials with SQ techniques and strategies which helped students to develop their own thinking strategies during the weekly quizzes.

The most prominent finding relevant to RQ2 which emerged from the statistical analyses of the t-test is that teaching CT skills through SQ strategies is beneficial which proved to be more productive. By implication, the researcher claims that students who are taught through SQ strategies are better able to develop and use the target skills in the five domains of CT.

Another important finding establishes the idea that continuous instructional practice leads to upward and steady growth in the effective use of CT which results in the internalization and gradual scaffolding of knowledge, skills, and competence of CT during operationalizing the intervention phase in an integrated, interactive social constructivist framework. Furthermore, an additional related finding also suggests that CT skills do develop through explicit instructional practice of SQ strategies at school level, where students have not yet become independent and autonomous learners as against the tertiary level students. Concerning the findings relevant to RQ3, based on the quantitative results which suggest that all the five students demonstrated a gradual improvement in the five domains of CT skills in their performance from quiz one to four.

Furthermore, the results of the qualitative data suggest that all the five students had developed an informed understanding and awareness about the SQ strategies which they had gained in order to develop and enhance their CT skills in all the five domains. Furthermore, the findings of triangulation of the data collected through the quantitative and qualitative instruments suggest that (a) at the end of the eight week intervention phase, the participants had developed the knowledge and awareness of, and the strategy-oriented abilities to use CT skills;(b) the performance of the students had improved consistently to more or less at a similar exit level in quiz 4;(c) the tone and tenor of the qualitative responses related to quiz 4 indicates towards a significant reduction in participants' frustration and lack of confidence, which prevailed at the level of quiz 1 and 2;(d) a significant raise was evident in the level of enthusiasm and confidence as all the five students placed themselves at a reasonably adequate level of competence in handling the CT skills in reading the texts.

Moreover, the consistency of the results obtained through the multi-method instruments used in this study and the scientific triangulation of results testify to the fact of acceptable degree of reliability and validity of the research design, the analytical procedures adopted, findings of the study, and also the validity of Paul and Elder's model and social constructivist approach to language learning and developing critical thinking which have been carefully selected and strategically used in this study.

REFERENCES

- Akkaya, N. & Demirel, M. 2012. Teachers' candidates' use of questioning skills in during-reading and post-reading strategies. *Procedia - Social and Behavioral Sciences*, 46:4301-4305. <http://dx.doi.org/10.1016/j.sbspro.2012.06.244>.
- Al-Busaidi, K. 1995. *English in the labor market in multilingual Oman with special reference to Omani employees*. Ph D, dissertation, University of Exter, U. K.
- Al-Darwish, S. 2012. The role of teacher questions and the Socratic method in EFL classrooms in Kuwait. *World Journal of Education*, 2 (4):76-84.
- Alder, J. 1982. *The paidea proposal*. New York: Macmillan.
- Al-Issa, A. S. & Al-Balushi, A. H. 2012 . English language teaching reform in Sultanate of Oman: A case of theory and practice disparity. *Educational research policy practice*. Doi.10.1007/s/10671-011-910-0.
- Al-Issa, A. S. 2010. Factors influencing critical thinking in the ELT context. In R. Al-Mahrooqi & V. Tuzlukova (Eds.) *The Omani Symphony: Maintaining linguistic and socio-cultural equilibrium*: Sultan Qaboos University Academic Publication Board. pp. 169-202.
- Al-Jardani, K. S. S. 2012. Evaluating a developed framework for curriculum evaluation in Oman. *International Journal of English Linguistics*, 2(6):1.25. doi.10.5539/ijel. v2n6p17.
- Al-Kindi, N. S. and Al-Mekhlafi, A. M. 2017. The practice and challenges of implementing critical thinking skills in post-basic EFL classrooms. *English Language Teaching*, 10 (12):116-133.
- Allen, J. M., Butler-Mader, C. & Smith, R. A. 2010. A fundamental partnership: The experiences

- of practicing teachers as lecturers in a pre-service teacher education program. *Teachers and Teaching: Theory and Practice*, 16 (5):615-632.
- Allwright, D. & Bailey, K. 1991. *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge: Cambridge University Press.
- Al-Mahrooqi, R. 2012. English communication skills: How are they taught in schools and universities in Oman. *English Language Teaching*, 5(4):125-130. Doi.10.5539/elt.v5n4p124.
- Angelo, T. & Cross, K. 1993. *Classroom assessment techniques: A handbook for college teachers*. USA: Jossey-Bass, Inc.
- Assessment Day, 2017. Free critical thinking test inferences: Solutions booklet. Available at atinfo@assessmentday.co.uk
- Atkinson, D. 1997. A critical approach to critical thinking in TESOL. *TESOL Quarterly*, 31(1):75-79.
- Austin, J. 1962. *How to do things with words*. MA: Harvard University Press.
- Backett, G. & Slater, T. 2005. The project framework: A tool for language content and skills integration. *ELT Journal*, 59 (2):108-116.
- Bailystock, E. (1990). *Communication strategies*. Oxford: Blackwell.
- Bardhan, S. K. 2013. Grammar-item focused materials for use in the ESL class: A pragmatic approach. *Yashashri: International Journal of English Language and Literature*, VI (2):93-99.
- Bauersfeld, H. 1995. The structuring of the structures: Development and function of mathematizing as a social practice. In L. P. Steffe & J. Gale (Eds.). *Constructivism in Education* (pp. 156 – 172). Hillsdale: Lawrence Erlbaum Associates Publishers.

- Baxter-Magolda, B. 2004. *Making their own way: Narrators for transforming higher education to promote self-development*. USA: Stylus Publishing, LLC.
- Baxter-Magolda, B. 2008. Three elements of self-authorship. *Journal of College Students Development*, 49(4):0897-5264.
- Beanmont, J. 2010. A sequence of critical thinking tasks. *TESOL Journal*, 1 (4):427-448.
- Beck, I. L. &McKeown, M. G. 2002. Text talk: Capturing the benefits of read aloud experiences for young children. *The Reading Teacher*, 55:10-20.
- Benesch, S. 1993. Critical thinking: A learning process for democracy. *TESOL Quarterly*, 31(3):545-547.
- Beretta, A. 1986. Toward a methodology of ESL program evaluation. *TESOL Quarterly*, 29, 144 – 155.
- Bialystok, E. 1990. *Communication strategies: A psychological analysis of second language use*. Oxford: Basil Blackwell.
- Bigelow, M. & Ranney, S. 2005. Pre-service teachers' knowledge about language and its transfer to lesson planning. In N. Bartles (Ed.). *Applied linguistics and language teacher education* (pp. 179-200). New York: Springer.
- Bloom, B., Engelehart, M., Furst, E., WILL, W. & Karthwhol, D. (Eds). 1956. *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain*. New York: David McKey.
- Borg, S. (Ed.). 2006. *Classroom research in English language teaching in Oman*. Sultanate of Oman: Ministry of Education.
- Brickhouse, T. & Smith, D. 2007) *Socrates on how wrongdoing damages the soul*. Available at lclark.edu/~ndsmith/45lsyl.htm.

- Cambridge University Press 2011. *The weighted mean*. Retrieved October 9, 2018, from http://www.cup.cam.ac.uk/resources/0521844282/4321_WeightedMean.pdf
- Carl, H. 2010. Cognitive education: A transactional metacognitive perspective. *Journal of Cognitive Education and Psychology*, 9(1):21-35.
- Carter, R. A. 1997. *Investigating English discourse*. London: Routledge.
- Cates, K. 2002. Teaching for better world: Global issues and language education. *Human Rights Education in Asian Schools*. Asia-Pacific Human Rights Information Center, Osaka 5:1-17.
- Chang, Y. 2010. *Students' perceptions of teaching styles and use of learning strategies*. Master of Science Dissertation, University of Tennessee.
- Chaudron, C. 1988. *Second language classrooms: Research on teaching and learning*. London: Cambridge University Press.
- Chenault, T. & Orsel, E. 2008. An act of translation: The need to understand students' understanding of critical thinking in the undergraduate classroom. *The Journal of Effective Teaching*, 8 (2):5-20.
- Cheng, M. M. H., Cheng, A. Y. N., & Tang, S. Y. F. 2010. Closing the gap between the theory and practice of teaching: Implications for teacher education programs in Hong Kong. *Journal of Education for Teaching*, 36 (1):91-104.
- Chin, C. 2006. Using self-questioning to promote pupils' process skills thinking. *School Science Review*, (321):113-123.
- Clarke, V. & Braun, V. 2013 Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The Psychologist*, 26(2):120-123.
- Creswell, W. (2012). *Educational research: Planning, conducting, and evaluating quantitative*

and qualitative research. Boston, UAS: Pearson Education, Inc.

Crookes, G. and Lehner, A. L. 1998. Aspects of process in an ESL pedagogy teacher education course. *TESOL Quarterly*, 32 (2):318-328.

Crossman, Ashley, 2018. *Understanding Purposive Sampling: An Overview of the Method and Its Applications*. Retrieved from <https://www.thoughtco.com/purposive-sampling-3026727> on January 2019.

Dagli, M. 2008. *Integrating critical thinking skills into planning and implementation of teaching Turkish: A comparative case study of three teachers*. A PhD Thesis: Middle East Technical University.

David, M., Baumfield, V., Steve, H., Mei, L. & Jen, M. 2004. *Thinking skill frameworks for Post-16 learners: An evaluation*. A Research Report for the Learning and Skills Research Centre, 1-160.

Deming, W. 1990. *Sample design in business research*. Vol. 23. London: John Wiley & Sons.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to deductive process*. Boston: Boston Heath.

Egege, S. & Kutieleh, S. 2004. Critical thinking: Teaching foreign notions to foreign students. *International Educational Journal*, 4 (4):75-85.

Eggen, P. & Kauchak, D. 2002. *Strategies for teachers: Teaching content and thinking skills* (4th Ed.). Needham Heights: M.A. Allyn and Bacon.

Elder, L. & Paul, R. 2010. *Critical thinking development: A stage theory*. Retrieved on 15 October 2018 from www.criticalthinking.org

Elliot, A. 2006. Hierarchical model of approach-avoidance motivation. *Kluwer Academic Publishers*, 30 (2):111-116.

- Enger, S. & Yager, R. 2009. *Assessing students' understanding in science: A standards-based K-12 handbook*. New York: Library of Congress Cataloging in Science.
- Ennis, C., Martin, R. & Sun, H. 2007. Situational interest: A curriculum component enhancing learning in physical education. *Nova Science Publishers, Inc.* pp. 51-77.
- Ennis, R. & Weir, E. 1985. *The Ennis-Weir Critical Thinking Test*. CA: Midwest, Pacific Grove.
- Ennis, R. H. 2008. Nationwide testing of critical thinking for higher education: Vigilance required. *Teaching Philosophy*, 31(1):1-26.
- Evans, C., Harkins, M. J. & Young, J. D. 2008. Exploring teaching styles and cognitive styles: evidence from school teachers in Canada. *North American Journal of Psychology*, 10 (3):567-573.
- Facion, P. 2007. *Critical thinking: What is and Why it counts?* California Academic Press, 1-20.
- Fleiss, J. L. 1981. *Statistical Methods for Rates and Proportions*. Second Edition. New York: Wiley
- Folsom, C. 2009. Teaching for intellectual and emotional learning (TIEL): Bringing thinking and moral-ethical learning into classrooms. In D. Ambrose & T. Cross (Eds.). *Morality, ethics and gifted minds (pp. 1 – 16)*. New York: Springer Publishing Company.
- Forsyth, D. R., Paul, S., Kelley, K., & McMillan, J. H. 2009. What causes failure and success? Students' perceptions of their academic outcomes. *Social Psychology of Education*, 12 (2):157-174.
- Foster, P. 1996. *Observing schools: A methodological guide*. London: Paul Chapman.
- Fox, H. 1994. *Listening to the world: Cultural issues in academic writing*. Urbana, IL: National Council of Teachers of English.

- Fraenkel, J. & Wallen, N. E. 2000. *How to design and evaluate research in education (4th Edition)*. Boston: McGraw-Hill.
- Friere, P. and Macedo, D. 1987. *Literacy: Reading the word and the world*. Connecticut: Bergin and Garvey.
- Gadzella, B. M., Stacks, J., Stephens, R. C., & Masten, W. G. 2005. Watson- Glaser critical thinking appraisal, Form-S for education majors. *Journal of instructional psychology*, 32 (1):9-12.
- Gambrill, E. 2005. *Critical Thinking in Critical Practice: Improving the Quality of Judgment and Decisions*. New Jersey: Wiley & Sons Press.
- Garret, W. 2006. *Developing Sixth Grade Gadflies: An Action research project presented to the Shawnee Mission Board of Education*. Bluejacket-Flint Elementary.
- Gayle, B. M., Preiss, R. W., & Allen, M. 2006. How effective are teacher-initiated classroom questions in enhancing student learning? In B. M. Gayle, R. W. Preiss, N. Burrell, & M. Allen (Eds.), *Classroom Communication and Instructional Processes* (pp. 279–293). Mahwah: Lawrence Erlbaum Associates.
- GianCarlo, C. 2004. Assessing secondary students' dispositions towards critical thinking. *Development of California Measurement*, 64 (2):347-364.
- Glaserfeld, E. V. 1987. *The construction of knowledge, contributions to conceptual semantics*. USA: Intersystems Publications.
- Goldberger, J. & Crow, J. 2010. Gender inequality within the U.S. land-grant agricultural sciences professoriate. *International Journal of Gender, Science and Technology*, 2(3), 23-34.
- Goddard, A. and Mean, L. 2009. *Language and Gender*. London and New York: Routledge.

- González, A. L. 2010. Researching classroom questioning. *Encuentro*, 19, 52-59.
- Gorard, S. 2002. Political control: A way forward for educational research? *British Journal of Educational Studies*, 50(3):378-389.
- Greene, C. 2005. The generative potential of mixed methods inquiry. *International Journal of Research & Method in Education*, 28(2):207-211, DOI: 10.1080/01406720500256293
- Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. C. 1995. *Multivariate data analysis (5th Ed.)*. New Jersey: Prentice-Hall International, Inc.
- Halliday, M. A. K. 1999. Language and reshaping of human experience. *Special Lecture*. Hyderabad: Central Institute of English and Foreign Languages.
- Halpern, F. 2002. Teaching for critical thinking: Four-part model. In Davis, S., Buskist, W. (Eds.). *The teaching of psychology* (PP 160 - 182). New York: Psychology Press.
- Halpern, F. 2014. *Critical thinking across the curriculum: A brief edition of thought & knowledge*. USA: Routledge.
- Hammond, J. (2006). High challenge, high support: Integrating language and content instruction for diverse learners in an English literature classroom. *Journal of English for Academic Purposes*, 5(4):269-283.
- Hashemian, M., Mirzaei, A. Abaszadeh, M. 2015. Relationship between critical thinking and compensatory strategies by Iranian upper-intermediate EFL learners in oral communication. *Mediterranean Journal of Social Sciences*, 6(6):11-18.
- Hayman, R. (1979). *Strategic questioning*. Englewood Cliffs, New Jersey, Prentice Hall, Inc.
- Hoa, N. 2008. Developing EFL learners' intercultural communities competences: A gap to be filled. *The Philippine ESL Journal*, 1, 29-56.

- House, J. and Kasper, G. 1981. Politeness markers in English and German. In F. Coulmus (Ed.) *Conversational Routine*, Vol. 2, The Hague (pp. 157-185). The Netherlands: Mouton Publishers.
- Impara, J. & Plake, B. (Eds.). 1998. *The thirteenth measurements yearbook*. Lincoln: The University of Nebraska Press.
- J. Doughty and Michael H. Long. 2005. *The handbook of second language acquisition*. California. Blackwell Publishing.
- Jariah-Mohd J. & Talif R. 2005. Questioning strategies and construction of context in classroom talk. *The English Teacher*, XXXIV, 76-89.
- Johnson, R., 2008. Making sense of "Informal Logic". *Informal Logic*, 16 (29):1-30.
- Justin A., DeSimone, P. D., Harms & Alice, J. 2015. Best practice recommendations for data screening. *Journal of Organizational Behavior*, 36 (2), pp. 171-181; doi: 10.1002/job.
- Kamali, Z. & Fahim, M. 2011. The relationship between critical thinking Ability of Iranian EFL learners and their resilience level facing unfamiliar vocabulary items in reading. *Journal of Language Teaching and Research*, 2 (1):104-111. Retrieved on 25.10.2017.
- Kasper, G. 2006. Beyond repair: Conversation analysis as an approach to SLA. *AILA Review*, 19 (1), 83-99.
- Kelly, K. 2009. Assessing students' critical thinking performance: Urging for measurement using multi-response format. *Journal of Science direct*, 4(1): 70-76.
- Keng, L. 1996. Critical thinking and Socratic inquiry in the classroom. Australian Association for research in Education. Paper presented at the joint ERA, AARE Conference, Singapore, 1996.
- Kenzic, D., Wubbels, T., Elbers, E. & Hajar, M. 2010. The Socratic dialogue and teacher

- education. *Teaching and Teacher Education Journal*, 26(4):1104-1111.
- Kip, A. 2002. Teaching for better world: Global issues and language education. *Human Rights Education in Asian Schools*, 1-17, Retrieved on January 11, 2017 from www.hurights.or.jp
- Kite, E., Stockedale, D., Whitley, E. & Johnson, T. (2005). Attitudes toward younger and older adults: An updated meta-analytic review. *Journal of Social Issues*, 61(2):241-266.
- Krashen, S. D. 1985. *Input in second language questions*. Oxford, UK: Pergamon Press.
- Kuhn, D. 2005. Education for thinking, library of congress cataloging-in-publishing data, Harvard College language learners. *The Modern Language Journal*, 83, 193-201.
- Kumar, r., and James, R. 2015. Evaluation of critical thinking in higher education in Oman. *International Journal of Higher Education*, 4(3):33-43.
- Leighton, J. & Geiri, M. 2007. Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' teaching processes. *Educational Measurement Journal*, 26(2):3-16.
- Liaw, M. 2007. Content-based reading and writing for critical thinking skills in an EFL context. *English Teaching and Learning*, 31(2):45-87.
- Lipman, M. 2003. *Thinking in education: Approaches in teaching for thinking*. Cambridge, UK: Cambridge University Press.
- Lombard, B. & Grosser, M. 2004. Critical thinking abilities among prospective educators: Ideas versus realities. *South Africa Journal of Education*, 24 (3), 212-216.
- Long, M., & Sato, C. 1983. Classroom foreigner talk discourse: Forms and functions of teachers' questions. In H. Seliger, & M. Long (Eds.). *Classroom-oriented research in second language acquisition* (pp. 268-286). Rowley, MA: Newsbury House

- Long, M. H. 1996. The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia Eds.) *Handbook of research on language acquisition*, 2:413-468. San Diego, CA: Academic Press.
- Lotter, C., Harwood, W. & Bonner, J. 2007. The influence of core teaching conceptions on teachers' use of inquiry teaching practice. *Journal of Research in Science Teaching*, 44(9):1318-1347.
- Luk, J. C. M. 2013. Forms of participation and semiotic mediation in board games for second language learning. *Pedagogies: An International Journal*, 8(4):352–368.
- Luk, J. C. M., & Lin, A. M. Y. 2007. Classroom interactions as cross-cultural encounters: Native speakers in EFL lessons. Mahwah: Lawrence Erlbaum Associates.
- M. Lightbown, Nina Spada. 1999. How languages are learned. Oxford. Oxford University Press.
- M. Teresa, Angel, D., Jose, L., Pedro, S., & Anthony, J. 2018. Revisiting the differences between mixed methods and multi-methods: Is it all in the name. *Qual Quant*. Retrieved from <https://doi.org/10.1007/s11135-018-0700-2>
- Mansoor, F., & Mohammad, B., 2012. Fostering critical thinking through Socrates' questioning in Iranian language institutes. *Journal of Language Teaching and Research*, 3(6):1122-1127.
- MaPeck, J. 1981. *Critical Thinking and Education*. New York, USA: Martin's Press.
- Martin, R. 2010. Bells that Still Can Ring: Systems Thinking in Practice. In: *1st International Workshop on Complexity and Real World Application Using the Tools and Concepts from Complexity Science to Support Real World Decision-Making Activities*, (pp. 21-23), July, England, UK.
- Marzano, R., Pickering, D. & Pollock, J. 2003. *Classroom instruction that works: Research-based*

strategies for increasing students' achievement. Library of Congress Cataloging-in-Publication Data, USA.

Masatoshi, S. & Susan, B. 2016. *Peer Interaction and Second Language Learning*. John Benjamins, vii, 399 pp.

Mason, J. 1996. *Qualitative researching*. Thousand Oaks, California: Sage Publications.

McCormick, K. 1994. *The culture of reading and the teaching of English*. Manchester: Manchester University Press.

Meckenzie, I. 2002. *Paradigms of reading*. Basingstoke: Palgrave Macmillan.

McCoy, M., 2008. *Plato on the Rhetoric of Philosophers and Sophists*. Cambridge: Cambridge University Press, UK.

McMillan. H. 1987. Enhancing college students' critical thinking: Download PDF

Mehta, S. R. & Al-Mahrooqi, R. 2014. Can thinking be taught? Linking critical thinking and writing in an EFL context. *RELC Journal*, 46 (1), 23-36. Bezemer, J., & Mavers, D. (2011). Multimodal transcription as academic practice: A social semiotic perspective. *International Journal of Social Research Methodology*, 14(3):191–206.

Mehta, S. R., Al-Mahrooqi, R., Denman, C. & Al-Aghbari, K. 2018. Assessing Omani university entrant's critical thinking skills with the Cornell class reasoning test form X. *Pertanika Journal of Social Sciences and Humanities*, 26 (4), 2229-2242.

Meskill, C., Anthony, N., Hilliker-VanStrander, S., Tseng, C. H. & You, J. 2006. Expert-novice teacher mentoring in language learning technology. *Teacher education in CALL*, 14, 283-298

Ministry of Education, Oman 2008. *Basic education, English language curriculum framework*. Curriculum Development Press.

- Ministry of Education, Oman. 2008. *International conference on the reform of secondary education*, 22-24 December 2008, Muscat, UNESCO, Conference papers.
- Miri, B., David, P. & Uri, Z., 2007. Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in Science Education*, 37(4):353-369, DOI: 10.1007/s11165-006-9029-2.
- Moor, L. & Rudd, R. 2002. Using Socratic questioning in the classroom. *ProQuest Education Journals*, 75(3):24.
- Moosa, A., Suyansah, S, and WardatulAkmam, D. 2015. Perceptions and use of question types by TEFL teachers in Oman: A pedagogical study. *European Journal of Social Sciences*, 49(3):287-306.
- Muspratti, S., Luke, A. & Leonards, P. 2009. *Constructing critical literacies: Teaching and learning textual practice. National Centre for Vocational Education Research* (pp. 102-155). Hampton Press.
- Norris, S. & Ennis, R. 1989. *Evaluating critical thinking*. Pacific Grove, CA; Midwest Publications.
- Norris, S., Leighton, J. & Philips, L. 2004. What is stake in knowing the content and capabilities of children's minds? *Journal of Theory and Research in Education*, 2 (3):282-308.
- Nunn, R. and Sivasubramaniam, S. (Eds.). 2011. *From defining EIL competence to designing EIL learning*. South Korea: Asian EFL Journal Press, A Division of Time Taylor International.
- Nunan, D. 2003. The Impact of English as A global language on educational policies and practice in Asia-Pacific Region. *TESOL Quarterly*, 37(4):589-613.
- Oakleaf, M. 2008. Dangers and opportunities: A conceptual map of information literacy assessment approaches. *The Johns Hopkins University Press*, 8(3): 233-253.

- Onuf, N. G. 2013. *Making sense, making words: Constructivism in social theory*. London: Routledge.
- Osborn, T. 2000. *Critical reflection and foreign language classroom*. Westport, Connecticut: Bergin and Garvey.
- Pallant, J. 2007. *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows version 15*. UK, USA: Open University Press Milton Keynes.
- Parvize, B. & Marzieh, B. 2010. The relationship between Iranian EFL teachers' critical thinking ability and their professional success. *English Teaching*, 3(2):1-17.
- Paul, R., & Elder, L. 2006. The art of Socratic questioning: The foundation for critical thinking. Retrieved on 25 March 2017 from www.criticalthinking.org
- Paul, R., Willsen, J., & Binker, A. J. A. 1993. Critical thinking: what every person needs to survive in a rapidly changing world. Revised third edition. Santa Rosa, CA: Foundation for Critical Thinking.
- Paul, S. 2010. A survey of attitudes towards critical thinking among Hong Kong secondary school teachers, *Journal of Thinking Skills and Creativity*, 6 (1):14-23.
- Pavlenko, A. and Lantoff, J. P. (2000). Second language learning as participation and the (re) construction of selves. In J. P. Lantoff (Ed.), *Sociocultural theory and second language learning* (pp. 155-177). Oxford: Oxford University Press.
- Pedhazur, E. J., & Schmelkin, L. P. 1991. *Measurement, design, and analysis: An integrated approach*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Pelech, J. and Pieper, G. 2010. *The comprehensive handbook of constructivist teaching*. Charlotte, N. C.: Information Age Publishing.
- Peterson, C. 2003. Bringing ADDIE to Life: Instructional Design at Its Best. *Journal of*

Educational Multimedia & Hypermedia, 12(3):227-241.

Perry, W. 1970. *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart.

Piaget, J. 1966. *The origins of intelligence in children*. New York: International Universities Press.

Piaget, J. 1972. *The principles of genetic epistemology*. New York: Basic Books.

Piaget, J. 1974. *To understand is to invent*. New York: Viking Press.

Preiss, D. & Sternberg, A. 2010. *Innovation in teaching psychology: Perspectives on learning, teaching, and human development*. New York: Springer Publishing Company.

Qing, Z., Jing, G. & Yan, W. 2010. Promoting pre-service teachers' critical thinking skills by inquiry-based chemical experiment. *Innovation and Creativity in Education*, 2(2):4597-4603.

Reagan, T. 1999. Constructivist epistemology and second/foreign language pedagogy. *Foreign Language Annals*, 32:413-425.

Rebecca, R., Berkes, E., Mosley, M., Hui, D. & Joseph, G. 2005. Critical discourse analysis in education: A review of literature. *American Educational Research Association*, 75 (3):365-416.

Redfield, D. L. & Rousseau, E. W. 1981. A meta-analysis of experimental research on teacher questioning behavior. *Review of educational research*, 51 (2), 237-245.

Renaud, R. & Murray, G. 2007. The Validity of Higher-Order Questions as a Process Indicator of Educational Quality. *Research in Higher Education*, 38(3):319-351, DOI: 10.1007/s11162-006-9028-1.

Research in Higher Education. A review of studies. 26(1):3–29.

Resnick, L., 2005. *Principles of Learning for Effort-Based Education*, version 3, 0. Pittsburgh: University of Pittsburgh Press.

Riffel, C. 2014. The Socratic method reloaded; How to make work in large classes?. *Canterbury Law Review*, 20, (125-135).

Robitaille, Y. P. and Maldonado, N. 2015. Teachers' experiences relative to successful questioning and discussion techniques. *American International Journal of Contemporary Research*, 5(1):7-16.

Rosenblatt, L. 1978. *The reader, the text, the poem: The transactional theory of the literary work*. Carbondale, Illinois: Southern Illinois University Press.

Rosenblatt, L. 1995. *Literature as exploration*. New York: The Modern Language Association of America.

Rupp, A., Ferne, T. & Choi, H. 2006. How assessing reading comprehension with multiple-choice questions shapes the construct: a cognitive processing perspective. *Language Testing Journal*, 23(4):441-474.

Sahamid, H. 2004. Socratic questioning in the teaching of short stories: An action research study. *The Reading Matrix*, 4(2):63-71.

Sahamid, H., 2016. Developing critical thinking through Socratic questioning: An action research study). *International Journal of Education and Literacy Studies*, 4(3):1-14.

Sasaki, M. 2005. Hypothesis generation and hypothesis testing: Two complementary studies of EFL writing processes. In P. K. Matsuda & T. Silva (Eds.), *Second language writing research: Perspectives on the process of knowledge construction* (pp.79-92). Mahwah, NJ: Lawrence Erlbaum.

- Savery, J. 2006. Overview problem-based learning: definitions distinctions. *The Interdisciplinary Journal of Problem-Based Learning*, 1(1):1-12.
- Scardamalia, M., & Bereiter, C. 2006. Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York: Cambridge University Press.
- Schuh, K., 2004. Learner-Centered principles in teacher- centered practice. *Teaching and Teacher Education*, 20(8):833-846.
- Schwarz, B., Hershkewitz, R. & Azmon, S. 2006. The role of the teacher in turning claims to arguments. In J. Novotna, H. Moraova, M. Kvatka & N. Stehlikova (Eds). Proceedings 30th conference of the international group for the psychology of mathematics education, 5(pp. 65-72). Prague: PME.
- Searle, J. R. 1969. Speech Acts. Cambridge: Cambridge University Press.
- Searle, J. 1979. Expression and Meaning. Cambridge: Cambridge University Press.
- Scollon, R. & Scollon, S. W. 2010. Discourse and intercultural communication. In Deborah Schiffrin, Deborah Tannen & Heidi E. Hamilton (Eds.) *The Handbook of Discourse Analysis* (pp. 538-547). USA, UK, Australia: Blackwell Publishing.
- Seifert, T. 1997. Academic goals and emotions: Results of a structural equation model and a cluster analysis. *British Journal of Educational Psychology*, 67:232-338.
- Seliger, H., & Shohamy, E. 1989. *Second language research methods*. Cambridge: Cambridge University Press.
- Sfard, A. 1998. On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27:4-13.
- Shomoossi, N. 2004. The effect of teachers' questioning behavior on EFL classroom interaction:

- A classroom research study. *The Reading Matrix*, 4(2):1-9
- Shulman, L.S. 1987. Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1):1-22.
- Siegel, H. & Biro, J. 2008. Rationality, reasonableness, and critical rationalism: Problems with the pragma-dialectical view. *Argumentation*, 22(2):191-202.
- Singh, P., Thomas, L., Mueller, E., Lims, G., Perkins, T. & Zhu, W. 2010. Open mind common sense: Knowledge acquisition from the general public. *Springerlink Library*, 2519:1223-1237.
- Sivan, E. 1986. Motivation in social constructivist theory. *Educational Psychologist*, 21(3):209-233.
- Stein, B. 2008. *Assessing critical thinking program, council for aid to education*. Master thesis, Tennessee Technological University, pp. 1-20.
- Sternberg, J., Roediger, L. & Halpern, F. 2007. *Critical thinking in psychology*. Cambridge: Cambridge University Press.
- Stevens, R. 1912. The questions as a means of efficiency in instruction: A critical study of classroom practice. New York. Teachers College, Columbia University.
- Shulman, L. 1986b. *Those who understand: knowledge growth in teaching in Educational Researcher*, 15(2):4-14.
- Sivasubramaniam, S. 2015. Maximizing EIL competence through students' agency, voice and inter-subjectivity. *ESBB: English Scholars Beyond Borders*, 1(1):74-108.
- Susana E. Franco-Fuenmayor, Yolanda N. Padrón & Hersh C. Waxman. 2015. Investigating Bilingual/ESL Teachers' Knowledge and Professional Development Opportunities in a

- Large Suburban School District in Texas, *Bilingual Research Journal*, 38(3):336-352.
- Swartz, R. 2009. Energizing learning. *Educational Leadership*, 65 (5):26-31.
- Tabachnick, B. G., & Fidell, L. S. 1996. *Using multivariate statistics* (3rd ed.). New York: HarperCollins.
- Tan, Z. 2007. Questioning in Chinese university EFL classroom. *RELC Journal*, 38 (1): 87-103.
- Taylor, L. K., Alber, S. R. & Walker, D. 2002. The Comparative effects of a modified self-questioning strategy and story mapping on the reading comprehension of Elementary Students with Learning Disabilities. *Journal of Behavioral Education*, 11(2):69-87. <http://dx.doi.org/10.1023/A:1015409508939>.
- Teddlie, C. & Tashakkori, A. 2009. *Foundations of mixed methods research*. Thousand Oaks: SAGE Publications, Inc.
- Thakur, V. S. 2002. Sociolinguistic perspectives of politeness in communication. Delhi: Mohit Publications.
- Thakur, V. S. & Al Mahrooqi, R. 2015. Orienting ESL/EFL students towards critical thinking through pectoral inferences and elucidation: A fruitful pedagogic approach. *English Language Teaching*, 8(2):126-133.
- Thakur, V. S. (2016). Promoting critical thinking as a social practice; Shaping students' voice, agency and inter-subjectivity in a cohesive frame work. *Paper presented at 3rd English Scholars beyond borders International Conference on Crossing borders in English Teaching; Publishing and academic Study*, Providence University, Taichung, 19, 22 May.
- Thayer-Bacon, B. J. 2000. *Transforming critical thinking: Thinking constructively*. NewYork: Teachers College Press.
- Theresa, E. 2007. *Assessing and enhancing critical thinking skills: enhanced peer review (EPR)*

with CAT instrument. PhD Dissertation., Educational Cognitive Therapy, Tennessee Technological University.

Tomlinson, C. A., & Strickland, C. A. 2005. *Differentiation in practice: A resource guide for differentiating curriculum, grades USA*: ASCD publishers, 9-12.

Tsui, L., (2002). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *Journal of Higher Education*, 37(3):1-8.

Tuzlukova, V. and Prabhukanth, K. U. (2018). Critical thinking and problem solving skills: English for science foundation program students' perspectives. Collection of Papers of the Faculty of Philosophy, XLVIII (3), 37-60. Retrieved from file:///C:/Users/moosa/Downloads/Critical thinking and problem solving skills Engli.pdf

Van Lier, L. 1988. The classroom and the language learner. London. Longman

Vickers, C. H. 2010. Language competence and the construction of expert novice in NS–NNS interaction. *Journal of Pragmatics*, 42(1):116-138.

Vygotsky, L. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Walker, J., Brophy, S., Hodge, L. & Bransford, J. 2006. Establishing Experience to Develop a Wisdom of Professional Practice. *New Directions for Teaching and Learning*, 49-53, doi: 1002/t/.255.

Watson, G. & Glasser, E. 2006. Watson-Glasser critical thinking appraisal short form manual. San Antonio, TX: Harcourt.

Watzke, J. 2007. Foreign language pedagogical knowledge: Toward a development theory of beginning teacher practice. *The Modern, Language Journal*, 91(1): 36-82.

Weber, F. F. 1998. Three models of power in David Memet's Oleanna. In Jonatnan Culpeper,

- Mick Short and Peter Verdonk (Eds). Exploring the Language of Drama: From Text to Context (pp. 112-127). London and New York: Routledge.
- Wegerif, R. (Eds.). 2007. *Dialogic Education and Technology Expanding in Space of Learning*. New York: Springer Science and Business Media LLC.
- Wenning, J. 2006. *Engaging students in conducting Socratic dialogues: Suggestions for science teachers*. Available on wenning@phy.ilstu.edu
- West, R. F., Toplak, M. E. & Stanovich, K. E. 2008. Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, 100(4):930-941.
- Weusijana, B. 2006. *A Socratic ASK system: Helping educators provide a Socratic tutor for learners*. A PhD Dissertation. Northwestern University, UMI No. 3238417.
- Whorf, B. L. 1956. Science and linguistics. In J. B. Carroll (Ed.) *Language, Thought and Reality*, MA: MIT Press.
- Wideen, M., Mayer-Smith, J., & Moon, B. 1998. A critical analysis of the research on learning to teach: Making the case for an ecological perspective on inquiry. *Review of Educational Research*, 68:130–178.
- Williams, M. & Burden, R. 1997. *Psychology for Language Teachers*. Cambridge: Cambridge University Press.
- Williams, M. & Burden, R. 1999. Students' developing conceptions of themselves as language learners. *The Modern Language Journal*, 83(2):193-201.
- Wood, A. & Anderson, C. 2001. The case study method: Critical thinking enhanced by effective teacher questioning skills. *Proceedings of International Conference of the World Association for Case Method Research and Application*, Lund, Sweden, June 17-20.

- Wortham, S. 2006. Learning identity: The joint emergence of social identification and academic learning. Cambridge: Cambridge University Press.
- Wragg, E. and Brown, G. 2001. Questioning in primary schools. Routledge Falmer, London.
- Xun, G. & Land, S. 2004. A conceptual framework for scaffolding III-structured problem-solving processes using questions promotes and peer interactions. *Educational Technology Research and Development*, 52(2):5-22.
- Yamashita, J. 2007. The relationship of reading attitudes between L1 and L2: An investigation adult EFL learners in Japan. *TESOL Quarterly*, 41(1):81-105.
- Yazdanpanah, L. 2015. Investigation of ESL teacher knowledge in Australian adult education. *TESL-EJ*, Lilly Monash University, Australia, 18(4):52-59.
- Zascavage, V., Masten, W. & Nichols, C. 2007. Comparison of critical thinking in undergraduates and graduates in special education. *International Journal of Special Education*, 22(1):1-7.
- Zohar, A. & Dori, Y.J. 2003. Higher order thinking skills and low achieving students. Are that mutually exclusive? *Journal of the Learning Sciences*, 12(2):145-183.
- Zohar, A. 2004. *Higher order thinking in science classroom, students' learning and teachers' professional development*. Dordrech: Kluwer Academic Publishers.
- Zohar, A. 2007. Science teacher and professional development in argumentation, science and technology. *Education Library*, 35(3):245-268.