

QATAR UNIVERSITY

COLLEGE OF EDUCATION

AN INVESTIGATION OF THE IMPACT OF METACOGNITIVE PROMPTS, AS
FACILATED BY THE EFAP-SRL MODEL, ON ACHIEVEMENT ON THE END |OF
UNIT MATH TEST AMONG 8th GRADE GIRLS IN QATAR PREPARATORY
SCHOOLS.

BY

DIMA ALI ALRAWASHDEH

A Thesis Submitted to
The Faculty of the College of Education
in Partial Fulfillment of the Requirements for the Degree of
Masters of Arts in Curriculum and Instruction

June 2019

© 2019 Dima. All Rights Reserved.

COMMITTEE PAGE

The members of the Committee approve the Thesis of Dima Ali
Alrawashdeh defended on 11/04/2019.

Dr. Mary Newsome
Thesis/Dissertation Supervisor

Dr. Abdel-Salam Gomaa
Committee Member

Dr. Gordon Brooks
Committee Member

Dr. Salaheldin Kassab
Committee Member

Approved:

Ahmed Al-Emadi, Dean, College of Education

ABSTRACT

ALRAWASHDEH, DIMA., Masters: June: 2019, Masters of Arts in Curriculum and Instruction

Title: An Investigation of the Impact of Metacognitive Prompts, as Facilitated by the EFAP-SRL Model, on Achievement on the End of Unit Math Test Among 8th Grade Girls in Qatar Preparatory Schools.

Supervisor of Thesis: Mary Newsome.

The aim of the present study is to investigate the impact of using metacognitive prompts as facilitated by the Enhanced Formative Assessment with Self-Regulated Learning (EFAP-SRL) model on students' achievement on the end of unit Math assessment among 8th grade girls. The study was conducted at two schools over a two-month period; each school had a control group and a treatment group (N=113). Students from the treatment group received quizzes with built-in metacognitive prompts, while students from the control group received quizzes not containing metacognitive prompts. Teachers in the treatment group received three training sessions on how to use the EFAP-SRL model; the researcher received training sessions given by a specialist in SRL in College of Medicine in Qatar. The results showed no significant difference on the end of unit assessment scores between the control groups and treatment groups; furthermore, assessment scores from all four groups were skewed. The results of the study suggest a need for summative assessments to be constructed using an item bank of good quality test items. Additionally, the study suggests that a replication study using metacognitive prompts over a longer period of time is needed to determine a possible impact.

DEDICATION

For the one who taught me how to be brave and ambitious, my father.

ACKNOWLEDGMENTS

I would first like to thank my thesis advisor Dr. Mary Newsome at Qatar University. She was there always, whenever I had a question about my research or writing. Her words and the way she believed in me were a major source of motivation. I would also like to thank all the committee members who were involved in the process of this study: [Dr. Abdel-salam Gomaa, Dr. Gordon Brooks, Dr. Salaheldin Kassab]. Without their passionate participation and their deep knowledge, the study could not have been successfully conducted. I would also like to give a special thanks to Professor Hussain Mohammad Al-Qassem at the Math, Statistics and Physics Department at Qatar University; his inspiring words always encouraged me through my Bachelor studies to continue higher education. I would also like to thank my friends, Haneen Jobara, Lina Bader, and my colleague Mohammad Mollazehi for their help in statistics. Finally, I must express my deep gratitude to my parents for their unconditional love and support, my lovable siblings, and to my spouse for believing in me and providing me with continuous encouragement throughout my years of study. It was impossible for me to reach this accomplishment without them. Thank you.

TABLE OF CONTENTS

DEDICATION	iv
ACKNOWLEDGMENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: LITERATURE REVIEW	6
2.1 Formative Assessment	6
2.2 Feedback.....	7
2.3 Formative Assessment and Self-regulated learning	7
2.4 SRL and Formative Assessment Studies.....	10
2.5 The Use of Metacognitive Prompts.....	13
CHAPTER 3: METHODOLOGY	16
3.1 Participants	16
3.2 Data collection.....	17
3.2.1 Method.....	17
3.2.2 The Model.....	18
3.2.3 Curriculum.....	18
3.2.4 Mathematics quizzes with metacognitive prompts.....	19
3.2.5 The SRL Math self-reflection and mastery learning form	20

3.2.6 End of unit test.....	21
3.3 Procedure.....	21
3.4 Data Analysis	24
3.5 Ethical Consideration	25
CHAPTER 4: RESULTS	27
4.1 Descriptive Statistics	27
4.2 Histograms	28
4.3 Normality and homogeneity:.....	30
4.4 Non-Parametric test.....	32
4.4.1 Mann-Whitney test:	32
4.4.2 Kruskal-Wallis Test.....	34
CHAPTER 5: DISCUSSION.....	36
5.2.1 The use of metacognitive prompts and their impact on achievement	37
5.2.2 Use of assessment and instructions to develop metacognitive skills among primary and preparatory school students.....	42
5.3.1 Limitations.....	44
5.3.2 Recommendations and suggestions for future research	46
REFERENCES	48
APPENDICES	54

LIST OF TABLES

Table 1. No. of Participants by group	16
Table 2. Descriptive Statistics for the Sample (N =113).	27
Table 3. Test of Normality: Shapiro Wilk	30
Table 4. Test of Homogeneity of Variances: Levene Statistic	32
Table 5. Mann-Whitney Test- Ranks (C1, T1).....	33
Table 6. Test Statisticsa(C1, T1).....	33
Table 7. Mann-Whitney Test- Ranks (C2, T2).....	34
Table 8. Mann-Whitney (C2, T2)	34
Table 9. Ranks – Kruskal- Wallis Test	35
Table 10. Kruskal- Wallis (C2,T2)	35
Table 11.Means of all participated groups.....	39
Table 12. Difficulty Index for Multiple choice questions.....	40
Table 13. Difficulty Index for Problem Solving questions.....	41

LIST OF FIGURES

Figure 1 Question sample of treatment group quiz.....	20
Figure 2. Mean Score Plot	28
Figure 3. School 1 Histogram	29
Figure 4 School 2 Histograms.....	29
Figure 5. Data Distribution.	31

CHAPTER 1: INTRODUCTION

1.1 Background

In education, the term *assessment* refers to the various tools and methods used by practitioners to evaluate and document educational needs, readiness, learning progress and/or skill attainment (Sumantri and Satriani, 2015). However, this definition does not explicitly acknowledge teachers' actions/reactions, which essentially direct the learning process, as one of these assessment tools. In fact, William (2011) points out that teachers' actions and reactions in class plays a major role in a student's understanding of where they are in relation to a learning goal despite the fact that it has not traditionally been counted as a means of assessment (William, 2011). More recently, there has been a growing interest to better understand such actions with more researchers recognizing them as an important means of assessment. From this perspective, then, there are two types of assessment: summative assessment and formative assessment. Summative assessment is concerned with measuring the final product of the learning, while formative assessment is concerned with evaluating the learning process (William, 2011). Formative assessment, or assessment for learning, is the process of collecting evidence to be used by practitioners and learners to determine the quality of learning (Sumantri and Satriani, 2015). Moreover, it evaluates whether learners are achieving the learning goals or not, it highlights learners' strengths and weaknesses, and it identifies students' educational needs in a constructive way. Formative assessment has been strongly related to self-regulated learning by a growing body of research, which is the primary concern of this investigation.

Self-regulated learning (SRL), sometimes referred to as self-directed learning, is an

intentional, psychological activity that learners direct and control with the aim of acquiring knowledge about a specific topic; it is described in the literature as the level to which learners are active participants in their own learning (Brinke and Boshuizen, 2015). Zimmerman described self-regulated learners by saying:

“They approach educational tasks with confidence, diligence, and resourcefulness. They are aware when they know a fact or possess a skill and when they do not. Unlike their passive classmates, self-regulated students proactively seek out information when needed and take the necessary steps to master it” (p. 4, 1990).

There is a growing body of empirical research supporting the importance of SRL strategies in academic achievement (Beekman, Brinke & Boshuizen, 2015; Zimmerman, 2000). For example, recent studies have shown that SRL leads to a better understanding of mathematical concepts (Beekman, Brinke & Boshuizen, 2015). Such research highlights the importance of giving students the chance to discuss and reflect on their problem-solving strategies, using metacognitive prompts as in the EFAP-SRL (Enhanced Formative Assessment Program with Self-Regulated Learning) approach. This approach is heavily borrowed from Zimmerman’s Cyclical Phases Model (Zimmerman, 2000). It is distinguished by multiple feedback cycles consisting of three main phases: planning, practice, and evaluation. The *planning phase* is where students learn how to review their previous efforts, analyze the problem, and use strategies that are aligned with the learning goal(s). Moreover, students are able to make self-efficacy judgments regarding their work. The *practice phase* is where students have the opportunity to implement the plan. This phase consists of monitoring and making adjustments to the plan, if needed. The *evaluation phase* is where students assess how effective their learning strategies were based on the

feedback provided. The response gathered from the evaluation will become the basis for the next planning phase for the repetition of the SRL cycle.

The researcher acknowledges that SRL involves other important constructs such as motivation and affect; however, the current investigation is strictly limited to the underlying metacognitive strategies involved during the learning of mathematics. Metacognitive prompts support and stimulate students' metacognitive reflection and play a vital role in enhancing the learning process in order to develop students' ability to generate hypotheses (Veenman, 1993). Metacognitive prompts may also justify students' actions and help them think about their thinking (Bodvarsson, 2005).

Other similar studies on self-regulation and metacognition have found that metacognitive skills are easier to be noted among secondary school students. Numerous studies have shown that primary school students lack metacognitive abilities (Cromley, et al. 2010). Consequently, only a few studies have focused on students aged 8 to 14 years. In fact, the overwhelming majority of research investigating SRL has been done among secondary and college-level students. Due to the lack of studies related to metacognition and SRL focusing on primary and preparatory students, this study focused on 8th grade preparatory students averaging in age from 13 to 14 years. This study aims to investigate how do metacognitive prompts, as facilitated by the EFAP-SRL model, impact achievement on the end of unit Math assessment among 8th grade girls.

1.2 Operational Definitions:

EFAP-SRL model: Enhanced Formative Assessment and Self-Regulatory Learning Program (EFAP-SRL). It is a psycho-educational model.

Formative assessment: Four EFAP-SRL quizzes for treatment groups and four normal

quizzes for control groups (Appendix A).

Metacognitive prompts: Questions built into the quizzes and self-reflection form that invites students to think about their thinking.

Achievement: The obtained grade on the end of unit summative test.

1.3 Significance of the Study

This study provides valuable information and investigation of metacognitive skills as part of SRL among students at preparatory schools. While there are many studies that associate metacognition with formative assessments at the secondary level and among university students, there is a shortage of research investigating a possible association between metacognition and formative assessment among younger students (Clark, 2012). The present investigation focuses on younger students' metacognitive skills because there is clear evidence that students who are able to self-regulate their learning may have better academic achievement (Clark, 2012; Cleary & Zimmerman, 2001; Zimmerman 2000). The study may also serve as a catalyst for the Ministry of Education and Higher Education in Qatar to systematically develop formative assessment strategies in order to enhance young students' metacognition skills. This study aims to investigate how do metacognitive prompts, as facilitated by the EFAP-SRL model, impact achievement on the end of unit Math assessment among 8th grade girls.

1.4 Research Questions

This study will serve to answer one central question and test one hypothesis in order to determine the impact of formative assessment and the use of metacognitive prompts on achieving the mathematics learning objectives measured by the end of unit test.

- 1- Research Question: How do metacognitive prompts, as facilitated by the EFAP-SRL model, impact achievement on the end of unit Math assessment among 8th grade girls?
- 2- Research hypothesis: Students who participate in EFAP-SRL have higher achievement on the end of unit summative assessment compared to students who did not participate in EFAP-SRL.

CHAPTER 2: LITERATURE REVIEW

This chapter will detail major studies relevant to the research question. Firstly, this chapter will discuss formative assessment and the importance of positive feedback. Secondly, the chapter will cover Self-Regulated Learning (SRL) and its relation to formative assessment through previous studies. Finally, it will highlight the role of metacognition as a part of self-regulated learning, particularly the use of metacognitive prompts.

2.1 Formative Assessment

Traditional school environments have been criticized for decades and continue to be so. For instance, Dewey (1938) is perhaps the most famous of such critics suggesting the way in which classroom desks were arranged in rows only served to force students into silence. Such environment, in this view, only discourages interaction and positions the teacher as the “knower” and students as passive recipients of knowledge. Building on Dewey’s concerns, Clark (2012) stated that such an environment destroys the student’s curiosity to learn. This and similar criticism led to a philosophical transition toward viewing students as active participants who design and contribute to their own learning rather than passive recipients of information. This transition to a more evolved perspective of classroom learning is essentially the foundation for the notion of formative assessment. Formative assessment refers to assessment tools concerned with the process of learning over the product of learning. Formative assessment helps in determining the quality and degree of the learning and aims to help students improve their own learning (Allal, 2010). According to research, formative assessment is used to provide feedback to both learners

and teachers during the learning process, while simultaneously being used as a tool to improve instruction and aid in the achievement of the learning objectives (Allal, 2010; Clark, 2012; Satriani & Syrafi, 2015).

2.2 Feedback

The literature on formative assessment suggests that effective feedback is an important element to support the learning process. For example, good feedback can be used by instructors to become aware of where instructional modifications are needed in order to improve their own teaching. Furthermore, effective feedback from assessments, or perhaps even classroom participation, can help students understand and improve their learning (Winne & Butler, 1995). However, all feedback is not created equal; a teacher who, for example, says to a student, “you have worked hard” is not offering anything up regarding how well the student has learned or what more the student can do to improve the learning despite the fact that the feedback was positive. In other words, the aim of the feedback should be to provide information to the learner regarding the quality and/or extent of the learning. Consequently, if the information given to the learner during formative assessment helps confirm, add to, or replace a concept in memory, then it is considered effective feedback (Gronlund & Linn, 1990; Winne & Butler, 1995). Students are expected to make constructive changes in how they learn, depending on effective feedback (Black & William, 2009). In this view, formative assessment uses effective feedback to help students “learn how to learn”.

2.3 Formative Assessment and Self-regulated learning

Formative assessment has been closely linked to self-regulated learning (SRL) by a growing body of research. Clark (2012), for example, described formative assessment as

a process with the potential to support learning beyond school years by developing learning strategies which individuals may rely on across their entire life span (p. 217). Other researchers have also hypothesized a relation between self-regulated learning and formative assessments (Black & William, 1998; Panadero & Jonsson, 2013). Such work has suggested that formative assessment allows students to contribute actively to the learning process, meaning they exercise SRL. Tay (2015) provided evidence of this relation in his study of 13 female students who were asked to do two tasks: The first task was a paper and pen in-class assignment in which students had to write a short essay, while the other were equivalent but in a live, online forum. The two tasks were presented to the students as formative assessments. The students filled a self-report questionnaire based on Zimmerman's 3-phase cyclical model of self-regulation. Also, the researcher gathered data through one-to-one interviews, and the quantitative analysis showed statistically significant differences between the two tasks. In the forum task for example, the students were more careful about their writing and they wrote the answers carefully because they took into consideration that other people could read their responses online. They tried to avoid grammatical mistakes and to write in a more formal way. In contrast, students who had the in-class assignment knew that only the teacher could read their answers. Consequently, the findings suggested that the context of the formative assessment had an impact on the use of SRL among participants.

In this view, the effective feedback received during formative assessment instills in the learning a way of thinking about and evaluating their learning that, in time, has the potential to become part of the individual's regular learning strategy. This psychological process has been coined "self-regulated learning". SRL has become a powerful theoretical

framework in educational and psychological research (Azevedo, 2010); it is an intentional psychological activity that learners direct and control with the aim of acquiring knowledge about a specific topic. In the literature, SRL is described as the extent to which learners are active participants in their own learning (Santriani & Serafi, 2015; Beekman, Brinke & Boshuizen, 2015). Studies have shown that formative assessment is an effective tool for facilitating self-regulated learning strategies (Allal, 2010); Formative assessment and the effective feedback that results, models how students can self-regulate their learning to acquire needed knowledge and skills. For example, formative assessment can raise self-awareness about how to plan individual goals, monitor progress, and reflect on the learning that occurred (Allal, 2010; Clark, 2012). In other words, formative assessment is a tool that can be used to support and promote SRL, which is indicative of a responsible and effective learner.

Similar to Bandura's (1986) Social Cognitive Theory (SCT) is grounded in Social Learning Theory (SLT) of the 1960s. Social Cognitive Theory proposed that learning takes place in a social context with an interaction among the person, environment, and his behavior. The skills involved in self-regulated learning are a significant part of SCT in that SCT emphasizes the importance of self-efficacy and metacognition, the basic tenets of SRL (Sasai, 2017). It is important to note that the goal of SCT is not only aligned with formative assessment; it is, in fact, encapsulated by its overall target. As Bandura pointed out, "The fundamental goal of education is to equip learners with self-regulatory capabilities that enable them to educate themselves." (1997, p. 174). Ultimately, students can be supported by teachers and through instruction on how to become aware of their learning and, in time, develop self-regulatory strategies (Bandura, 1986; Zimmerman, 2002; Pintrich 1999).

2.4 SRL and Formative Assessment Studies

The majority of formative assessment and SRL studies are concerned with secondary level and college students (Zimmeman, 2000; Hudesman, et al., 2013). Secondary education students demonstrate higher metacognitive skills since SRL behavior and the monitoring of one's learning require abstract thinking, which improves with age (Fletcher & Shaw, 2012). The challenge of investigating the impact of formative assessment and SRL or metacognition prompts on students' academic achieving is addressed by the Enhanced Formative Assessment and Self-Regulatory Learning Program (EFAP-SRL).

The EFAP-SRL approach is based on a model of self-regulated learning put forth by Zimmerman (2000, 2002, 2006) and Grant (2003; Grant & Green, 2001). As a leader in SRL, Zimmerman has developed three SRL models (Panadero and Tapia, 2014): The first model is known as the Triadic Analysis of SRL, the second model is called the Multi-Level model, and the third one is Cyclical Phases model. We focused on the Cyclical Phases model because it has been tested in a series of studies (Cleary & Zimmerman, 2001; Kitsantas & Zimmerman, 2002; Cleary et al., 2006; DiBenedetto & Zimmerman, 2010). The first investigation explored the SRL skills among teenage boys who were professional, non-professional, and novice basketball players, which concluded that those players who were professionals performed more SRL actions. The second study was similar to the first study but the participants were college women. The third study, Cleary et al. (2006) coached 50 college students in basketball free-throws and the results showed a linear trend: the more phases trained the better the score. Finally, the fourth study was among 51 high school seniors during Science courses. The results found that higher achievers showed

more use of the sub processes of Zimmerman's model. Cyclical Phases model is a psycho-educational model characterized by a cycle of continuous feedback. Each cycle is broken into three phases: planning, practicing, and evaluating (Zimmeman, 1998; Hudesman, et al., 2013). Administering EFAP-SLR consists of 5 steps: (1) Instructors build special quizzes to assess both content knowledge and SRL skills; (2) Instructors review students' answers and grade papers by providing feedback to the students; (3) Students fill the self-reflection form for each incorrect answer where they have the chance to improve their performance; (4) Teachers review and mark the self-reflection form to determine the extent to which the student has mastered both content and metacognitive skills; (5) Teachers discuss the results with the students.

The self-reflection form consists of different metacognitive prompts to motivate students to think about their thinking. The reflection form has three parts. Part one focuses on preparation time before the quiz such as the number of problems solved while studying and preparing for the quiz at home and the time needed. Part two is concerned with how the student solved the problem; why the problem-solving strategy went wrong and what could be done to correct the mistake. The third part gives the student the chance to re-solve the same problem before being asked to solve a similar problem in order to demonstrate mastery of the learning objective. Finally, the teachers are expected to review and measure the degree to which the student has mastered the objective. In many cases, instructors use the feedback provided on the self-reflection form for class discussions (Hudesman, et al., 2013).

One group of researchers conducted two quantitative studies and used the EFAP-SRL model mentioned above (Hudesman, et al., 2013). Both were conducted during

summer sessions at Urban College of Technology among 1500 students in 2005 and 2009. Of the 1500 participants, 53% were female, and 47% were male. The main research question was concerned with investigating whether students enrolled in EFAP-SRL developmental mathematics classes were more likely to pass a post-course administration of the COMPASS portion of the ACT than students enrolled in comparison sections of the course (Hudesman, et al., 2013). COMPASS is a series of computerized placement tests developed by ACT, Inc. to evaluate skills and place students in appropriate courses. The study compared the academic achievement of students enrolled in EFAP-SRL developmental math classes with the academic progress of the students in the control group. The results showed that students enrolled in EFAP-SRL passed the developmental mathematics course with higher pass rates compared to the control group ($M=73.18$ and $M=58.03$ for EFAP-SRL and control groups respectively, $F = 9.96, p < .05$). Additionally, participants in this study attained higher grades on the COMPASS than the control group. Despite obvious limitations such as funding and the experience level of mathematics instructors, the overall outcomes were positive. (Hudesman, et al., 2013).

Supporting the previously mentioned research, similar studies among primary students found that SRL leads to a better understanding of mathematical concepts (Beekman, Brinke, & Boshuizen, 2015; Santriani & Serafi, 2015). Likewise, Pintrich and De Groot (1990) investigated the relationship between self-regulation (use of metacognitive strategies) and learning performance among 7th grade English and Science students using the Motivated Strategies for Learning Questionnaire (MSLQ); they found that SRL skills (in specific: management, metacognition) correlated positively with students' performance; moreover, it predicted performance. However, contrary findings

were revealed from yet another similar study at two languages institutes in Iran. In this investigation, the aim was to determine the relationship between SRL and second language achievement among 130 students using a questionnaire that consisted of 46 items assessing self-regulated learning. After running the frequency analysis, the researchers concluded there was no significant relationship between SRL and L2 performance (Mahmoodi, 2014). In general, studies have documented that primary education students do not control their learning as much or as well as secondary students.

2.5 The Use of Metacognitive Prompts

EFAP-SRL provides prompts to help students improve their regulating abilities in three primary phases: planning, practicing, and evaluating. In the planning phase, students are promoted to read each item of the assessment and make a judgement regarding confidence in their ability to solve the items correctly. They choose on a scale of 0%, 25%, 50%, 75%, or 100% how confident they are in their ability to solve the problem. After that, students set goals and practice addressing the best solving strategy to use in order to reach this goal. After they finish solving the problem, students are then prompted to evaluate their work and make a second self-judgement regarding their confidence that they had answered the question correctly by choosing on the scale again. After finishing the quiz, for each incorrect answer the students were asked to fill a SRL Math self-reflection and mastery learning form. This form assists the students to connect their content knowledge and the use of metacognitive skills depending on the metacognitive prompts. (Hudesman, et al., 2013).

Metacognitive prompts can boost SRL by motivating regulatory activities, which, in turn, leads to improvement in learning outcomes (Bannert, 2015). Recent research in

metacognition and SRL reveals that learners usually do not spontaneously use metacognitive skills while learning resulting in lower achievement of learning outcomes (Azevedo, 2009; Costa, 2013; Winne & Hadwin, 2008). Such research also suggests that students must be prompted to use metacognition until it becomes a natural learning strategy for them. Current research suggests that a learner's control and awareness of the learning process is important and demonstrates valuable consequences from using metacognitive prompts (Azevedo, 2009). Other findings have also shown significant differences in performance for students supported by metacognitive prompts (Bannert & Sonnenberg, 2015).

More recent work conducted by Bannert & Sonnenberg (2015), investigates the effect of a new type of metacognitive prompt, self-directed metacognitive prompts, on learning outcomes. The study investigated 70 undergraduate students (predominately female 82.9%) studying media communications and human-computer systems at a German university and was concerned with whether the use of metacognitive prompts during learning had an impact on the achievement of learning outcomes. Students in the experimental group received guidance sessions on the use of metacognitive prompts to guarantee adequate application during the learning. The prompts asked students to reflect, monitor, and control their learning explicitly. They helped students stay focused on their own thoughts and to understand the activities they were engaged in throughout the learning process. The prompts appeared as a pop-up window placed in the screen, and each prompt consisted of a list of options for strategies to regulate the learning process. The learning performance was measured by three knowledge tests; the tests were on different levels of Bloom's Taxonomy (Bloom, 1956). In summary, the findings revealed that learning

outcomes were achieved and the students' performance were enhanced by the use of metacognitive prompts. Students who were in the experimental group showed a higher performance ($M = 116.43$, $SD = 45.97$) than students in the control group who did not have the prompts ($M = 98.49$, $SD = 36.72$).

Metacognitive skills are essential to improving achievement in mathematics. While there are many studies connecting formative assessments to metacognition, there are no studies investigating the impact of EFAP-SLR among students in the Gulf region. Furthermore, the results from previous studies are not generalizable to learners in Qatar because of significant cultural differences that surely influence both teaching and learning; therefore, such a study is needed in Qatar.

This study provides insight for educational policy makers, teachers, and school administrators who are looking for best practices in helping students develop their understanding of mathematics and achieve specific learning goals. By giving students the chance to discuss and reflect on their problem-solving strategies, such a model will help them realize that success is related to experimenting with different strategies and not simply spontaneous ability. Therefore, this study will investigate the impact of Self-Regulated Learning as facilitated by EFAP-SLR model on students' achievement on the end of unit Math test among 8th grade girls at two preparatory schools in Qatar.

CHAPTER 3: METHODOLOGY

This chapter presents the participants' information, data collection method, and data analysis procedures. Finally, it presents the ethical considerations implemented in the study.

3.1 Participants

A convenience sample of 113 students from two preparatory schools in Qatar participated in this study; all of the participants were 8th grade females (13-14 years old). The 113 participants were from 4 different classrooms selected on a random basis; the teachers for each class were randomly assigned by the school at the start of the year. Of the four groups (from two schools), the researcher randomly identified one treatment group and one control group from each school. In other words, two random classes from two different schools formed the treatment groups and the other two classes formed the control groups. The groups were coded as T1, C1, T2, and C2. (see Table 1).

Table 1. No. of Participants by group

Group	N
C1	28
T1	29
C2	27
T2	29
Total	113

Participants were not made aware to which group they belonged (treatment vs. control). Additionally, participants were informed they were participating in a research study involving multiple schools in Qatar investigating different Math learning strategies, which allowed the researcher to control for possible contamination across groups. For example, participants being aware that the study was concerned with investigating learning strategies across multiple schools may reduce the likelihood that students would discuss any differences between quizzes as differences would be expected in this case. Additionally, it is important to point out that students in treatment groups are not receiving any special instruction other than the individual differences among the instructional approaches of the randomly assigned teachers. In the control groups, students were simply asked to “call their attention” to their learning in order to solve a quiz item.

3.2 Data collection

3.2.1 Method

A quantitative approach was employed in this study. Since the purpose of this study is to investigate how do metacognitive prompts, as facilitated by the EFAP-SRL model, impact achievement on the end of unit Math assessment among 8th grade girls. The researcher chose a quantitative approach and used the end of unit test scores as the measure of achievement. Additionally, similar studies investigating the implementation of EFAP-SRL model such as those conducted by Zimmerman (2000, 2002, 2006), Grant (2003; Grant & Green, 2001), and Hudesman, (2013) followed the same approach. We did not use the pre-test/post-test approach to ensure that any improvement in achievement at the end of unit test was not the result of particular unit being easier or harder than another.

3.2.2 The Model

The approach guiding this investigation is rooted in Zimmerman's (2000) model of self-regulated learning (SRL), which was later modified by Grant and Green (2001) and adapted by Hudesman (2013) to suit the Mathematics curriculum where it was given the name, EFAP-SRL (Hudesman et. al, 2013). Zimmerman (2001, 2012) and Kisantas (Kisantas & Zimmerman, 2002) developed various micro-analytic measures to assess the validity of SRL cyclical phases model. The model consists of three phases: Planning, practicing, and evaluating. In the Planning phase, the students analyze the problem, set goals and plan how to achieve them by depending on number of metacognitive prompts to energize this process. In the practicing phase, students execute the solution. Finally, in the evaluating phase students assess how they performed the solution and make attributions about their points of strength and weakness. These phases will influence the approaches used in the upcoming tasks. The quiz content part in the model will be modified for the study as needed based on the Qatari mathematics curriculum and its learning goals. The modifications will depend on the topic of the unit and the number of questions in each quiz only (Appendix A). The researcher did not change/modify any part pertaining to the SRL metacognitive prompts. Zimmerman's models all have empirical evidence supporting the validity of their main aspects. (Panadero, 2017). Regarding reliability, SRL cyclical phases model is found to have 0.76 – 0.93 percent agreement under reading and studying tasks (Callan, Cleary & Zimmerman, 2012).

3.2.3 Curriculum

The 8th grade mathematics textbooks used in the classroom throughout this study were

translated from Pearson Books. Textbook selection and translation of materials is coordinated by the Ministry of Education, and, therefore, is the same across groups. The language of instruction is bilingual (Arabic-English) where students use both languages to learn concepts. However, English numbers are used to solve Math problems. Quizzes are unified among schools since they cover the same learning objectives. The addition of metacognitive prompts to the quizzes provided to the treatment groups was reviewed and approved by The Ministry of Education because it was agreed that the prompts did not change the objectives being assessed. The two main topics covered through class instruction during this study were integer indices and roots, and irrational numbers.

3.2.4 Mathematics quizzes with metacognitive prompts

Students from the experimental group completed specially formatted quizzes that included metacognitive prompts at least once a week throughout the duration of the study (approximately two months). Each quiz consisted of 4 to 5 Math questions and required no more than 15-20 minutes to be solved. This procedure was followed across one unit of instruction ending with the end of unit assessment. For quiz items, students were required to apply different metacognitive judgments before and after solving the problem. Before solving the problem, students had to estimate how confident they were in their ability to solve the problem on a scale of 1 to 100%. After solving the problem, students were asked to choose again out of 100% how confident they were of their answer and the strategies used (see Figure 1). The control groups were administered the same quizzes over the same duration minus the metacognitive prompts.

Before solving each problem, how confident are you that you can solve it correctly?	REMEMBER! Show all your work. Simplify all your answers.	After you have solved each problem, how confident are you that you solved it correctly?
0% 25% 50% 75% 100%	1. Factor completely: $10x^4 y + 4xy^2 - 2y =$	0% 25% 50% 75% 100%

Figure 1 Question sample of treatment group quiz

3.2.5 The SRL Math self-reflection and mastery learning form

For each incorrect answer, students were asked to complete a separate self-reflection and mastery learning form (Appendix B). This form is designed to give the student further assistance in assessing the relationship between their knowledge of the mathematics content and their ability to apply metacognitive skills depending on the added metacognitive prompts. In the first section of the form, students were asked to (a) do a comparison between their predicted quiz score and the actual quiz score, then justify any significant discrepancy; (b) evaluate how accurate their academic confidence judgments were; (c) depending on the feedback provided by the teacher, students were asked to indicate which strategy was incorrect on the first attempt. In the second section of the form, students were asked to solve the original problem again and include a description of mathematics strategies used in their work. Then, they were asked to solve another similar problem using the same strategy.

3.2.6 End of unit test

At the end of the unit, students from the experimental and control groups were asked to complete the end of unit test. The end of unit test was prepared by one teacher in each school following a descriptive table sent from the evaluation institute at Ministry of Education. After the teachers build the exam, each mathematics supervisor at each school revised it and send it to the evaluation institute to be approved or modified. As a result, all the exams covered the same learning goals and mathematics standards. And both experimental and control groups in each school had the same end of unit test.

3.3 Procedure

This investigation was conducted at two girls' preparatory government schools in Qatar during the first semester of the 2018-19 school year. Prior to conducting the study, the researcher received the approval of the Ministry of Education (MOE) (Appendix C). as well as the approval of the Institutional Review Board (IRB) (Appendix D). at Qatar University. Moreover, the researcher met with both school principals to explain the purpose of the research and to present the informed consent to be given to the participants. The researcher also introduced the specialist who gave training sessions for teachers in self-regulated learning. Four classes with four different teachers were identified to participate in the study; the classes were randomly divided into two control groups and two treatment groups. There were approximately N=30 students in each class with N=113 in total. The two control groups were administered standard quizzes, while the two treatment groups had the EFAP-SRL quizzes for a period of two months. The EFAP-SRL quizzes contained the same mathematics problems but included metacognitive prompts. Over a two-month period, students had 4 quizzes, two each month. Possible ethical issues concerning why

Math quizzes differed slightly among groups were addressed in several ways. Firstly, the objectives assessed on each quiz were the same across all groups. Secondly, the individual quiz items for all quizzes were identical and constructed by a testing committee at the parent school. The only difference among the quizzes across groups is that those in the treatment group were asked to think about their approach to solving an item as well as to reflect on how they answered the problem after the fact (i.e. the addition of metacognitive prompts in the quizzes of the treatment groups). It is unlikely that students would discuss the metacognitive prompts across group, but rather the quiz items and answers because the metacognitive prompts are not part of the actual “correct” answer. That is, the quiz score is determined by the number of correct answers and not the quality of their reply to a metacognitive prompt.

Teachers were not informed of whether they were in the treatment or control group in order to mitigate any bias that could be introduced in the teaching itself. Additionally, as with the student participants, teachers were informed the study was concerned with investigating different learning approaches across schools so as not to emphasize any instructional differences among teachers. Teachers in the treatment groups, however, were asked to participate separately in two training sessions on the implementation of EFAP-SRL quizzes provided by the researcher, which may allow those teachers to deduce that they are among the treatment groups. In order to mitigate the impact of such possibility, a level coordinator was asked to “spot check” the treatment and control groups by observing the classroom and reporting on whether the use of metacognitive prompts like those integrated into the quizzes were being “coached” in classroom instruction.

The treatment is not concerned with having teachers integrate metacognitive

prompts into classroom instruction; the treatment is concerned with using metacognitive prompts on quizzes to help students recall their learning in order to answer a problem correctly and to lead them to incorporate metacognition as a regular test-taking strategy. Although it could be claimed that the possibility of higher scores among the treatment groups may be a result of the quality of teachers in those groups as opposed to the integration of metacognitive prompts on the formative assessments, due to the random assignment of teachers to groups, it is unlikely that the two different teachers from two different schools are of better quality, by chance, than those in the control groups. This is particularly true since there is a treatment teacher and a control teacher from each school, which diminishes the likelihood that one school may have significantly better-quality teachers than the other. Furthermore, all of the teachers across groups had received “meets expectations” on the performance evaluation administered by the schools, and all of the teachers hold a bachelor degree, which further suggests that the quality of teachers was not a mitigating factor in the quiz results across groups. Treatment group teachers received training in EFAP-SRL by a specialist to ensure their familiarity with the modified quiz format; they followed these steps:

Step 1: Teachers construct EFAP-SRL quizzes that assess both content knowledge and metacognitive skills.

Step 2: Teachers review students’ answers, mark papers, and provide effective feedback.

Step 3: Students fill self-reflection form for each incorrect answer and enhance responses using teacher feedback.

Step 4: Teachers review and mark the self-reflection form to determine the extent to which the student has mastered both content and metacognitive skills.

Participants were assessed three times using short quizzes throughout this study. At the end of the unit of instruction, participants completed the end of unit summative assessment, which was used to measure achievement of the learning goals and answer the research question. The data from the end of unit assessment was delivered to the researcher by the classroom teachers; For privacy, students' scores were coded as (T1:P1, T2:P1; C1:P1, C2:P1) where T stands for treatment, C for control, P for participant. The researcher is the only person who had access to the coded data.

3.4 Data Analysis

Descriptive statistics including means, standard deviations, and frequencies were calculated for each group individually in order to investigate the effect of implementing EFAP-SRL model supported by the use of metacognitive prompts. It was done to provide summary statistics for the control and treatment group scores. The statistical Package for the Social Sciences (SPSS) version 24 was used for the statistical analysis. After the students had completed the end of unit test, the teachers corrected the papers and they assigned a score out of 30 for each student. The names of students were coded as (T1:P1, T2:P1; C1:P1, C2:P1) where T stands for treatment, C for control, P for participant. Parametric tests such as t-test and *ANOVA* was used if our data met the assumptions to be normally distributed and homogenous. Moreover, t-test was employed to compare two means (Control group and treatment group) and indicates whether they are different than each other, also it illustrates how significant the differences were. We considered p-value to be 5% (.05). We used one-way *ANOVA* to compare four groups (the classes) on one variable (Test score). The main formulas used for the one-way *ANOVA were*:

1- Between-groups variance:

$$SS_{between} = \sum n_j (\bar{X}_j - \bar{X})^2$$

2- Within-groups variance:

$$SS_{within} = \sum (X_i - \bar{X}_j)^2$$

Otherwise if the data was not normally distributed we would do Non-Parametric tests such as Mann-Whitney test and Kruskal-Wallis test. Mann-Whitney is the non-parametric alternative test to the independent sample t-test, and it is generally used with ordinal data. It was conducted to analyze the effect that participation in treatment had on the participants. More specifically, the Mann-Whitney test was used to compare the control groups' end of unit test score with the scores of the treatment group. Results were illustrated using Histograms for a clearer vision. The assumption that supports the Mann-Whitney-Wilcoxon test is there is a difference in the scores for the treatment group. And regarding the Kruskal-Wallis test, it is a test by ranks, or one-way ANOVA on ranks. A non-parametric method that analyzes whether samples are originated from the same distribution. In our data, the assumption of independent observation is met and the outcome is ordinal. Therefore, we conducted the test to compare the four classes

3.5 Ethical Consideration

All ethical considerations were taken into consideration in this study. The participants had the option of withdrawing from the study at any time. Noteworthy, a strict confidentiality was assured for all participants, and all scores were signed as an anonymous student, no name or any identifiable information was stated when scores were delivered to

the researcher by the classroom teachers.

CHAPTER 4: RESULTS

This chapter presents the results of end of unit mathematics test for 113 females 8th graders who participated in this study. It also reports the differences between the treatment groups and control groups in two preparatory girls' schools using the Statistical Package for the Social Science software (SPSS).

4.1 Descriptive Statistics

This chapter discusses the results from N=113 scores from 8th grade Math students' end-of-unit test. The purpose of this study is to investigate how do metacognitive prompts, as facilitated by the EFAP-SRL model, impact achievement on the end of unit Math assessment among 8th grade girls. Table 2 presents the descriptive statistics for the sample for school 1 (57 students) and school 2 (56 students) under control group (C) and treatment group (T), from this table as seen the minimum score was 5 for control group of the first group, while the maximum score was 30 for both schools.

Table 2. Descriptive Statistics for the Sample (N =113).

	N	Minimum	Maximum	Mean	SD
C1	28	5	30	22.65	6.2566
T1	29	8.75	30	24.81	5.6155
C2	27	11.50	30	22.10	5.3853
T2	29	14.75	30	24.37	5.4952

The Mean plot is shown in Figure 2. From the figure, it is noticeable that the mean of treatment groups in each school is higher than the control groups.

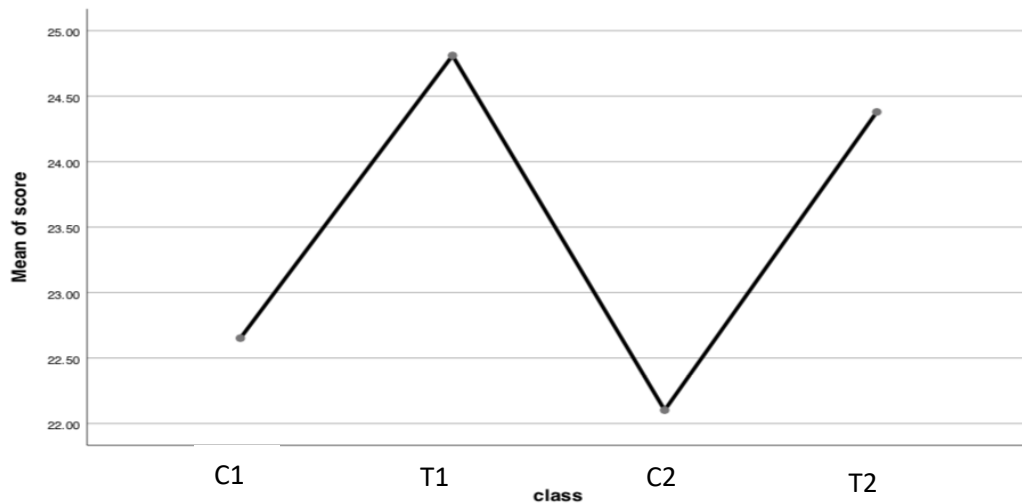


Figure 2. Mean Score Plot

4.2 Histograms

The Histogram plot represents the distribution of the sample graphically. In all four sub samples, we noticed skewness to the left. Figure 3 shows the distribution of school 1; the control group has a constant increase in score starting from 5 as the minimum up to 30 with the plot being skewed to the left overall. Although the treatment group is skewed to the left as well, but in different distribution, from the graph we can see the scores are formulated into three groups that show an increase in frequency in an unstable manner.

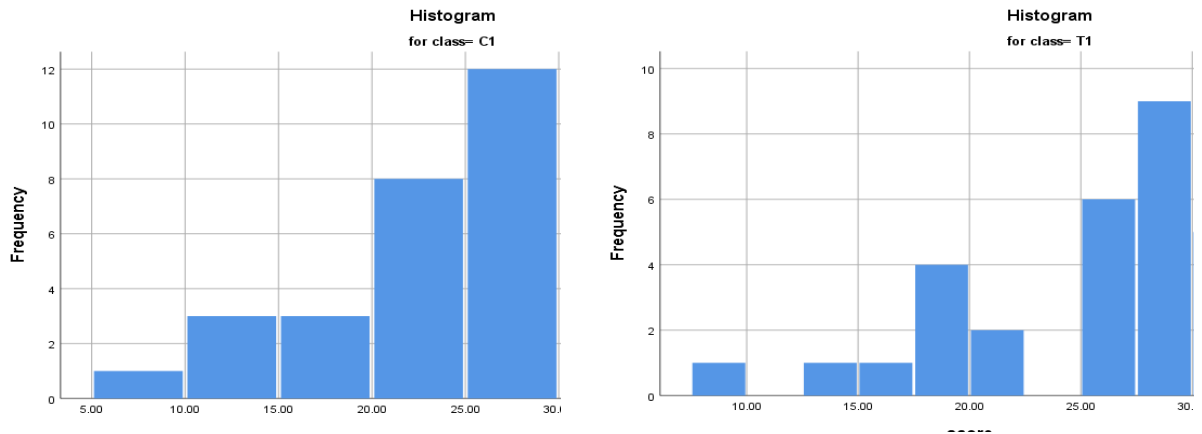


Figure 3. School 1 Histogram

School 2 distribution as shown in Figure 4 for both sub samples is skewed to the left. The control group has a higher median at 25, while the treatment group is more distributed.

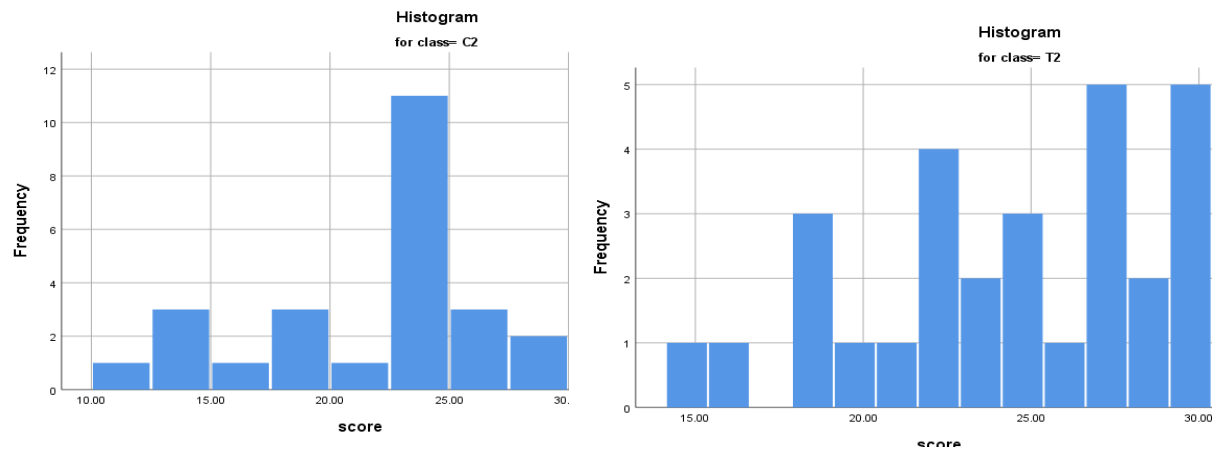


Figure 4 School 2 Histograms

4.3 Normality and homogeneity:

The data analysis of the four groups was initiated by testing the normality and homogeneity using the Shapiro-Wilk test and Levene's test to ensure that our sample is normal, considering the small sample size we had. The null hypothesis test assumes that the population is normally distributed ($\alpha = .05$). Based on the results, we rejected the null hypothesis test and we indicated that there is evidence that our data is not following a normal distribution ($p\text{-value} < .05$). Table 3 presents the results of the Shapiro-Wilk test. The Histogram is shown in Figure 2. From the figure, it is noticeable that the data is not normally distributed.

Table 3. Test of Normality: Shapiro Wilk

	Statistic	df	Sig.
score	0.917	113	0.000

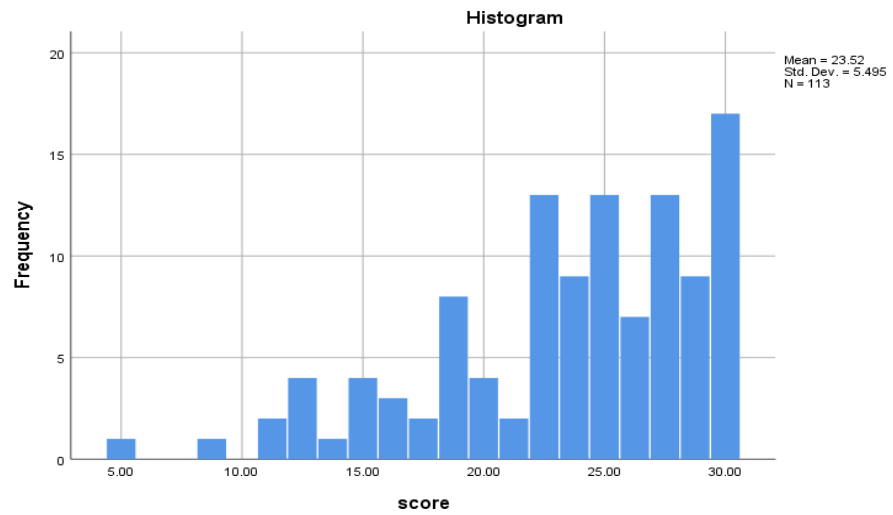


Figure 5. Data Distribution.

As illustrated below, it was found that the sample is homogenous, meaning the variances within the groups are equal (Table 4: p-values > .05). However, we were not able to conduct parametric test as t-test and ANOVA because of the normality. Our data considered not normally distributed, distribution-free tests are better because they don't assume that your data follow a specific distribution. Therefore, we conducted the Non-parametric tests as explained below.

Table 4. Test of Homogeneity of Variances: Levene Statistic

	Levene Statistic	Sig.
score	Based on Mean	.599
	Based on Median	.822
	Based on Median with adjusted df	.822
	Based on trimmed mean	.697

4.4 Non-Parametric test

4.4.1 Mann-Whitney test:

The Mann-Whitney test was used to compare the control and treatment sample means. Since the data did not represent a normal distribution, the Mann-Whitney test is suitable for this data to compare each school mean score separately. The null hypothesis of Wilcoxon rank sum (or Mann-Whitney) test assumes that it is equally likely that a randomly selected value from one sample will be less than or greater than a randomly selected value from the second sample. Table 4.3 presents the mean and sum rank of each group in school 1. The total Mann-Whitney score was 302.5, which suggests that there is no difference between control and treatment means where $p\text{-value}=0.098$ (see Table 5). Therefore, we fail to reject the null hypothesis.

Table 5. Mann-Whitney Test- Ranks (C1, T1)

class	N	Mean Rank	Sum of Ranks
C1	28	25.30	708.50
T1	29	32.57	944.50
Total	57		

Table 6. *Test Statistics^a*(C1, T1)

Test	score
Mann-Whitney	302.500
Wilcoxon W	708.500
Z	-1.654
Asymp. Sig. (2-tailed)	0.098

a. Grouping Variable: class

The same test was repeated with the second school. The results, shown in Table 6, indicate that there is no difference between the control and treatment group in school 2 as well (p-value=0.131) with a total Mann-Whitney score of 299.5 (see Table 7).

Table 7. Mann-Whitney Test- Ranks (C2, T2)

class	N	Mean Rank	Sum of Ranks
C2	27	25.09	677.50
T2	29	31.67	918.50
Total	56		

Table 8. Mann-Whitney (C2, T2)

Test	score
Mann-Whitney	299.500
Wilcoxon W	677.500
Z	-1.510
Asymp. Sig. (2-tailed)	0.131

a. Grouping Variable: class

4.4.2 Kruskal-Wallis Test

The Kruskal-Wallis test by ranks, or one-way ANOVA on ranks, is a non-parametric method whether samples are originated from the same distribution. In our data, the assumption of independent observation is met in the second school, and the outcome is ordinal. Therefore, we conducted the Kruskal-Wallis test to compare the four classes. The test showed that the four group mean rank was 52.91 for C1, 67.03 for T1, 47.17 for C2

and 60.07 for T2 as seen in Table 8 whereas the p-value was not significant (0.119) indicating no statistically significant differences in Math test scores between the four groups (see Table 9)

Table 9. Ranks – Kruskal- Wallis Test

class	N	Mean Rank
C1	28	52.91
T1	29	67.03
C2	27	47.17
T2	29	60.07
Total	113	

Table 10. Kruskal- Wallis (C2,T2)

score	
Kruskal-Wallis H	5.850
df	3
Asymp. Sig.	0.119

- a. *Kruskal Wallis Test*
- b. *Grouping Variable: class*

CHAPTER 5: DISCUSSION

This chapter provides a discussion of the results of the data analysis presented in chapter four. The chapter is divided into three main sections: summary, discussion of results, and recommendations for future research.

5.1 Summary of the Study

The main purpose of this study was to investigate how do metacognitive prompts, as facilitated by the EFAP-SRL model, impact achievement on the end of unit Math assessment among 8th grade girls. In particular, the researcher investigated the impact of metacognitive prompts, as facilitated by the EFAP-SRL model, on achievement on the end of unit Math assessment among 8th grade girls in two preparatory schools in Qatar. The hypothesis was as follow:

Research hypothesis: Students who participate in EFAP-SRL have higher achievement on the end of unit summative assessment compared to students who did not participate in EFAP-SRL.

Null hypothesis: There is no difference in achievement on the end of unit summative assessment between students who participate in EFAP-SRL and those who do not participate in EFAP-SRL.

To answer the main question, the EFAP-SRL model was implemented among two classes of 8th grade female students at different schools with $n = 29$ in each class. There were two control groups at each school with $n = 28$ in the first school and $n = 27$ in the second school. The two treatment group teachers participated in 3 training sessions regarding the implementation of the EFAP-SRL model.

5.2 Discussion of the results

5.2.1 The use of metacognitive prompts and their impact on achievement. A comparison of the means from school 1 revealed $M = 22.65$ for the control group and $M = 24.81$ for the treatment group; Similarly, school 2 revealed $M = 22.10$ for the control group and $M = 24.37$ for the treatment group. It was found that the treatment group performed slightly better compared to the control group; however, the results obtained were not significantly different. The analysis was done using non-parametric tests because the data was not normally distributed based on the results of Shapiro-Wilk test ($p\text{-value} < .05$) (refer to Figure 2). The results of the non-parametric tests suggest that the use of metacognitive prompts is not a mitigating factor when it comes to achievement on the end of unit assessment. We used the Mann-Whitney test to compare the control and treatment sample means at school 1; the total Mann-Whitney score was 302.5 and $p\text{-value}$ was 0.098, which indicated no significant difference between control and treatment means. We repeated the same test for school 2, and the results indicated that there is no significant difference between control and treatment group ($p\text{-value} = 0.131$) with 299.5 total of Mann-Whitney score. A possible explanation might be that time factor had an impact. Time is needed in order to notice the effectiveness of metacognition prompts as facilitated by EFAP-SRL model, and to improve students' achievements in mathematics. After analyzing the data using the Mann-Whitney test to compare control and treatment sample means, all the results indicated that there was no significant difference between treatment group and control group students on the test scores. However, using the results obtained by comparing the mean scores, we can conclude that the majority of the students received very high scores on the end of unit test, which lead to an increase in the means.

However, it could not be guaranteed if the difference between the mean scores were affected by the EFAP-SRL model or from students' efforts in revising the units covered. Moreover, the end of unit test itself may have affected the results. Based on the histograms shown in Figure 4 and Figure 5, all groups had left-skewed score distribution meaning, the majority of students received very high scores, between 25-30 out of 30. Only a few students failed or received low scores. Such high scores across four groups from two schools raises concerns regarding the difficulty of the test items. Additionally, it raises concerns as to whether the test covered all of the learning goals for the unit as well as the extent to which the test addressed the various levels of thinking (i.e. evaluation, synthesis, analysis, etc. Having high means because of the exceptionally high scores among all the groups may suggest that the items on the end of unit test do not differentiate between those students who have met the learning objectives and those who have not (see Table 11). The end of unit test consisted of six multiple choice items with each item having four choices. Each question had a weight of 1 point leaving the total with 6 points. The second part of the test consisted of nine problem solving questions with total weight of 24 points.

Table 11. Means of all participated groups

	N	Mean
C1	28	22.65
T1	29	24.81
C2	27	22.10
T2	29	24.37

The total points possible on the test is 30 points. At school 1 86% of the students in the control group passed the exam and 93% passed the exam in the treatment group. For school 2, 92% passed from the control group and 97% passed from the treatment group. This may suggest that the exam was too easy and did not effectively cover some of the higher order thinking skills such as analysis, synthesis and evaluation, which are well known to help differentiate those students who have met learning goals from those who have not. Because of the high scores across the board, the researcher decided to perform on item-analysis on the end of unit assessment. The items analysis provided the difficulty index (p) for each item. The proportion of students who answered an item correctly indicates the level of difficulty of a given item. The more students who answer N item correctly, the less difficult this item IS (Kelley, 1939). Table 5.2 presents the Difficulty index for each multiple-choice item. It is noticeable that the highest difficulty index was 1.00, which means all the students were able to solve the question. Perhaps the distractors used in these items were not plausible enough to sound like possible answers for the

students. However, having the lowest difficulty index equal only 0.64 indicates that the multiple-choice items as a whole were too easy. Based on the rule of thumb, we consider the question as a difficult question if it has a difficulty index of 0.25 or less. (Nitko & Brookhart, 2011). If the difficulty Index is between 0.75 – 1 then the item is considered as an easy question. Finally, a difficulty index of 0.25-0.75 indicates that the item is of an average difficulty.

Table 12. Difficulty Index for Multiple choice questions

	C1	T1	C2	T2
MC1	0.75	0.93	0.92	0.86
MC2	0.89	0.93	0.88	0.97
MC3	0.71	0.90	0.64	0.69
MC4	0.64	0.83	0.80	0.79
MC5	0.86	1.00	1.00	0.97
MC6	0.86	0.86	0.92	0.72

The same Item analysis was done for the nine problem-solving questions (See Table 5.3). All the questions in all schools and groups had a difficulty index higher than 0.65, which also indicates that those items were too easy.

Table 13. Difficulty Index for Problem Solving questions

	C1	T1	C2	T2
Q#7	0.79	0.79	0.96	0.86
Q#8	0.82	0.86	0.92	0.97
Q#9	0.82	0.86	0.92	0.97
Q#10	0.93	0.93	0.88	0.97
Q#11	0.68	0.66	0.64	0.72
Q#12	0.75	0.93	0.92	0.90
Q#13	0.89	0.93	0.96	0.97
Q#14	0.79	0.93	0.80	1.00
Q#15	0.86	0.93	0.92	0.76

These results shed some insight on the possible problems in the assessment itself, which are likely to be impacting the results of the study. Math tests are intended to be a kind of assessment tool that evaluates a student's skill in different areas of mathematics. If the test items are too easy and barely on the average, this will lead to difficulty in using these evaluations to inform instructions and to improve the achievement of students. Moreover, it does not differentiate between those students who met the objectives and those who did not. The quality of the end of unit test in the current study may have affected the results of the study leading to the negative results. However, these findings have opened up various doors into the assessment system and the validity of tests in mathematics

curriculum in Qatar, which will be discussed in more detail later in the suggestions and recommendations for future research.

The available findings of research on the understanding of metacognitive skills and SRL by young students from 11-14 years are dispersed and the discussion seems to be full of assumptions instead of significantly grounded arguments. Veenman et al. (2006) once argued that metacognitive skills start to develop at the age of 8-10. His assumption faced disagreement by Whitebread et al. (2009), who suggested that there are visible metacognitive skills in young children at age of 3-6 years old. Consequently, later studies confirmed the advantages of training self-regulation and metacognition in early grades (Hattie et al., 1996). The results of the current study support the notion that metacognitive skills are easier regulated among older students. Research on self-regulation and metacognition has found that primary school students seem to lack the significant strategies for SRL (Cromley, et al. 2010). Therefore, the majority of self-regulation research has focused on high school students. Finding no significant differences between the control groups and treatment groups across the two schools used in this study may indicate that students at this age have difficulty utilizing metacognitive strategies to support their learning. This will be discussed in more depth in the next section.

5.2.2 Use of assessment and instructions to develop metacognitive skills among primary and preparatory school students. Findings assures the importance of developing metacognitive skills at young ages because of the lack of significant SRL strategies and metacognitive skills at this age. For example, related studies have shown that in preparatory education, metacognition and SRL programs should be focused on constructing a strategy repertoire (Dignath & Buttner, 2008). Young students' metacognitive knowledge is still

developing, and therefore they may benefit more from instruction that combines effective formative assessment and models or scaffolds the use of metacognitive prompts. Combining instructional strategies and formative assessments with metacognitive prompts can make distinct contributions to developing students' learning and achievement in young ages (Black & William, 2003, 2009). The findings of this study introduce various questions regarding the training of self-regulation and metacognition in the primary and secondary classroom. Do teachers use instructions that develop metacognitive skills? Are the instructional strategies used in the class aligned with the assessment tools to improve higher-order thinking skills in students?

This paper may align with previous research suggesting that secondary school students demonstrate richer metacognitive skills and regulation compared to primary and preparatory school students (Dignath & Buttner, 2008). It is important that students gain experience in strategy use in order to change their strategic behavior. Additionally, metacognitive prompts and interventions seem to be more effective the longer they are implemented.

Besides the duration of the intervention, the poor knowledge of self-regulation and metacognition training by schoolteachers may have affected the study. Many teachers may need to receive training regarding ways of developing SRL and metacognition among students. (Woeytens et al., 2002). In the present study, the researcher provided three sessions to the teachers of the treatment groups. However, this may have not been enough for teachers who had never dealt with the ways of improving SRL and metacognition skills in their students before. In order to conduct such a study in Qatar, teachers may need an

expanded period of preparing and training to get used to the implementing of metacognition prompts.

The study took one semester to be done (Almost three months). The time period is short comparing to other previous related studies and this may have affected the results. More time should be allocated to train the students and get them to be familiar with metacognition prompts. A long-term investigation comparing the use of metacognition or SRL among students who were taught to do so at a younger age and those who were not would be very interesting.

5.3 Limitations and suggestions for research

5.3.1 Limitations

Sample size. Quantitative studies with a larger number of participants have proved to produce more statistically reliable results. The main results usually have 95% confidence intervals (CI). The sample size affects the width of these intervals. In other words, studies with a larger number of participants produce narrow intervals, which leads to more precise results (Pocock, 1983). The sample size used in this study was $n = 113$, a relatively small number in comparison to other related studies; comparing to similar studies, this is a relatively small number. For example, in a study of first year college students conducted by Hudesman et. al. (2013), the sample size was $n = 1198$. (Hudesman et al., 2013). This larger sample size makes it easier to achieve statistical significance. On the other side, however, one could argue that a very large sample size may be viewed as a kind of manipulation of the data in that the sample size alone has been shown to be related to statistical significance. Despite this argument, a larger sample size was not achievable in

this study for several reasons. Firstly, increasing the sample size would have meant including more teachers and possibly more schools, which would have presented additional uncontrollable variables. Moreover, enlarging the sample size would have also involved more time for training teachers in the EFAP-SRL model as well as significantly more time on data collection and analysis. Regardless, this study may be useful to support larger confirmatory studies in the future.

Time Factor. One of the major limitations faced during this research was time. Students being unfamiliar with the metacognitive prompts did take up some of the time that was allocated for the implementation. The study was conducted over one semester (approximately three months). The time period is short comparing to other previous related studies and this may have affected the results. If this study or similar one was done again, a longer time period would be better, at least one academic year in order to notice the change in students' metacognitive skills and self-regulation.

Paper Reflection forms. The researcher depended on the self-reflection form drawn from the EFAP-SRL model from Zimmerman's (2003) work to stimulate students' metacognition; this method was also used in various other studies (Hudesman et al., 2013). However, the age of our participants is younger and an oral explanation of thinking about thinking could work better. Students perform better orally compared to written tests at the primary and preparatory levels because oral language may supplement body language and carry more emotional charge Carter (2008).

5.3.2 Recommendations and suggestions for future research

The main objective of this study was to investigate how do metacognitive prompts, as facilitated by the EFAP-SRL model, impact achievement on the end of unit Math assessment among 8th grade girls. and use the results to investigate the effectiveness of such programs in enhancing student learning and achievements. Results of this study have a few implications for teachers, educators and stakeholders. In addition, educational researchers can consider this study as a starting point for extra research examining the impacts of metacognitive prompts at young ages in Qatar. Based on the data analysis, findings and discussion, we offered the following recommendations.

Since one of the limitations was students being unfamiliar with the metacognitive prompts used did take up some of the time that was allocated for the implementation. For future studies, it is recommended to have more time to be spent on introducing the program to the students and hopefully the study in the future would yield significant result. A highly recommended study that could be looked at in the future would be to examine the impact of formative assessment on self-regulation skills, and metacognition in primary and preparatory schools. In addition, a deeper research is needed to investigate having a training program that is constructed to develop metacognition that focus on developing a strategy repertoire.

Moreover, to ensure the validity of the test in upcoming studies, we suggest that the researcher and a supervisor of the evaluation institute at MOE (Ministry of Education) discuss the complexity of the problems and build a Test Descriptive Table that guarantees that the exam includes items that measure s higher-order thinking. Teachers can follow this descriptive table and build a valid exam. This option was not allowed in our study because

of the time restrictions and some existing MOE policies. Another suggestion is having an “Item Bank” constructed by highly qualified mathematics teachers and supervisors. This may help in reducing the possibility of having easy tests that affect the study results. At a minimum, teachers could receive professional development on performing item analysis on their assessments and developing their own item bank to help ensure they are preparing strong assessments. As a one in kind study in Qatar, we recommend having a replication study that fills the gaps we highlighted and focuses on metacognition skills among young ages. This can be done for a longer time period in order to train students and teachers in how to deal with metacognitive prompts included in formative assessments such as quizzes. And, of course, to use a valid summative test as a measurement tool.

Finally, as we mentioned before, oral explanation of thinking about thinking could work better in young ages. Studies found that students perform better orally compared to written tests at primary and preparatory schools age because oral language may supplement body language and carry more emotional charge (Carter, 2008). Therefore, we recommend using oral questioning to let the students in young ages express their way of thinking in a better way, which may lead to better results.

REFERENCES

- Allal, L. (2010). Assessment and the regulation of learning. *International encyclopedia of education*, 348-352. Oxford: Elsevier.
- Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning*, 4(1), 87– 95. doi:10.1007/s11409-009-9035-7
- Azevedo, R., Johnson, A., Chauncey, A., and Burkett, C. (2010). Self-regulated learning with MetaTutor: advancing the science of learning with metacognitive tools. *New Science of Learning*, 225–247.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1997). Self-efficacy: The exercise of control. New York.
- Bannert, M. (2015). Promoting Self-Regulated Learning Through Prompts. *Zeitschrift Für Pädagogische Psychologie*, 23(2), 139-145. doi:10.1024/1010-0652.23.2.139
- Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, 52, 293–306. doi:10.1016/j.chb.2015.05.038
- Beekman, K., Brinke, D., & Boshuizen, H. (2015). Developing young adolescents' self regulation By means of formative assessment: A theoretical perspective. *Cogent Education*.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Education*

Assessment, 21(1), 5-31.

Bloom, S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York.

Bodvarsson, M., (2005). Prompting students to justify their response while problem solving: A nested, mixed-methods study. *ETD collection for University of Nebraska*.

Brinke, D., & Boshuizen, H. (2015). Effects of formative assessments to develop self-regulation among sixth grade students: Results from a randomized controlled intervention. *Studies In Educational Evaluation*, 51, 126-136.
doi:<http://dx.doi.org/10.1016/j.stueduc.2016.10.008>

Butler, D., & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245-281.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational psychological reviews*, 205-249.

Cleary, T., Callan, G., & Zimmerman, B. (2012). Assessing Self-Regulation as Cyclical, Context-Specific Phenomenon: Overview and Analysis of SRL Microanalytic Protocols. *Education Research International*, 2012, 1-19.
doi:10.1155/2012/428639

Cromley, J., (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Education Psychology*, 96(3), 523-535.

Dewey, J. (1938). What Is Social Study? *The Modern Language Journal*, 23(3), 219.
doi:10.2307/317271

Fletcher, A., & Shaw, G. (2012). How does student-directed assessment affect learning?

- Using assessment as a learning process. *International Journal of Multiple Research Approaches*, 6(7), 245-263.
- Grant, A. M. (2003). The impact of life coaching on goal attainment, metacognition, and mental health. *Social Behavior and Personality*, 31(3), 253-264.
- Grant, M., & Greene, J. (2001). *It's your life. What are you going to do with it?* Cambridge, MA: Perseus Publishing.
- Greene, J. A., Dellinger, K. R., Tüysüzoğlu, B. B., & Costa, L. (2013). A Two-Tiered Approach to Analyzing Self-Regulated Learning Data to Inform the Design of Hypermedia Learning Environments. *International Handbook of Metacognition and Learning Technologies Springer International Handbooks of Education*, 117-128. doi:10.1007/978-1-4419-5546-3_8
- Gronlund, N., & Linn, R. (1990). *Measurement and evaluation in teaching (6th ed.)*. New York: Macmillan.
- Hadwin, A. F., Oshige, M., Gress, C. L., & Winne, P. H. (2010). Innovative ways for using gStudy to orchestrate and research social aspects of self-regulated learning. *Computers in Human Behavior*, 26(5), 794-805. doi:10.1016/j.chb.2007.06.007
- Hudesman, J., Crosby, S., Flugman, B., Isaac, S., Everson, H., & Clay, D. (2013). Using Formative assessment and metacognition to improve student achievement. *Journal of Developmental Education*, 37(1).
- Kelley, L. (1939). The selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*, 30, 17-24.

- Kitsantas, A., Zimmerman, B. & Cleary, T. (2000). The role of observation and emulation in the development of athletic self-regulation. *Journal of Educational Psychology, 92*(4), 811-817.
- Mahmoodi, M. H., Kalantari, B., & Ghaslani, R. (2014). Self-Regulated Learning (SRL), Motivation and Language Achievement of Iranian EFL Learners. *Procedia - Social and Behavioral Sciences, 98*, 1062-1068. doi:10.1016/j.sbspro.2014.03.517
- Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students*. Boston, MA: Pearson/Allyn & Bacon.
- Panadero, E., & Andrade, H. (2018). Fusing self-regulated learning and formative assessment: A roadmap of where we are, how we got here, and where we are going. *Australian Educational Researcher, 45*(1), 13-31. doi:10.1007/s13384-018-0258-y
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129–144.
- Panadero, E., & Tapia, J. (2014). How do students self-regulate? Review of Zimmerman’s cyclical model of self-regulated learning. *Anales de Psicología, 30*(2), 450–462.
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74- 98. doi:https://doi.org/10.1016/j.edurev.2017.08.004
- Pintrich, P., & DeGroot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1),

33–40.

- Sasai, L. (2017). Self-Regulated Learning and the Use of Online Portfolios: A Social Cognitive Perspective. *Journal of Educational and Social Research*, 7(2), 55-65.
doi:10.5901/jesr.2017.v7n2p55
- Syarif, M., & Satriani, R. (2016). The effect of formative testing and self-directed learning on mathematics learning outcomes. *International Electronic Journal of Elementary Education*, 8(3), 507-524.
- Tay, H. (2015). Setting formative assessments in real-world contexts to facilitate self-regulated learning. *Educational Research for Policy and Practice*.
- Veenman, F., (2003). Alternative assessment of strategy uses with self-report instruments: a discussion. *Metacognition and learning*, 6(2), 205-211.
- William, D. (2011). What is assessment for learning. *Studies in Educational Evaluation*. 37(2011), 3-14.
- Zimmerman, B. (1998). Academic studing and the development of personal skill: A self-regulatory perspective. *Educational Psychologist*, 33(2), 73-86.
doi:10.1207/s15326985ep3302&3_3
- Zimmerman, B. (2000). Attaining self-regulation: A social cognitive perspective. San Diego, CA: Academic Press.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70.
- Zimmerman, B., Moylan, A., Huesman, J., White, N., and Flugman B. (2011). Enhancing self-reflection and mathematics achievement of at-risk urban technical college students. *Psychological Test and Assessment Modeling*, 53(1), 108-127.

Zimmerman, B.J. (2006). Integrating classical theories of self-regulated learning: A cyclical phase approach to vocational education. *Self-regulated learning in vocational education*, 7-48

APPENDICES

Appendix A

Before solving each problem, how confident are you that you can solve it correctly?	<table border="1"> <tr> <td>School</td> <td>الاسكندرية Prep. School</td> <td>المدرسة</td> <td>الاعدادية للبنات</td> </tr> <tr> <td>Exam</td> <td>Math</td> <td>تقديم</td> <td>رياضيات</td> </tr> <tr> <td>Level</td> <td>8</td> <td>لن</td> <td>المستوى</td> </tr> <tr> <td>Academic Year</td> <td>2018/2019</td> <td>التاريخ</td> <td>2018/2019</td> </tr> <tr> <td>Semester</td> <td>first</td> <td>الفصل الدراسي</td> <td>الاول</td> </tr> </table>	School	الاسكندرية Prep. School	المدرسة	الاعدادية للبنات	Exam	Math	تقديم	رياضيات	Level	8	لن	المستوى	Academic Year	2018/2019	التاريخ	2018/2019	Semester	first	الفصل الدراسي	الاول	Before solving each problem, how confident are you that you can solve it correctly?
School	الاسكندرية Prep. School	المدرسة	الاعدادية للبنات																			
Exam	Math	تقديم	رياضيات																			
Level	8	لن	المستوى																			
Academic Year	2018/2019	التاريخ	2018/2019																			
Semester	first	الفصل الدراسي	الاول																			
0% 25% 50% 75% 100%	<p>الإسم: : الصف : التطبيق الاول</p> <p>السؤال الاول:</p> <p>استعمل خواص الأسس لكتابة مقدار مكافئ لكل مقدار من المقادير التالية:</p> <p>1) $6^4 \times 6^3$ 2) $7^3 \times 2^3$ 3) $(9^2)^5$ 4) $4^{10} \div 4^4$ 5) $(189345)^0$</p> <p>السؤال الثاني:</p>	0% 25% 50% 75% 100%																				
0% 25% 50% 75% 100%	<p>أكتب كل مقدار مما يلي باستعمال أسس موجبة:</p> <p>1) 9^{-4} 2) $\frac{1}{3^{-5}}$</p>	0% 25% 50% 75% 100%																				

Appendix B

SRL Math Revision Sheet, Quiz # ____ Item # ____

Student: _____

Date: _____

Instructor: _____

Now that you have received your corrected quiz, you have the opportunity to improve your score. Complete all sections thoroughly and thoughtfully. Use a separate revision sheet for each new problem.

PLAN IT

8pts

1. a. How much time did you spend studying for this topic area? _____
b. How many practice problems did you do in this topic area _____ in preparation for this quiz?
(circle one) 0-5 / 5-10 / 10+
c. What did you do to prepare for this quiz? (use study strategy list to answer this question)

2. After you solved this problem, was your confidence rating too high (i.e. 4 or 5)? yes no

3. Explain what strategies or processes went wrong on the quiz problem.

PRACTICE IT

8pts

4. Now re-do the original quiz problem and write the strategy you are using on the right.

Definitely not Not confident Undecided Confident Very confident
confident

5. How confident are you now that you can correctly solve this similar item?

1 2 3 4 5

6. Now use the strategy to solve the alternative problem.

4pts

7. How confident are you now that you can correctly solve a similar problem on a quiz or test in the future?

1 2 3 4 5

4pts