



2014-03-17

# The Effect of Formative Assessments on Language Performance

Brian W. Radford

*Brigham Young University - Provo*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Educational Psychology Commons](#)

---

## BYU ScholarsArchive Citation

Radford, Brian W., "The Effect of Formative Assessments on Language Performance" (2014). *All Theses and Dissertations*. 3978.  
<https://scholarsarchive.byu.edu/etd/3978>

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

The Effect of Formative Assessments on Language Performance

Brian W. Radford

A dissertation submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Richard R Sudweeks, Chair  
David D. Williams  
Stephen C. Yanchar  
Joseph A. Olsen  
Andrew S. Gibbons

Department of Instructional Psychology & Technology

Brigham Young University

March 2014

Copyright © 2014 Brian W. Radford  
All Rights Reserved

## ABSTRACT

### The Effect of Formative Assessments on Language Performance

Brian W. Radford

Department of Instructional Psychology and Technology, BYU

Doctor of Philosophy

This study sought to improve the language learning outcomes at the Missionary Training Center in Provo, Utah. Young men and women between the ages of 19-24 are taught a foreign language in an accelerated environment. In an effort to improve learning outcomes, computer-based practice and teaching of language performance criteria were provided to missionaries in an effort to allow them to progress at their own pace outside of the classroom. The effect of computer-based practice and the teaching of language performance criteria were assessed in a 2x2 factorial design. The dependent variable was speaking proficiency in the Spanish language. This variable was assessed in two different ways: (a) human-rated speaking proficiency and (b) computer-scored speaking proficiency. Results suggest that the teaching of language performance criteria increases speaking proficiency over those who are not taught the criteria. Missionary trainee responses also indicate that the training of criteria helped the trainees to evaluate their own performance and the performance of other language speakers. Missionary trainees also reported that this training helped them to see their own progress and to set appropriate learning goals.

Keywords: formative assessments, assessment for learning, teacher feedback, student feedback, language performance

## TABLE OF CONTENTS

ABSTRACT .....	ii
List of Tables .....	vi
List of Figures .....	vii
Chapter 1: Introduction .....	1
Background .....	1
New Challenges.....	2
Areas of Focus.....	3
Statement of Purpose.....	4
Research Hypotheses.....	4
Chapter 2: Review of Literature .....	6
Introduction .....	6
Improving Learning Through Assessment.....	7
Summative assessment .....	8
Formative assessment .....	8
Evolution of Formative Assessment .....	11
Assessments for learning .....	11
Alternative assessment .....	12
Dynamic assessment.....	12
Self-Assessment as Formative Assessment .....	13

Self-assessments .....	14
Self-assessments influence on self-efficacy .....	15
Self-monitoring.....	16
Formative Assessments in the Second Language (L2) Classroom .....	17
Language Learning Challenges with Formative Assessment .....	18
Moving Formative Assessment Research Forward in L2 Environments.....	19
Chapter 3: Method .....	22
Design.....	22
Participants .....	25
Instrumentation Overview.....	26
Instrumentation details .....	28
Instrumentation example .....	29
Procedure.....	31
Analysis.....	31
Chapter 4: Results .....	34
Assessments of the Various Components .....	34
Multivariate Analysis of the Relationship Between Treatments and Language Performance..	38
Missionary Attitudes Towards Language Learning.....	40
Overall language learning experience .....	42
Computer assisted rater instruction experience.....	43

Effect of CARI on trainees' goals and study patterns .....	44
Language speaking assessment experience .....	46
Chapter 5: Discussion .....	49
Research Hypotheses.....	49
Language performance practice.....	49
Understanding language performance criteria.....	50
Interaction between language performance practice and understanding criteria .....	50
Recommendations .....	51
Conclusion.....	52
References.....	54
Appendix A: Handouts for Missionaries .....	64
Appendix B: Language Criteria Rubric .....	68
Appendix C: Entrance Questionnaire .....	69
Appendix D: Exit Questionnaire.....	73

## List of Tables

Table 1: Experimental Groups and Treatment Conditions .....	23
Table 2: Description of 2 x 2 Factorial Design.....	24
Table 3: Number of Missionaries by Type of Instruction Provided .....	25
Table 4: Prior Language Experience by Experimental Group.....	26
Table 5: Formative Language Assessments.....	27
Table 6: Formative Language Assessment Usage .....	28
Table 7: Descriptive Statistics for the Four Groups on the Pretest.....	35
Table 8: Descriptive Statistics for the Four Groups on the Posttest .....	36
Table 9: Descriptive Statistics for Minutes Spent in Practice for Groups 2 and 4 .....	37
Table 10: Results of ANCOVA on the Posttest with a Pretest Covariate .....	38
Table 11: Results of HLM on the Posttest with a Pretest Covariate and Random Class Effect..	39
Table 12: Results of HLM on the Posttest with Random Effect and no Covariate .....	40
Table 13: Missionary Questionnaire Responses .....	40
Table 14: Descriptive Statistics for Attitudinal Scores by Group by Item .....	41
Table 15: Overall Language Learning Experience — Missionary Responses .....	42
Table 16: CARI Experience — Missionary Responses.....	43
Table 17: Missionary Response on the Effect of CARI on Goals and Study.....	45
Table 18: Language Speaking Assessment Experience — Missionary Responses.....	47

## List of Figures

Figure 1: Example instructions for language grammar assessment.....	29
Figure 2: Example of questions that are context dependent. ....	30
Figure 3: Example feedback with description of correct response.....	30



## **Chapter 1: Introduction**

### **Background**

The missionary trainees at the Latter-day Saint (LDS) Missionary Training Center in Provo, Utah are quite unique. These trainees volunteer 18 months (females) or 24 months (males) of their life to serve as missionaries for their church. Since many of these missionaries are called to serve in foreign countries, they are required to learn a foreign language. With little to no previous experience speaking the language in which they are called to serve, students are highly motivated to learn the language before arriving in their assigned country. Their language learning experience spans 8-14 weeks depending on the language to which they are assigned. The teachers are non-professionals: they are typically college students in varying fields of study who for the most part qualify to teach only because they previously served as a missionary in a geographic area where the language which they are now expected to teach was spoken.

With the limited training time before these missionaries are sent to their assigned countries, these missionaries must learn basic building blocks that will enable them to understand, speak, read, and listen in a new language and to also continue their learning throughout their mission. The training program is focused on allowing the trainees to progress at their own pace with the challenge that these missionaries are organized into classes typically consisting of 8-12 students, who are grouped together merely because of the common target language and country in which they are assigned to serve. The levels of existing language skills vary greatly within each class and the students' abilities to gather and process the instruction vary greatly as well. These challenges make it difficult for a non-professional teacher to adjust classroom teaching to meet individual needs. Also, the missionaries are challenged with knowing how to monitor and assess their own learning. These challenges make it difficult to

assist the missionaries with the tools and the capacity to continue their learning process once they leave the formal training facilities.

In past years, the missionaries were expected to memorize lessons in their target language. The approach now is for the missionaries to learn basic language concepts and skills that allow them to speak more proficiently in their own words. Aspects of memorization remain in practice with regards to gaining vocabulary in the target language, but memorization of content and flow in a predefined order has been discontinued. Memorizing the lesson content had allowed for missionaries to chunk or group the content into logically related topics. This chunking allowed the students to assess where they were in their progress and how much more they needed to memorize before being prepared for their service. As a result of memorization, students were teaching only the words they had memorized and were not expanding their language abilities or their understanding of the content they were teaching. This memorization allowed most students to feel too comfortable with their level of learning and their teaching was done in a rote manner with limited additional effort to improve. Even though recent changes have addressed these latter issues, there has not been an obvious replacement provided yet for the built-in chunking and built-in progress tracking.

### **New Challenges**

As part of the missionary training program, language materials are presented in the classroom along with practice time. Additional time is then provided outside of the classroom to allow missionaries to either (a) process what they have been taught and catch up with their understanding or (b) study extra materials and push their learning beyond that of what was presented in the classroom. Language learning tools are also provided in computer labs that allow for variety in missionary learning. One of the language learning tools available is a means

of formative assessment in hopes to better assist the students in identifying weaknesses and strengths to enable them to better use this unstructured block of time.

The application of formative assessments to language performance has not been tested yet, but there have been some promising indications that it can enhance language performance learning outcomes as well (Radford, 2010). Previous research at the Provo Missionary Training Center (MTC) has demonstrated that the use of formative assessments (Radford, 2010) enhances learning outcomes in non-language areas. These formative assessments have helped to provide feedback to both missionaries and teachers and have enhanced the learning process. This formative approach differs greatly from the summative approach that is used in many teaching environments.

The rationale for conducting this study is based on the assumption that formative assessments will (a) enable missionaries to identify their own strengths and weaknesses by allowing them to compare their speech samples with native and proficient speakers, and (b) enable missionaries to receive rater feedback which identifies specific strengths and weaknesses in order to use their time wisely to advance their learning. The researcher hypothesized that both internal (self-comparison) and external (rater provided) feedback will enable missionaries to adjust the levels of time and effort that they invest into different language areas. This adjustment in study time and effort should also end up increasing their base language performance levels beyond that of what they would be able to do without the formative assessments.

### **Areas of Focus**

The MTC administrators expect the missionaries to adjust their study activities to meet their individual needs. Each class varies greatly in their prior knowledge and understanding of the languages being taught, as well as how quickly they learn what is being taught. Teachers are

expected to teach at the pace and depth of content coverage to match the readiness and abilities of the missionaries in each class. This is a huge challenge because the teachers only receive minimal training before they begin teaching. Through the use of formative assessments, missionaries will be enabled to focus on their own individual strengths and weaknesses during their out of class time.

### **Statement of Purpose**

The purpose of this study was to investigate the effects of providing formative feedback to missionary trainees regarding their language performance abilities. The researcher hypothesized that the use of formative assessments and frequent feedback would improve achievement of the intended learning outcomes. The researcher also hypothesized that missionaries will reach higher achievement of the intended learning outcomes if they are taught relevant criteria by which they can judge their own performance and progress.

### **Research Hypotheses**

A language speaking assessment (LSA) provided written questions to which the missionaries responded verbally and the assessment recorded the audio of that response. The LSA provided native and proficient speaking sample responses to each question and allowed the missionaries to compare their responses to these samples. The LSA also allowed for missionaries to re-practice and compare their responses. Raters were also able to provide formative feedback to the LSA responses. A generalizability study allowed for the testing of different facets and their interaction. The levels of variables were no feedback (control group), internal feedback (missionary comparison to provided samples), and external feedback (rater provided).

For the purpose of this study, language performance was measured by speaking ability in the language that the missionaries are learning. The LSA provided feedback in the categories of fluency, grammar, pronunciation, and vocabulary used in their spoken responses. These were the categories used to rate the missionaries' language performance. The study focused on the following hypotheses:

1. Missionary trainees who practice their language performance with a computer, comparing their performance to native and proficient speakers will obtain higher posttest scores on average than missionaries who did not have the opportunity for computer-based practice.
2. Missionary trainees who are taught the criteria by which their language performance is to be rated will obtain higher posttest scores on average than missionaries who are not taught the criteria by which their language performance was rated.
3. Missionary trainees who both practice their language performance with a computer and are taught the criteria by which their language performance was rated will obtain higher posttest scores on average than missionaries who experienced only one or neither of these treatments.

## Chapter 2: Review of Literature

### Introduction

Black and Wiliam (2006) asserted that “assessment in education must, first and foremost, serve the purpose of supporting learning” (p. 9). For over 20 years, studies have shown that assessments can and do serve the purpose of learning. Researchers have worked using assessments for closing the gap between what the student has already learned and what potentially may be learned. Although this research has shown that assessments can support learning, there is still much work to do in the area of using this approach in the second language classroom.

This review will discuss the evidence supporting the use of formative assessments as a means of facilitating learning outcomes in classroom contexts. Several benefits of formative assessments will be discussed such as (a) empowering learners to recognize the areas in which they can improve themselves, (b) improving learner motivation, and (c) raising a level of awareness for both the learner and teacher during the learning process. After summarizing the benefits of formative assessments, there will be discussion of how self-assessments are a type of formative assessment that can be used to assist students in their learning process and can have the same effects as other types of formative assessment. Finally, there will be discussion on how these same approaches can be used in the second language classroom, but little research has been done in this area. Some challenges of applying a formative assessment approach in a second language classroom will also be reviewed.

Key articles were used to begin the process of establishing a baseline of the benefits of formative assessments. This effort resulted in much research and so key articles were identified through review with research area experts. These articles included Black and Wiliam (1998a),

Crooks (1988), and Fontana and Fernandes (1994). The literature search was conducted by several methods. First, these key articles were reviewed followed by a citation search on these articles. As this was not the key point of the literature review, this search was not exhaustive, but did try to identify key benefits and studies regarding formative assessments. A second approach was to search by keywords in the ERIC, ProQuest, and EBSCO databases. In an effort to make this an efficient approach, tools were used to identify common research terms in an effort to better identify key terms used in publication. A third approach was to review dedicated journals to these topics from the language testing community, these journals were *Language Testing* which devoted a special issue in 2001, and *Language Assessment Quarterly* devoted a special issue in 2007. Finally, articles that were referenced in these journals were reviewed and sources that cited these articles as well.

The next two sections deal with establishing a baseline that formative assessments enhance learning outcomes. These sections are meant as an overview of some key studies and key benefits that have resulted from these studies. Although, much work has been done in this area, the second section shows that this area is still evolving and there is still more growth and understanding to be obtained.

### **Improving Learning Through Assessment**

Black and Wiliam (1998a) reviewed studies of assessment practices in the areas of mathematics, science, and other general subjects in the UK, and they reviewed representative examples from 250 studies identified as ‘sufficiently important’ from an original 681 publications on the topic. This study claimed that formative assessments are both an essential component to classroom work, and that development in this area can raise standards of achievement. This paper reviewed several studies that found formative assessments yielded

greater learning gains than that of the conventional teacher-dominated summative assessment practices.

***Summative assessment.*** Summative assessment refers to an assessment given at the end of a learning period to determine if learning occurred, and often to place some value (score) on how much learning had occurred or to quantify how much a learner knows about the subject matter. Learning may be a secondary benefit as a result of taking a summative assessment, but the primary purpose is to *measure* learning and to make informed inferences about the learner's ability or level of achievement (Atkin, Black, & Coffey, 2001).

Often teachers' intentions for using summative assessments do not match that of their actual practices. Some teachers intend to use assessment as a means of assisting the learning process, but due to the timing and approach of summative assessments the practice contradicts intention. Bol and Strage (1996) showed this by interviewing 10 high school biology teachers individually and then comparing their teaching philosophies and practices to that of their intentions. These teachers wanted their students to develop a (a) general interest in biology, (b) general understanding of biology and its real-world applications, and (c) higher order study skill of interpreting information. Their practices did not, however, support these goals. Nearly two thirds of test and practice items were simply recognition items and interviews showed that most of these teachers were not aware of the contradiction between their goals and their assessment practices.

***Formative assessment.*** In contrast to a summative assessment in which the teacher or institution judge the achievement, formative assessments allow the learner to judge their own learning achievement. Summative assessments are often administered at the end of the teaching of a learning block as an effort to provide an overall impression of how well the topic(s) was



learned, while formative assessments are administered throughout the learning process as an effort to inform both teacher and learner during the learning process. The Assessment Reform Group (ARG) defines formative assessment as *the process of seeking and interpreting evidence for learners and their teachers to decide where the learners are in their learning, where they need to go, and how best to get there* (Assessment Reform Group, 2007).

Several studies have shown that formative assessments can play a critical role in the process of empowering the learner and enhancing classroom learning (Bachman, 2005; Black & Wiliam, 1998a, 1998b; Fontana & Fernandes, 1994; Geeslin, 2003; Harlen & Winter, 2004; Shohamy, 2001; Van de Watering & Van der Rijt, 2006; Whiting, Van Burgh, & Render, 1995). Studies have also shown that formative assessments can both (a) motivate learners and (b) result in greater gains in assessment scores (Brookhart & Durkin, 2003; Harlen & Winter, 2004; Rea-Dickins & Gardner, 2000; Weeden & Winter, 1999). Motivation of learners, as a result of assessments, was seen as Brookhart and Durkin (2003) reviewed a single-case, descriptive study regarding a teacher-researcher in an urban high school where their entire teaching load was studied. The teacher-researcher taught courses on 10th-grade world cultures, honors 11th-grade US history, and philosophy which included mostly 12th-grade students. There were 96 students observed across these courses who participated in 12 classroom assessment events, four across each course. The results evidenced a positive correlation as a result of performance assessments and suggested that both internal and external sources of motivation (Ames, 1992; Covington, 1992) can result from formative assessments administered by a classroom teacher.

Additional studies have shown that formative assessments help teachers identify the needs of the learner and foster learning in the classroom environment (Shohamy, 2001; Triggs, Weeden, Winter, & Broadfoot, 2000; Weeden & Winter, 1999). By using smaller more frequent

assessments, teachers are able to collect and review feedback throughout the learning process and not solely as a summative distinguishing factor upon completion of a topic. This frequent assessment feedback allows both the teacher and the student to gain insights on progress during the learning process. Chickering and Gamson (1991) showed that frequent feedback as a result of formative assessments can keep students on task by helping to identify areas in which the student is not performing well. This prompt feedback informs students while they are planning their individual study plans and strategies.

A central premise to formative assessments is the goal of assisting the learner in establishing an index to their own learning. Probably the most common and almost intuitive form of formative assessment is that of question-and-answer during the teaching of a lesson. This allows a teacher and a student to gain instant feedback on understanding and learning. In this sense formative assessment is obviously not a new concept. Socrates' preferred way of teaching was to question the learner, using questions to promote higher order thinking and foster learning (Gareis, 2006). Formative assessment as an instructional method enables learners to contribute to their own process of learning by providing a quick index to their learning during the instruction process. Students in this situation are then enabled to ask questions and express non-understanding of the lesson while the teacher is still discussing the topic.

As many researchers and teachers view these benefits from formative assessments there has been a focus shift from the conventional summative testing methods towards a focus on formative assessments. Part of this shift has resulted from a shift of centralized control and authority to that of joint teacher and learner control of the learning process (Davidson & Lynch, 2002; Lynch, 2001, 2003; McNamara, 2001). Although benefits have been seen from formative assessment approaches, there are some caveats in that fairness and accuracy depend on informed

and responsible teachers and also on the experience of learners in the area of self-assessment (Ross, 1998).

Assessment practices and studies have progressively moved from a focus on objective mastery of content to a formative assessment of the effort and contribution that the learners make to the process of learning. This shift appears to be a result of educators valuing the contribution of the learner to the process of learning (Boston, 2002; Chatterji, 2003). This can be seen as more educators are gradually integrating formative assessments into their conventional summative testing efforts (Davison, 2004).

### **Evolution of Formative Assessment**

Over the years, formative assessments have evolved into a dynamic and reflective approach called *Assessment for Learning* (AFL) which is an interactive, learning-focused pedagogy (Colby-Kelly & Turner, 2007). This AFL approach requires a teacher to encourage the use of student assessment in order to provide useful feedback to the learner and adapt lesson planning to the needs of the students that are identified through the assessment process (Black & Wiliam, 1998a; 1998b; 2005; 2006).

*Assessments for learning.* The term Assessment for Learning was crafted by the ARG (Olson, 2005) and supported by Stiggins (2001, 2002) as an effort to clarify the intention of formative assessments. The main idea was to not focus on the frequency of formative assessments, but on the outcome of promoting learning and empowering the learner to take control of their own learning process.

In an effort to help clarify the definition and intentions of Assessments for Learning, Stiggins provided seven strategies (Stiggins, Arter, Chappuis, & Chappuis, 2004). Stiggins used these strategies also to ensure systematic student involvement in the formative assessment

process (Stiggins, 2007). Recent studies have also used these strategies in an effort to help clarify the definition of Assessments for Learning that will be used in their studies (Forbes, 2007). These seven strategies are:

Where am I going?

1. Provide clear learning targets
2. Use examples of strong and weak student work

Where am I now?

3. Offer regular descriptive feedback
4. Teach students to self-assess and set goals

How can I close the gap?

5. Design lessons to focus on one aspect at a time
6. Teach students focused revision
7. Engage students in self-reflection, and let them keep track of and share their learning

***Alternative assessment.*** With a conventional assessment, a learner chooses a response from a given list, e.g. multiple-choice, true/false, or matching. Alternative assessment is any type of assessment in which the learner *creates* a response when presented with a question or task, e.g., short-answer, performance assessments, oral presentations, or demonstrations. Alternative assessment is a process involving both learner and teacher in making judgments about progress as opposed to using strategies which only involve the teacher making the judgments (Hancock, 1994).

***Dynamic assessment.*** Dynamic Assessment (DA) also appears much in language testing literature. DA is derived from Vygotsky's theory of the Zone of Proximal Development (ZPD) (Poehner & Lantolf, 2005; Vygotsky, 1998). DA is concerned with both the person's performance with assistance from someone else and also the extent to which the person can benefit from this assistance. This benefit is observed not only as the person completes the same task, but also as the person transfers the assisted performance to different tasks. The main difference between a dynamic assessment and that of others is whether or not the experience of

taking the assessment is intended to change performance during the experience itself.

Development, or learning, for Vygotsky was not specific to an individual test or task, but must include the person's ability to transfer what has been internalized through mediation beyond that of the individual test or task to other tests or tasks.

AFL, alternative assessment, and dynamic assessment are types of formative assessment. These types of formative assessments have been used in many areas and through many different means. There are different ways to use formative assessments in the learning process and many have shown increased learning outcomes.

Formative feedback is provided to students via formative assessments and there are different types of feedback that can be provided. This feedback can be provided through teachers, automation (computers), or from the student themselves. These different types of feedback have shown to assist the learning process by providing a basis for correcting mistakes or misconceptions ascertained through the learning process (Clariana, 1993; Cohen, 1985).

### **Self-Assessment as Formative Assessment**

One type of formative assessment involves self-assessment by individual learners. Self-assessments are not commonly used in schools, and most students view assessments as only a summative tool used by teachers to form judgment (Weeden & Winter, 1999). Self-assessments can be helpful in assisting the learner to gain insights to their own progress or lack of progress. Self-assessments can increase learning outcomes. Butler and Jiyoung (2010) found positive effects of self-assessments through quantitative analyses on 254 young learners of English as a foreign language. This study looked at sixth-grade students in South Korea who were asked to self-assess on a regular basis throughout a semester of English classes.

***Self-assessments.*** Self-assessment is an assessment of learner performance in which the learner plays an active role in evaluating and monitoring his or her own abilities. A self-assessment may be in the form of a summative or formative assessment, but the intent is one of which the learner interprets the outcome as opposed to the teacher doing the interpretation of the outcomes.

Fontana and Fernandes (1994) showed that self-assessment methods can result in greater learning gains. They showed this as they trained 25 Portuguese teachers of mathematics in self-assessment methods on a 20-week part-time course. These trained teachers put the self-assessment methods into practice as they taught 246 students of ages 8 and 9 and another 108 older students of ages between 10 and 14. There was also a control group of another 20 Portuguese teachers who were not trained in self-assessment methods. Both groups, experimental and control, were given pre- and posttests of mathematics achievement, spending the same amount of total time on mathematics. The experimental group's mean gain was a significant difference which was about twice the control group's mean gain.

McDonald and Boud (2003) showed that students trained in self-assessment techniques gained significant learning outcomes over a control group that was not trained. High school teachers were trained in how to develop a student's ability to self-assess their own work. Their study included 256 students in a treatment group that were trained, over an academic year, on self-assessment techniques in general curriculum subjects. This treatment group was matched with a control group that was not trained in the same techniques. Significant empirical differences were found with regards to learning outcomes in each curriculum area observed for those who were in the treatment group. This study demonstrated that self-assessment training can result in learning outcome gains.

Self-assessments enable learners to play a crucial role in their own learning process. Oscarson (1989) observed through the use of self-assessment in the classroom that students as well as teachers acknowledge assessment as a mutual responsibility, and not solely the responsibility of the teacher. Oscarson (1997) also observed many benefits to self-assessment including student involvement throughout the learning process, increased motivation, and the development of study skills that continued past the period in which they were used. Alderson and Banerjee (2001) found similar results from self-assessments and added the benefit of the learner's confidence in their own judgment, meaning that the learner felt more confidence in their own ability to understand their progress in the learning process.

***Self-assessments influence on self-efficacy.*** Bandura (1986, p. 391) defined self-efficacy as “people’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances.” Students with high self-efficacy have higher scholastic achievement, experience less anxiety, stick on their tasks longer, and search for deeper meaning in their learning tasks (Bandura, 1997; Cubukcu, 2008; Joet, Usher, and Bressoux, 2011; Multon, Brown, & Lent, 1991; Pajares, Miller, & Johnson, 1999; Pajares & Schunk, 2005; Yang, 2007).

Just as self-efficacy of students predicts a student's achievement at learning tasks, self-efficacy also corresponds to the manner in which students can implement self-regulated strategies in their learning situations (Joet, Usher, & Bressoux, 2011; Zimmerman, 2008). This is referred to as *self-efficacy for self-regulated learning* and self-regulatory beliefs have been shown to positively influence student achievement, motivation, and risk of dropout (Caprara et al., 2008; Usher & Pajares, 2008a, 2008b; Zimmerman, 2002). Bandura (1997) identified four primary sources of self-efficacy: (a) mastery experience, (b) vicarious experience, (c) social

persuasions, and (d) physiological and emotional states. According to Bandura's (1986, 1997) social cognitive theory, the most powerful source of self-efficacy comes from a student's interpretation of their own performance, or *mastery experience*. In this study, we will use self-assessments as a form of the mastery experience in which the student is able to assess their own progress. This assessment will enable the students to self-regulate and to monitor their own progress and understanding of their language speaking abilities (Eriksson Gustavsson & Samuelsson, 2011). Studies have shown success with this type of approach in disciplines other than second language speaking performance, and so this study will contribute to the application of these principles in the area of second language speaking performance.

The ability for a student to self-assess their own progress and understanding of a topic may both improve learning of content (Sadler, 1998) and enable a student to learn how to learn, which enables the student to become a lifelong learner (Boud, 2000). Although these benefits may result from self-assessment, much of the study has been done in comparing self-assessment results to that of teacher provided evaluations (Boud & Falchikov, 1989) in the same subject and not to the benefits of self-assessment training effects on courses taught by different teachers (McDonald & Boud, 2003).

***Self-monitoring.*** In addition for the learners to who need training on self-assessment, learners will need to learn how to self-monitor (Dickinson, 1987; Dickinson & Carver, 1980). Self-assessment can answer questions such as "How am I doing?" (Harris, 1997) and encourage learners to become part of the whole process of language learning. Self-assessment and self-monitoring enable a learner to make judgments on their own accuracy and on the appropriateness of their performance (Finch, 2001).



Miller and Ng (1996) and Harris (1997) showed that learners can be quite accurate in judging their own learning and that self-monitoring through self-assessment can be a practical tool for making students more active in their learning process and can assist the students in the daunting task of learning another language. Although students can be accurate in their self-assessments, students will need to know how to validly assess their behavior if they are not provided with native speech samples and criterion with which they can judge themselves.

### **Formative Assessments in the Second Language (L2) Classroom**

Ross (2005) used a mixed methods approach, which included self-assessment as a form of formative assessment, to study Japanese undergraduates ( $n = 2215$ ) who were enrolled in a two-year, sixteen course English for academic purposes program. Analyses indicated that formative assessments yielded higher language proficiency growth than those who were assessed by conventional summative assessments only. Ross's findings also revealed that although formative assessment can produce substantive increases in achievement and proficiency growth, this impact may be domain-dependent, e.g. language listening comprehension improvement.

In fact, the second language classroom can benefit from formative assessments just as in many other subjects. Geeslin (2003) highlighted benefits of self-assessment as applied to the quality of the language-learning environment, and discussed how self-assessment can act as an important mechanism through which learners interact with teachers in order to gain skills to evaluate their own process. Blanche (1988) found a general correspondence between self-assessed ratings and teacher evaluations of the same learners.

Recent contributions to the literature indicate that there is a transition from the conventional summative assessment approach of language learning outcomes to gradually integrating formative assessments, but that this is a continuing process (Davison, 2004). Also,

focus in recent literature is concerned less with reliability and validity and more on formative assessment uses and processes (Leung & Mohan, 2004; Moss, 1994; Rea-Dickins, 2001; Teasdale & Leung, 2000). Examples of this recent literature are Rea-Dickins and Gardner (2000) and then repeated by Rea-Dickins (2001) as they studied the uses and processes on formative assessments without reporting information on the reliability and validity of the studies reviewed.

### **Language Learning Challenges with Formative Assessment**

Although learning outcome gains have been seen in studies of formative assessments, there are many challenges that must be overcome. Teasdale and Leung (2000) addressed the epistemic and practical issues and challenges of formative assessments in the L2 classroom and suggest that the lack of standardization in this area provides a great challenge in moving forward the research. Another challenge is that of teaching students the ability to self-assess and set goals, which was one of the seven strategies of Assessment for Learning (Stiggins, et al., 2004).

Another key issue that is yet to be resolved is that of the possibility of weak reliability or internal consistency due to subjective observations involved with formative assessments (Brindley, 1994, 2000). If students are not trained well on rubrics or how to evaluate their own work, then outcomes may vary greatly amongst students.

Ross (2005) suggests that since formative processes are so dynamic, conventional experimental methods may be unlikely to detect effect in achievement and proficiency and those innovative methods will be needed in order to measure formative assessment effects. In an effort to evaluate the formative assessment effects in Ross's study, he used a mixed methods approach in an effort to provide innovative methods.

Clearly defined standards are necessary so that student, teacher, and formative assessment expectations are in alignment. Teachers need to know what concepts are most important and why they are important in order to contribute to the learning process in which the learner is engaged. In turn, this should provide purpose, assistance, and feedback (Atkin, et al., 2005). The assessment itself must also be aligned with the same standard in order to assist in the self-assessment process.

If students feel ownership and clearly understand these standards then they are more likely to stay focused on achieving those standards (Schmoker, 2001). Students will achieve goals that ‘stand still’ and that they clearly understand and see (Stiggins, 2005a, 2005b). Assessments should focus on specific outcomes (Brindley, 2001) in order to enable students to achieve those goals.

More work can still be done with regard to performance based self-assessments and how these might help in the learning process. Researchers are calling for empirical research on the impact of formative assessment in the area of second language acquisition and how it effects learner morale and achievement (Ross, 2005). This study reports on these effects through focus groups and feedback from the missionary trainees on their experience with their self-assessments in addition to the empirical data that is provided.

### **Moving Formative Assessment Research Forward in L2 Environments**

Self-assessments of foreign language proficiency have shown in limited situations to be helpful in the learning process for the student. Part of the challenge with self-assessment is that of how experienced the learner is with the process (Ross, 1998). Self-assessments of foreign language proficiency have also been shown in limited situations to be helpful to teachers in assisting the teacher to better instruct the students. More work still needs to be done in applying

these concepts to L2 performance in addition to acquisition of vocabulary and grammar skills. Ross (2005) identifies that there is little literature that addresses formative assessment and its effect on learner reflection and cooperative learning as well as its effects on language learner achievement and proficiency over time.

Some work has been done to evaluate the *usefulness* of formative assessments (Bachman & Palmer, 1996; Bachman, 2005) as defined and measured by Bachman's (2004, 2005, 2006a, 2006b, 2006c, 2007) Assessment Use Argument, but additional work still needs to be completed to evaluate this *usefulness* with regards to the second language classroom and learning (Colby-Kelly & Turner, 2007) which is an area of *usefulness* that has yet to be evaluated. Although *usefulness* was evaluated by this study, it was beyond the scope of the study to provide evidence as to whether learning had taken place. Work can and should move beyond that of *usefulness* and into evidences of the learning benefits.

Although studies have been done in the area of Assessment *for* Learning (AFL), much of the focus has been in the areas of mathematics and science (Harlen & Winter, 2004; Rea-Dickins & Gardner, 2000; Triggs et. al., 1999). Little work has also been done to apply these concepts to Second Language Acquisition (Colby-Kelly & Turner, 2007; Rea-Dickins, 2004). Recent focus has been given on the use of assessments in the area of second language acquisition as seen in special issues dedicated to this topic in the journals of *Language Testing* in 2001 and *Language Assessment Quarterly* in 2007.

Although this topic has been the focus of these special issues, much of the language testing community continues to make calls for more study into the area of applying AFL concepts to second language learning (Brookhart, 2005; McNamara, 2001a, 2001b; Poehner & Lantolf, 2005; Rea-Dickins, 2004; Shohamy, 2004). Rea-Dickins (2004) stated that "assessment,

with specific reference to teaching and learning in the language classroom, has remained, until recently, relatively unresearched” (p. 249). Ross (2005) also acknowledged that the long-term impact of formative assessments on language learning has not been examined empirically.

The effects for applying self-assessments and formative assessments to teaching and learning in the language classroom are an area of research need. This study uses the teaching of language criteria as a form to facilitate self-assessment and uses computer-based practice as a method of formative assessment in which the missionary trainees can evaluate their own performance while comparing them to the responses of others. Those who are taught language criteria are also able to apply those criteria in their self-assessment in a more directed way than those who were not explicitly taught language criteria. This study contributes empirical data on language speaking performance as a result of these two treatments.

### **Chapter 3: Method**

The reviewed studies have shown that (a) formative assessments can enhance learning outcomes in the classroom, (b) formative assessments have evolved into tools for learning, (c) self-assessment can be an effective form of formative assessment, (d) use of formative assessments in the second language classroom has been shown to also enhance learning outcomes, and (e) although some second language classroom research has been performed there is still much to do. While there is a need for more research to be done with regards to formative assessments and second language classrooms, there are challenges that need to be addressed as well.

For second-language assessments, Oscarson (1997) concluded that:

- Accuracy depends to a considerable degree on the purpose of the assessment.
- Assessments are more accurate when based on task content that is closely related to students' situations as potential users.
- Assessment is easier (and more accurate) when concerned with narrowly defined situations.

#### **Design**

The two independent variables that were studied included (a) computer-based practice, and (b) formal instruction about the criteria by which the missionaries' language performance was rated.

The two dependent variables in this design are (a) human-rated language speaking ability and (b) a computer-rated speaking assessment. These two dependent variables were compared in order to report on the reliability of the human raters used in this study.

The experimental design consisted of a 2 by 2 factorial design which allowed for the simultaneous testing of Hypotheses 1, 2, and 3. The 2 by 2 factorial analysis of variance using the General Linear Model procedure in SPSS permitted assessment of the potential interaction between the independent variables.

Intact classes of missionaries were the experimental unit used in this study. Four randomly selected classes were selected for 4 consecutive weeks from a total of 40 classes entering the MTC over a 4 week period. These 16 classes were randomly assigned to one of the four experimental conditions shown in Table 1. Therefore, the experimental unit consisted of classes rather than individual missionaries. There were 15-34 missionaries included in each experimental group.

Each experimental group was given a handout as shown in Appendix A in order to provide a description of the expectations for their assigned group. Two groups received opportunities for computer-based practice, and two groups received an opportunity for online training on the criteria by which their language performance was evaluated. An example of the language performance criteria rubric that was taught to the missionaries is included in Appendix B. All groups were administered a final test before their departure from the MTC.

Table 1

*Experimental Groups and Treatment Conditions*

Group	Computer-Based Practice	Taught Criteria
1	No	No
2	Yes	No
3	No	Yes
4	Yes	Yes

To account for how well the computer-based practice is implemented in this study, information for two covariates, (a) number of times the computer-based practice was used by each missionary and (b) amount of time spent in seconds reviewing feedback, were computer-recorded and used in the data analysis. Reviewing feedback for the purpose of this feedback is defined as time spent re-practicing and comparing self-samples to the audio samples from others that were provided.

In order to determine to what extent the four experimental groups were equivalent prior to exposure to the various treatment conditions, an elicited imitation pretest was administered to each group. Scores on this pretest were used as a covariate to account for any initial differences in the four groups.

The pretests and posttests measured different variables. Consequently, they were not treated as repeated measures of the same variable. A 2 x 2 repeated-measures design was considered but determined to not be an appropriate design due to the differing tests. The pretest was used as a means of checking to determine if the random assignment of classes to treatment conditions actually resulted in randomly equivalent groups prior to their exposure to the respective treatment. The two independent variables and the assignment of each treatment group are described in Table 2.

Table 2

*Description of 2 x 2 Factorial Design*

Practice Opportunity	Instructed on Evaluative Criteria	Not Instructed on Evaluative Criteria
No Practice Provided	Group 1	Group 3
Practice Opportunity Provided	Group 2	Group 4



## Participants

The participants in this study were 128 randomly selected missionaries assigned to learn Spanish as a second language. All of the missionaries were enrolled as trainees at the Missionary Training Center (MTC) in Provo, Utah. Each missionary began the study having never received prior training at the MTC.

As the missionaries were randomly selected, the corresponding language teachers were randomly assigned and so these teachers were identified and tracked as part of the analysis to be performed. The counts per treatment group are listed in Table 3. There were 3 participants that self-selected themselves out of the study during the orientation. No additional follow-up was performed to investigate why these missionary trainees self-selected themselves out of the study.

Table 3

*Number of Missionaries by Type of Instruction Provided*

Computer-based Practice Condition	Criteria Condition		Combined Groups
	No Criteria Taught	Criteria Taught	
No Practice Provided	33	34	67
Practice Provided	32	29	61
Combined Groups	65	63	128

Prior language experience was reported during a questionnaire administered within the first few days of each missionary's training. Table 4 shows responses indicating whether the missionary spoke the language before attending their training, whether they had lived in the assigned country, and whether any formal language study had occurred.

Table 4

*Prior Language Experience by Experimental Group*

Question	Group 1		Group 2		Group 3		Group 4	
	No	Yes	No	Yes	No	Yes	No	Yes
Did you study or speak your Mission Language before you came to the MTC?	12	28	10	15	20	18	9	22
Did you use your mission language as a common means of communication with your family or friends while you were growing up?	37	3	25	0	38	0	31	0
Have you lived in a foreign country where your mission language is spoken?	40	0	25	0	38	0	30	1
Have you done any formal study of your mission language (e.g. classes, self-study)?	15	25	11	14	25	13	8	22

**Instrumentation Overview**

A questionnaire was administered to each missionary before their training began so that information on prior language experience could be collected. A specimen copy of this instrument is shown in Appendix C.

The Language Speaking Assessment (LSA) used to assess speaking proficiency in the target language was previously developed by the Research and Evaluation Department at the MTC. This task-based instrument presents narrowly defined situations to which the student is expected to respond. Although the study focused on task-based items, the measure can be used to measure performance improvements both for that particular task and as an independent measure on general language ability.

Currently at the MTC, several formative assessments are used to assist the missionaries in learning a language. These assessments along with their purpose and availability are listed in Table 5. Each of these assessments is made available to missionaries as often as they would like to take them. The missionaries are given time each week during which they can take these assessments during their study time. These formative assessments were provided as an optional complement to the language practice treatment.

Table 5

*Formative Language Assessments*

Assessment	Description of Assessment
Language Speaking Assessment - Long	7-8 computer presented tasks during which missionaries are able to practice responses and compare themselves to others.
Language Speaking Assessment - Short	2-3 tasks to which the missionary verbally responds and records themselves to practice and compare themselves to others.
Formative Grammar Assessments	These are short assessments in which the missionary is presented with a task and context-based multiple choice, multiple select, and short answer responses.
Summative Grammar Assessment	These are more in-depth assessments in which the missionary is presented with 30-40 random items that seek to make a summative judgment on how well the missionary has learned the grammar.
Language Study Checklist	A self-assessment learning activity types that the missionary is doing, and intends to suggest additional activities that may help their learning.
Computer Assisted Rating Instruction	This is a computer-based training course on the criteria by which language performance will be rated. It provides examples that are rated by the student and then compared to previous rater scores. Upon completion of this instruction, the students receive a score on how well they learned the criteria.

The usage details for these formative assessments are listed in Table 6. The counts per assessment and by group are shown. Groups 1 and 3 did not take the Language Speaking Assessments as part of their treatment, and so the zero values are as expected.

Table 6

*Formative Language Assessment Usage*

Assessment	Assessment Counts			
	Group 1	Group 2	Group 3	Group 4
Language Speaking Assessment - Long	0	44	0	52
Language Speaking Assessment - Short	0	47	0	31
Formative Grammar Assessments	67	24	72	25
Summative Grammar Assessment	40	25	38	31
Language Study Checklist	62	52	50	32

***Instrumentation details.*** Each assessment was composed of several parts: (a) instructions, (b) questions, and (c) feedback. The instructions for the assessment consist of the purpose of the particular assessment, instructions on how to take the assessment, and possibly an example question for the assessment. An example of the instructions is displayed in Figure 1, which contains the instructions for the Summative Grammar Assessment. The instructions for the remaining assessments were similar, but adjusted to describe that particular assessment. The intention of providing instructions is to reduce the amount of variance due to missionary trainee computer experience and its influence on their assessment experience.

Assessment Tools

Current Assessments ▶ Grammar Test | Displaying of 50 |
Exit

Grammar Test

The Spanish Grammar Test is designed to measure a wide range of language ability. MTC teachers should be able to answer all or almost all of the items on the test. Missionaries - unless they have had a lot of previous experience in Chinese - will generally know only part of the items. Consequently, do not be discouraged if you don't know everything. Just do the best you can. On this test you will encounter some common missionary situations. Several multiple-choice items are associated with each situation. For each item, please select the alternative that best expresses the meaning of the underlined word(s) in English.

Example:

**In this situation, Elder Jones is teaching María, an investigator, about God.**

		(0)
María:	Quien es Dios? Who is God?	<input type="radio"/> nosotros <input type="radio"/> nuestro <input type="radio"/> nos <input type="radio"/> somos
Elder Jones:	Dios es <u>(0)</u> Padre Celestial. God is <u>our</u> Heavenly Father.	

For blank Number 0, you need a Spanish word that means our, as in God our Father. Looking at the options listed at the right for Number 0, you would pick 'nuestro'.

**If you don't know the answer to a particular item, do not hesitate to mark the alternative that says, "I don't know the answer to this question." You will not be shown an overall score on the test, so checking this box will not affect the results you receive. After completing the grammar test, you will receive feedback on all of the items missed. The correct answer, as well as an explanation of why the answer is correct, will be shown to you. Please review the feedback carefully so that you can learn from the test. You may print the feedback page if you wish.**




Figure 1. Example instructions for Language Grammar Assessment.

**Instrumentation example.** An example of how context-based items are displayed is shown in Figure 2. After completion of an assessment, the missionary is presented with feedback on their response and specific references on how to arrive at the correct response. This feedback is shown in Figure 3. Each question also includes the response of “I do not know the answer to this question” since all items require a response before a student can move on through the assessment.

In this situation, Elders Davis and Hart are becoming acquainted with Mateo and Silvia, a member couple.

<p>Mateo: <u>(1)</u>, élderes. Acabo de llegar del trabajo.  <u>Come in</u>, elders, I just got home from work</p>	<p>(1)</p> <p><input type="radio"/> 1. Entran</p> <p><input type="radio"/> 2. Entren</p> <p><input type="radio"/> 3. Entre</p> <p><input type="radio"/> 4. Entramos</p>
<p><input type="radio"/> I do not know the answer to this question.</p>	
<p>Elder Davis: Gracias. Con permiso.          Thanks. With permission.</p> <p>Elder Hart: ¿Dónde <u>(2)</u> ?          Where <u>do you work</u> ?</p>	<p>(2)</p> <p><input type="radio"/> 1. trabaja Ud.</p> <p><input type="radio"/> 2. hace Ud. trabajo</p> <p><input type="radio"/> 3. Ud. trabajo</p> <p><input type="radio"/> 4. hace Ud. trabajar</p>
<p><input type="radio"/> I do not know the answer to this question.</p>	

Figure 2. Example of questions that are context dependent.


<p> Elder Davis: Gracias. Con permiso.          Thanks. With permission.</p> <p>Elder Hart: ¿Dónde <u>(2)</u> ?          Where <u>do you work</u> ?</p>	<p>(2)</p> <p>1. trabaja Ud.</p> <p>2. hace Ud. trabajo</p> <p>3. Ud. trabajo</p> <p>4. hace Ud. trabajar</p>
<p><b>Correct response:</b> trabaja Ud.</p> <p><b>Your response:</b> trabaja Ud.</p> <p>The present tense is used to describe actions that are currently taking place. The word "you" translates to Usted and requires an -a ending for present tense -ar verbs. The words 'do,' and 'does' used in English questions are not translated into Spanish. For example, "Where does he eat?" would be translated as, ¿Dónde come él? Although the verb hacer means to make or do, it is not used as a helping verb in Spanish.</p>	
<p>Mateo: Trabajo en la ciudad. <u>(3)</u> carros usados.          I work in the city. <u>I sell</u> used cars.</p>	<p>(3)</p> <p>1. Vendio</p> <p>2. Vendió</p> <p>3. Vendo</p> <p>4. Vendó</p>
<p><b>Correct response:</b> Vendo</p> <p><b>Your response:</b> Vendió</p> <p>The present tense is used to describe actions that are currently ongoing. To form the Yo form (first person present tense) drop the -er, -ir, or -ar ending of the infinitive verb. Then add -o.</p>	

Figure 3. Example feedback with description of correct response.

**Procedure**

All assessments used in this study were delivered as web-based assessments. Missionaries from each class took the assessments in onsite labs that contain 12 computer workstations each. All missionaries from a given class were able to take the assessments in the same lab room. The times were scheduled by the missionaries themselves. How long the missionaries spent in the lab taking the assessments was determined by each missionary. The assessments were voluntary, and so not only was the event of taking an assessment dependent on the missionary, but also the number of times that each missionary took each assessment was also up to the missionary themselves.

**Analysis**

The data were gathered by the assessments that are delivered through web-based tools. The web-based tool was a computer program developed specifically for this purpose, and data was collected on amount of time spent by each missionary in the assessment as well as the responses provided by each missionary. The amount of elapsed time spent by each missionary using each assessment was especially helpful in determining how much time was spent in the Language Speaking Assessment listening to the missionary's own response as compared to other responses. We could also determine from the data whether the missionary re-practiced their responses.

The Many-Facets Rasch Model was used to assess the reliability of the ratings and to adjust for any observed rater effects. The design was used with sufficient linkage to have connected subsets. The Facets software was used to implement the Many-Facets Rasch Model. The Facets software generated an adjusted score for each missionary. These adjusted scores, reducing rater effects, were used as the input to a factorial analysis of covariance.

A two-way factorial analysis of covariance was used to assess each of the main effects plus the interaction. This two-way factorial analysis of covariance allowed for the testing of each of the three hypotheses.

If the  $F$ -test for the interaction was statistically significant, then the means of the four treatment groups would be plotted in order to describe the nature of the observed interaction. A separate graph would be constructed for each dependent variable. Mean scores on the dependent variable were to be plotted on the y-axis. The two levels of the computer-based practice variable would be represented on the x-axis, and the two levels of the other independent variable would be represented by lines in the graph. If the  $F$ -test for the interaction is not statistically significant, then the graphs would not be provided.

If an interaction effect was reported, a PostHoc test would be run in order to better determine the causes of the interaction effect. The means of Group 4 would be compared to the combined means of Groups 1-3.

Because classrooms were used as the experimental unit, Hierarchical Linear Modeling (HLM) was used to analyze data to account for nesting within higher levels of analysis. When units are grouped at higher units of analysis, e.g. companions, classrooms, etc., such data are considered to be nested. This nesting can occur between subjects (students within classrooms, companionships within classrooms, etc.) and/or within subjects (repeated observations on the same individuals over time.)

Regular regression analysis on nested data can increase Type I error resulting in model misspecification, and miss opportunities to examine potentially interesting contextual questions. HLM analysis was performed in addition to the 2 by 2 factorial analysis in order to examine



these contextual questions and detect whether nesting within the higher levels of analysis provides for any insights on resulting scores.

HLM was implemented through SPSS mixed models. Hierarchical models are those in which data collected at different levels of analysis (e.g. class, wave, occurrence) may be studied without violating assumptions of independence in linear multiple regression. For example, the fact that students respond together and have the same exposure within a classroom or wave means that responses from students within each classroom or wave are likely not independent from one another. Multilevel modeling accounts for these dependencies by estimating variance associated with group differences in average response and group differences in associations between predictors. Declaring intercepts and/or slopes to be random effects accomplishes this.

## Chapter 4: Results

Due to the nature of the two between-subjects variables and the one within-subject variable, which this study used, the following hypotheses were tested. This chapter presents the results of the tests performed to gather information on the following hypotheses:

1. Missionary trainees who practiced their language performance with a computer, comparing their performance to native and proficient speakers will obtain higher posttest scores on average than missionaries who did not have the opportunity for computer-based practice.
2. Missionary trainees who were taught and understand the criteria by which their language performance was to be rated will obtain higher posttest scores on average than missionaries who are not taught the criteria by which their language performance was rated.
3. Missionary trainees who both practiced their language performance with a computer and were taught the criteria by which their language performance was rated will obtain higher posttest scores on average than missionaries who experienced only one or neither of these treatments.

### Assessments of the Various Components

The pretest was administered to all missionary trainees within the first 2-3 days of arriving at the Provo MTC in order to measure their language speaking ability before language training begun. Descriptive statistics for each treatment group on the pretest are provided in Table 7. Descriptive statistics for the combined groups for the criteria condition and computer-based practice condition are also provided.

The assumption is that the treatment groups were statistically equivalent at the beginning of the study because the initial differences between all treatment group means were not statistically significant,  $F(3, 124) = 2.429, p = 0.069$ . The initial differences between the means

of the two practice conditions were also not statistically significant,  $F(1,124) = 1.948, p = 0.165$ .

The initial differences between the means of the two criteria conditions were not statistically significant,  $F(1,124) = 2.591, p = 0.110$ .

Table 7

*Descriptive Statistics for the Four Groups on the Pretest*

Computer-based Practice Condition	Criteria Condition								
	No Criteria Taught			Criteria Taught			Combined Groups		
	<i>n</i>	M	SD	<i>n</i>	M	SD	<i>n</i>	M	SD
No Practice Provided	33	23.88	8.70	34	28.74	7.86	67	26.34	8.58
Practice Provided	32	24.13	10.06	29	24.21	7.81	61	24.16	8.99
Combined Groups	65	24.00	9.32	63	26.65	8.10	128	25.30	8.81

An additional assumption is that there is no class clustering effect at the time of the pretest. The pretest was administered within the first few days of training, and thus there was not sufficient time for class assignment to have a clustering effect. Clustering effect was included as part of the posttest analysis that was performed.

The pretests were an elicited imitation test rated by the computer while the posttests were verbal responses that were human-rated and so it is noted that they were different tests, which rated different variables. The pretest and posttest were also on different scales and so the results were transformed to match the pretest scale.

Posttests revealed statistically significant effects to support the second hypothesis, but the posttests did not reveal significant main and interaction effects on the dependent variables that would support the first and third hypotheses. The descriptive statistics for the posttest are provided in Table 8.

The effect of whether or not criteria were taught to the missionaries was statistically significant,  $F(1, 121) = 4.820, p < 0.030$ . The effect of computer-based practice was not statistically significant,  $F(1, 121) = 0.103, p < 0.748$ . The interaction effect of computer-based training and the teaching of criteria was also not statistically significant,  $F(3, 121) = 1.701, p = 0.170$ . Since no interaction effect was found, a PostHoc analysis was not required to further investigate the source of this effect.

Table 8

*Descriptive Statistics for the Four Groups on the Posttest*

	Criteria Taught Condition								
	No Criteria Taught			Criteria Taught			Combined Groups		
Computer-based Practice Condition	<i>n</i>	M	SD	<i>n</i>	M	SD	<i>n</i>	M	SD
No Computer-based Practice Provided	33	46.86	7.78	34	49.17	6.58	67	48.05	7.23
Computer-based Practice Provided	32	46.45	9.26	29	50.52	8.72	61	48.38	9.16
Combined Groups	65	46.66	8.47	63	49.78	7.59	128	48.21	8.17

Note however that the variance of the scores for the computer-based practice group ( $9.16^2 = 83.9$ ) is more than 50% larger than the variance of the no practice groups ( $7.23^2 = 52.3$ ). The range for the computer-based practice groups (41.9) is also greater than the range for the no practice groups (37.2). This difference is not statistically significant, but it may be an indication that the computer-based practice had a positive effect on some missionaries, but not on others. This potential differential effect would not be detected by simply comparing the group means. The variance of the two groups on the pretest was much more similar.

The amount of time spent by the missionaries in Groups 2 and 4 practicing their language performance on a computer was recorded and is reported in Table 9. All of the participants (32 out of 32) in Group 2 participated in the computer-based practice. All but one of the participants (28 out of 29) in Group 4 participated in the computer-based practice. Group 2 missionary trainees repeated the practice from 1-5 times each with the distribution (a) 3 trainees practicing 1 time, (b) 8 trainees practicing 2 times, (c) 5 trainees practicing 3 times, (d) 8 trainees practicing 4 times, and (e) 8 trainees practicing 5 times. Group 4 missionary trainees repeated the practice from 3-5 times each with the distribution (a) 2 trainees practicing 3 times, (b) 25 trainees practicing 4 times, and (c) 1 trainee practicing 5 times.

Table 9

*Descriptive Statistics for Minutes Spent in Practice for Groups 2 and 4*

Computer-based Practice Condition	Criteria Taught Condition								
	No Criteria Taught			Criteria Taught			Combined Groups		
	<i>n</i>	M	SD	<i>n</i>	M	SD	<i>n</i>	M	SD
Computer-based Practice Provided	106	12.89	6.70	111	17.54	7.82	217	15.27	7.64

### Multivariate Analysis of the Relationship Between Treatments and Language Performance

The ANOVA calculations above do not account for the pretest as a covariate nor account for random class effect, but the results do show a statistically significant increase in posttest scores when language performance criteria were taught. Hierarchical Linear Modeling (HLM) was used in order to determine the relationship between the treatments and language performance after controlling for the variables class / teacher and companionship.

HLM takes into account the fact that there are correlated error terms between missionary trainees who have the same teacher. If a teacher is extraordinarily good at language instruction, then this would positively affect the scores for his or her students. Given that there is a potential correlation between the missionaries and these factors, HLM is needed. Results of HLM are interpreted in a manner to correctly take into account the fact that correlated error terms may exist among groups of missionary trainees.

In order to account for group differences at the time of the pretest, HLM was used with pretest as a covariate with no random class effects. The pretest effect was modest, but resulted in the teaching of criteria no longer being statistically significant. The HLM results of using pretest as a covariate are displayed in Table 10.

Table 10

*Results of ANCOVA on the Posttest with a Pretest Covariate*

Parameter	Estimate	Std. Error	df	<i>t</i>	<i>p</i>
Practice	-2.106	2.067	120	-1.019	0.310
Criteria	-4.007	2.081	120	-1.926	0.056
Practice*Criteria	2.563	2.886	120	0.888	0.376
Pretest	0.178	0.083	120	2.146	0.034

The pretest effect was modest, but is statistically significant ( $t = 2.146, p < 0.034$ ). This is not surprising due to the small sample size. The study has relatively low power to detect effects. The study also has relatively low power to detect effects particularly when testing the random class effect using pretest as a covariate. The standard error also increases when we add the random class effect into the HLM calculations as shown in Table 11.

Table 11

*Results of HLM on the Posttest with a Pretest Covariate and Random Class Effect*

Parameter	Estimate	Std. Error	df	<i>t</i>	<i>p</i>
Practice	-1.936	3.735	12.209	-0.518	0.613
Criteria	-3.883	3.722	12.119	-1.043	0.317
Practice*Criteria	2.607	5.259	12.034	0.496	0.629
Pretest	0.206	0.074	111.747	2.775	0.006

Accounting for random effect resulted in only a 4-point difference in posttest scores among the treatment groups. Adding the covariate and random class effect makes the pretest a statistically significant factor, but does not dramatically affect the outcome nor reduces the error. The conditional intra-class correlation (ICC) based on the variance estimates was 0.302 while the unconditional intra-class correlation was 0.247. The ICC calculations show that about one-third of the posttest variance comes from the clustering effect of the class.

Removing the covariate of the pretest from the calculations and only accounting for the random effect also did not show any significant changes in the outcomes. The results of these HLM calculations are below in Table 12.

Table 12

*Results of HLM on the Posttest with Random Effect and no Covariate*

Parameter	Estimate	Std. Error	df	<i>t</i>	<i>p</i>
Practice	-1.038	3.648	11.944	-0.285	0.781
Criteria	-3.841	3.650	12.058	-1.052	0.313
Practice*Criteria	1.605	5.142	11.839	0.312	0.760

### **Missionary Attitudes Towards Language Learning**

Only 84% of the missionaries responded to the *Exit Questionnaire*. The distribution of responses by treatment group is indicated in Table 13. These responses provided valuable feedback in terms of the trainees' attitudes towards language assessments, confidence in their ability to speak in the target language, and general feedback regarding their involvement with the study.

Table 13

*Missionary Exit Questionnaire Responses*

Treatment Group	Count
1	28
2	15
3	28
4	34



Specific items from an *Exit Questionnaire* are listed in Table 14. The full set of items administered in this questionnaire is shown in Appendix D. The responses to items particular to this study were averaged by group to illustrate differences in responses between the treatment groups. The missionaries' responses to each item indicated the degree to which they agreed or disagreed with the statements in the left column. The responses were coded on a scale from 1 to 5. The standard deviation is also provided in order to indicate the deviation for each treatment group.

Table 14

*Descriptive Statistics for Attitudinal Scores by Group by Item*

Item	Statistics	Group 1	Group 2	Group 3	Group 4
I feel confident in my language speaking abilities for my mission.	Mean	2.14	2.60	2.68	2.54
	St. Dev.	1.08	0.74	0.88	0.74
I feel confident in evaluating excellent and poor language performance.	Mean	2.07	2.67	2.82	2.86
	St. Dev.	0.77	0.62	0.72	0.71
I understand criteria that determine high language performance.	Mean	2.46	2.80	3.24	3.11
	St. Dev.	1.00	0.41	0.55	0.69
Taking language assessments motivate me to work harder.	Mean	2.54	2.53	2.24	2.75
	St. Dev.	1.04	0.99	0.43	0.52
Taking language assessments makes me feel frustrated or discouraged.	Mean	2.18	1.40	1.85	1.75
	St. Dev.	1.19	0.63	0.50	0.89
How helpful was the Language Speaking Assessment for you personally?	Mean	2.40	2.40	2.26	2.71
	St. Dev.	0.92	0.63	0.75	0.81
How helpful was the Computer Assisted Rater Instruction (CARI) for you personally?	Mean	n/a*	n/a*	2.64	2.68
	St. Dev.			1.01	0.86

\*Not applicable, the treatment group did not receive this treatment and so no responses were recorded

***Overall language learning experience.*** The missionaries were asked “How would you describe your overall language learning experience at the MTC?” The responses in Table 15 are characteristic responses for each treatment.

Table 15

*Overall Language Learning Experience — Missionary Responses*

Group	Characteristic Responses
1	<p>I feel like I’ve learned a lot since I have been here but I know that I still have a really long way to go.</p> <p>Alright. It was very fast paced and I learned a lot. I’m not very good at learning language stuff, but it was good.</p> <p>It was ok, I think we need to be immersed in the language more.</p>
2	<p>It was wondertastic. I just wish that I had worked harder.</p> <p>I loved the learning experience.</p> <p>Impressively fast to learn. I don’t know everything but I would survive in Mexico.</p>
3	<p>I learned so much, I can speak Spanish decent and have the ability to carry out my purpose as a missionary.</p> <p>It was very good. I learned and used a lot of Spanish. Definitely much better than high school.</p> <p>It was very good, I learned more Spanish than I thought was possible here.</p>
4	<p>Great I’ve learned so much. I never thought that I could get this good this fast.</p> <p>I feel that my district as well as I have been given a great opportunity and a higher level of fluency than many. Though I of myself can say there is always room for improvement, and I am not fluent yet, I am well on my way to fluency very few months into my mission. The teachers, teaching, and CARI-LSA combined with individual and group efforts have strengthened the language skills far beyond expected. This program has given great success in language.</p>

***Computer assisted rater instruction experience.*** The missionaries were asked “How helpful was CARI in learning the language?” The responses below in Table 16 are characteristic responses for each treatment. This question was only asked to the two groups who experienced the CARI training.

The overall feedback on the missionary’s experience with CARI was that it was helpful and provided insights that assisted in their language learning experience. There were a small handful of missionaries who felt that were already aware of the language criteria, but for the most part those who participated benefited from the training.

Table 16

*CARI Experience — Missionary Responses*

Group	Characteristic Responses
3	<p>I felt that CARI helped me improve my fluency. I noticed it was irritating when nonnative English speakers had really long pauses. It also made me want to pronounce words better. I feel that CARI made me improve overall in the various categories because I try and think about how I would do if I was being tested.</p> <p>I enjoyed it because it helped humble me. I know I can speak better than the people that are trying to learn the language. However, I'm the exact same way if it were me trying to speak Spanish. That was most helpful to me. However, I enjoyed using CARI because its cool to be able to see where you're at in the Spanish language. There isn't really much more than that, but I am grateful for being able to take part in using CARI.</p> <p>CARI helped me realize the places that I needed to work on and that I was actually better than I thought.</p> <p>I personally didn't like the whole thing. But I don't like TALL either so that tells a lot about what I liked with the computer stuff.</p> <p>I felt that it helped me recognize what I would need to be working on. I also showed me a lot of areas that I could use work in. I showed how to tell if it was good or bad language which reflects back to how I may sound.</p>

(Table 16 continues)

(Table 16 continued)

Group	Characteristic Responses
4	<p>I didn't really feel like it related very much to learning the language. I didn't understand the connection. I feel like before I took the CARI assessments I already had a pretty solid understanding of how to rate the fluency of a language. I would suggest trying to have people begin rating without reading through all of the instructions and just telling them to read the descriptions while they are listening as they hover their mouse over the criteria.</p> <p>I got to hear others speak the language and that helps me to understand and learn how I should be saying words better. It also helped me to see some of the things I should not do.</p> <p>I felt like it was a positive impact, though a small one. It could help those being rated though. I didn't receive any feedback from my LSA tests, but I could it being very helpful to take the LSA and see your results from the CARI evaluation</p> <p>I feel like CARI was greatly helpful in my language learning ability. By rating others I was able to rate myself and others and see where improvements were and where they were needed. It also helps to know what in rating grades your level of fluency.</p> <p>By hearing people who are completely fluent, I am more motivated to be able to answer questions like they can. It's nice to see how fluent we should sound every once in a while because in the MTC, we're surrounded by broken spanish so it's easy to feel comfortable when you're in that type of environment but I think we all need to be reminded of what we really need to be working towards</p> <p>I loved the CARI. It was nice to see it broken up into the different parts. That language is not only about knowing the words. There are so many other parts to it. With the pronunciation, fluency, vocabulary... A person can be good at one part, but there is room for improvement and when you are ranked high in all parts, you can be fully understood by all people.</p>

*Effect of CARI on trainees' goals and study patterns.* The missionaries were asked "Did CARI change your goals and how you study?" The responses below in Table 17 are characteristic responses for each treatment. This question was only asked to the two groups who experienced the CARI training.

Although there were some negative responses pointing out that the treatment could be enhanced by connecting the treatment to their learning experience, most of the missionaries involved with the CARI training felt that the training enhanced their learning experience by emphasizing areas in which they could personally improve.

Table 17

*Missionary Responses on the Effect of CARI on Goals and Study*

Group	Characteristic Responses
3	<p>It made me want to work harder on things like speaking better and made me want to work on improving my language skills to sound more like 7 rather than a 3.</p> <p>It just makes me want to learn faster and learn more in a less period of time.</p> <p>Not really, I thought it was somewhat of a waste of time and eachtime I did do it I dreaded doing it. But I can see how it was helpful</p> <p>Not really, however it helped me with my fluidity and how to properly pronounce words. It showed me that I really need to improve. I like the concept and idea of using CARI, however I don't really now how it has helped me change my goals and how I study.</p> <p>I started to read aloud in the Book of Mormon to work on my pronunciation of words. I also have paid more attention recently to the structure of my sentences.</p> <p>Not really just because I didn't get to experience it enough. I think if I had the option to use it more often and it was a little more interactive it could change my study.</p> <p>Yes, I wanted to improve my fluency and accent after using CARI, so I didn't change my language study plan too much, I would just spend more time on reading aloud or something like that to help my accent and fluidity in speaking spanish. Besides that, I didn't change too much about my studies.</p> <p>CARI made me want to try harder to increase fluency and vocabulary. It is easier to address areas that are lacking when you can rate where you are at.</p>

(Table 17 continues)

(Table 17 continued)

Group	Characteristic Responses
4	<p>CARI changed my study through trying to speak more and goals I just knew by the grading what areas I needed to work on. So in stride, in that way combining it with LSA helped too.</p> <p>It made me want to work on my pronouncitaion and accent to sound more native but thats about it. Though it is a fairly short program and wouldn't be too hard to implement and I think it would be effective tool to use to press the importance of pronunciation.</p> <p>No. I think if I had understood the ""why"" behind CARI it would have been more helpful.</p> <p>Yes it did, because I am now looking at how I pronounce things (more than before) and I am seeing how focusing on the specific areas can really help me find balance in my studies. Meaning that I focus on more than just learning the word but also the other areas of making it sounds right, along with being able to explain it and finding ways to speak it right.</p> <p>It did a little, in that I thought of new techniques to implement in my study such as memorizing scripture, or figuring out how to share a personal experience.</p> <p>YES!!! because i need to realize i will never be a fluent speaker with out working for it. I enjoy my studying more now.</p>

***Language speaking assessment experience.*** The missionaries were asked “In what ways was the Language Speaking Assessment helpful to you?” The responses in Table 18 are characteristic responses for each treatment. This question was only asked to the two groups who experienced the Language Speaking Assessment (LSA) as part of their training.

From the feedback, it appears that the Language Speaking Assessment (LSA) was helpful to the missionaries. There was very little negative feedback with regards to their LSA experience, and many of the missionaries reported that it helped them improve in both study and performance.

Table 18

*Language Speaking Assessment Experience — Missionary Responses*

Group	Characteristic Responses
2	<p data-bbox="370 443 1409 583">It allowed me to see areas where I need to improve and work on for my language study plan. It allowed me to think about how I would respond to situations that I could encounter in the mission field and be efficient in my answer so I didn't take forever</p> <p data-bbox="370 625 1110 659">here accents, and know how to say certain things correctly</p> <p data-bbox="370 701 1016 735">It helps me see where I'm at in my language study.</p> <p data-bbox="370 777 883 810">it helped me realize i forgot future tense</p> <p data-bbox="370 852 1084 886">It helped me to evaluate where my language skills were.</p> <p data-bbox="370 928 1409 1024">It helped me to want to be more focused on trying to understand what people are saying when they speak the language and it prepared me to what I will be hearing when I am out in the field.</p> <p data-bbox="370 1066 1425 1129">it helped me because i was able to have a topic by surprise and i had to just talk and say what they want me to say</p>
4	<p data-bbox="370 1213 1425 1354">At first I was really frustrated with it because I didn't know very much. But then the second time I did it, I saw a lot of improvement. This was very encouraging for me and I wanted to push myself to do even better the the next time. I was never able to but I know that I wish I would have done more sooner.</p> <p data-bbox="370 1396 1425 1537">It was an eye opener to see what I think I sound like, and how that will sound to the natives in my mission. I have a better idea of what to work on in my language study to make sure my language speaking skills won't make people uncomfortable or distracted when I teach</p> <p data-bbox="370 1579 1409 1682">It helped me to see what to look for in someone that sounds and speaks well. and helped me to see where i am at and in what areas i need to work on the most and what needed the most work.</p>

(Table 18 continues)

(Table 18 continued)

Group	Characteristic Responses
	<p>The Language Assessment was helpful to me to understand how much I needed to improve. Not only how much, but also what I should do to improve. There are aspects of language speaking that I did not think were important for the understanding of the listener when communicating. I have learned how much I need to improve my fluency, vocabulary, and grammar and not so much my pronunciation.</p>
	<p>It helped me to track my progress and give me encouragement when i could see how far I have progressed. It also let me know how fast I am progressing and if I start to slack off I can see that my language isn't improving as well as it should be.</p>
	<p>It was nice to be able to compare my language skills to that of a natives. It was nice to afterwards go back and listend to all the speakers in that section to know where I fell. And to try and understand their accent.</p>



## Chapter 5: Discussion

As discussed in Chapter 2, studies have shown that formative assessments (Ross, 2005) can be an effective tool for increasing second language learning outcomes for students.

Although this study resulted in observable gains where students spent time practicing with the computers and receiving formative feedback, there were no statistically significant interaction results between computer-based practice and receiving instruction on the criteria by which the students would be rated.

In Chapter 2, it was also discussed that students must understand the criteria by which they will be rated in order to stay engaged and focused on the proper outcomes (Schmoker, 2001). The results from this study showed that those who were familiar with the criteria did in fact score statistically significantly higher on the posttest.

### Research Hypotheses

Although data collected provided statistically significant support for only the second hypothesis, there was helpful data collected to shed light on benefits from focusing on the first and third hypotheses. Missionary trainee responses to the questionnaire provided positive feedback on the use of computer-based practice during their learning process and in the opportunity to compare their responses to the responses of others.

*Language performance practice.* The first hypothesis asserted that practicing language performance with a computer would increase posttest language scores. The results from this study did not support this hypothesis. The overall attitude of missionaries tends to be positive with regard to language learning in general, but most missionaries who were given the opportunity to practice their language performance with a computer stated that the experience was helpful in their learning process. This computer-based practice opportunity influenced their

study and goal setting as well, and the missionaries stated that the opportunity to listen to other native and fluent speakers assisted in the language learning process.

***Understanding language performance criteria.*** The second hypothesis asserted that understanding language performance criteria would assist the missionaries in their language learning process and result in higher posttest scores. The posttest scores for these groups were statistically higher on the average than the scores of those who were not taught the language performance criteria. Thus, this hypothesis was supported.

Missionaries not only scored higher on their posttest on the average, but the comments in the exit questionnaire also showed that they were more confident than those who were in the treatment group. Missionaries reported that going through the computer-assisted rater instruction and learning the criteria by which language performance is rated allowed them to better identify areas in which they needed to personally improve. Missionary comments also indicated that they adjusted their study, practice, and interactions with classmates to respond to the areas in which they needed to improve.

***Interaction between language performance practice and understanding criteria.*** The third hypothesis asserts that there would be an interaction effect between language performance practice on the computer and understanding the criteria by which language performance is rated. No statistically significant results were found to support the hypothesis.

Overall confidence in the missionary trainees' language learning and performance abilities were increased over the treatment group, but there was not an increase over the results found without the interaction of the two treatments.

## **Recommendations**

A follow-up study should be performed to gather longitudinal data in order to investigate whether language performance for each treatment group has a long-term effect. This follow-up study should focus on whether language performance continues to increase among the treatment groups. This follow-up study should also focus on whether those missionaries who were taught the criteria by which language performance is rated continue to apply the principles they were taught in order to continue their learning.

If evidence is obtained in future studies showing that the results of participation are beneficial, then a cost-benefit analysis should be performed in order to determine whether providing computer-based language practice opportunities can help reduce the face-to-face time required by teachers with each missionary. If computer-based training can help reduce face-to-face requirements from the teachers, then teacher time may be better utilized for those missionaries who need the additional help as opposed to working with the entire class. It may be difficult to identify methods to determine the cost-to-benefit ratio in order to determine if the effort is worth the costs, but such a study will help future language instructional designers identify whether they would like to further develop computer-based language performance practice tools.

Future studies should be adjusted to investigate heterogeneity of treatment fidelity across groups within experimental conditions. Accounting for treatment fidelity across groups may help to clarify or further distinguish difference in posttest scores across the treatment groups.

The results of this study showed that it was underpowered. The study should be repeated with a (a) at least 32 classes of missionary trainees, (b) increase in sensitivity of the measure, and (c) lengthened and required time that missionary trainees spend using the treatment. The

repeated study should be conducted either during the summer months when the number of trainees is higher than during this study or for an extended period of time spanning at least one year.

The method of training missionary trainees on the language criteria should be re-evaluated and much improvement can be made on this approach. Some of the missionaries commented on not understanding how the criteria connected to their language experience. This is due to the fact that the training used was for raters of language, and the learning strategy was not designed for a missionary trainee. This criteria instruction should be redesigned and targeted on the missionary trainee and how they are expected to use the criteria throughout their learning experience.

The computer-based practice should also be better integrated into the curriculum of the missionary trainee. With this study, the use of the computer-based practice was voluntary and follow-up was not performed to enforce use. Future studies should require the use of the computer-based practice and integrate its use with the classroom curriculum.

## **Conclusion**

The purpose of this study was to examine the benefits of providing computer-based language performance practice, teaching the criteria by which language performance is rated, and the interaction between the two treatments. Although statistically significant quantitative data was only found to support the teaching of criteria by which language performance is rated, missionary trainee responses were found that support the further use and exploration of these treatments. The findings are preliminary but promising in showing a favorable increase in language performance outcomes when being taught language performance criteria and given the opportunity to practice language performance on a computer.

It is assumed that the measurements used did not identify all of the benefits to the missionaries for each of these treatments, and thus there may be benefits that encourage the continued use of both of these treatments. Qualitative feedback was strong enough that further research in these areas would be encouraged.

For MTC specific goals and implementation, the MTC administration will continue its use of the Language Speaking Assessment and in fact is evaluating whether this tool should continue with the missionaries beyond their formal language training experience. Missionary responses and encouragement for the use of the Language Speaking Assessment show that there might be value in continuing its use as the missionaries continue their language learning on their own.

MTC administrators may also be encouraged to evaluate whether the CARI instruction should become part of the missionary training process. The current curriculum is already more than most missionaries can understand within the limited training period at the Missionary Training Center, and thus it would be recommended that a shorter and more interactive training be developed to teach these principles. This training would also be helpful to the missionaries beyond the formal training period, and may be provided as a refresher course as the missionaries begin to manage the language learning on their own.

## References

- Alderson, J., & Banerjee, J. (2001). Language testing and assessment (Part i). *Language Testing*, 34, 213-236.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261-271.
- Atkin, J., Black, P., & Coffey, J. (2001). *The Relationship Between Formative and Summative Assessment—In the Classroom and Beyond*. Retrieved on April 12, 2010, from Classroom Assessment and the National Science Education Standards: [http://books.nap.edu/html/classroom\\_assessment/index.html](http://books.nap.edu/html/classroom_assessment/index.html).
- Atkin, J., Black, P., & Coffey, J. (Eds.) (2005). *Classroom Assessment and the National Science Education Standards*. Committee on Classroom Assessment and the National Science Education Standards, Center for Education, National Research Council. Washington, DC: National Academy Press.
- Assessment Reform Group. (2007). *Assessment for learning*. Retrieved on March 11, 2010, from <http://www.assessment-reform-group.org/ASF.html>.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge, England: Cambridge University Press.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2, 1-34.
- Bachman, L. F. (2006a). Generalizability: A journey into the nature of empirical research in applied linguistics. In M. Chalhoub-Deville, C. A. Chapelle, & P. Duff (Eds.), *Inference and generalizability in applied linguistics: Multiple perspectives* (pp. 165-207). Amsterdam: Benjamins.

- Bachman, L. F. (2006b, April). *Linking interpretation and use in educational assessments*. Paper presented at the National Council for Measurement in Education, San Francisco.
- Bachman, L. F. (2006c). *A research use argument: An alternative paradigm for empirical research in applied linguistics?* Paper presented at the Annual Meeting of the American Association for Applied Linguistics, Ottawa, CA.
- Bachman, L. F. (2007). What is the construct? The dialectic of abilities and contexts in defining constructs in language assessment. In J. Fox, M. Wesche, & D. Bayliss (Eds.), *What are we measuring? Language testing reconsidered*. Ottawa: University of Ottawa Press.
- Bachman, L. F., & Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- 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, NY: Freeman.
- Black, P. & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, 5, 7-73.
- Black, P. & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*. 80, 139-144.
- Black, P. & Wiliam, D. (2005). Changing teaching through formative assessment: Research and practice. In CERI (Ed.), *Formative assessment: Improving learning in secondary classrooms* (pp. 223–240). Paris: Organisation for Economic Co-operation and Development.
- Black, P. & Wiliam, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning: Theory, practice and policy* (pp. 9–25). London: Sage.

- Blanche, P. (1988). Self-assessment of foreign language skills: implications for teachers and researchers. *RELC Journal*, 19, 75-96.
- Bol, L. & Strage, A. (1996). The contradiction between teachers' instructional goals and their assessment practices in high school biology courses. *Science Education*, 80, 145-163.
- Boston, C. (2002). The concept of formative assessments. *Practical Assessment, Research and Evaluation*, 8, 9-14. Retrieved March 11, 2010, from <http://pareonline.net/getvn.asp?v=8&n=9>.
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society, *Studies in Continuing Education*, 22, 151-167.
- Boud, D. & Falchikov, N. (1989). Quantitative studies of student self assessment higher education: A critical analysis of findings, *Higher Education*, 18, 529-549.
- Brindley, G. (1994). Competency-based assessment and reporting in language learning programs: Some issues and questions. *Prospect*, 9, 41-55.
- Brindley, G. (2000). Task difficult and task generalizability in competency-based writing assessment. In G. Brindley (Ed.). *Studies in Immigrant English Language Assessment* (vol. 1, pp. 45-80). Sidney: National Centre for English Language Teaching and Research. Macquarie University.
- Brindley, G. (2001). Outcomes-based assessment in practice: some examples and emerging insights. *Language Testing*, 18, 393-407.
- Brookhart, S. M. (2005, April). *Research on formative classroom assessment*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.



- Brookhart, S. M., & Durkin, D. (2003). Classroom assessment, student motivation, and achievement in high school social studies classes. *Applied Measurement in Education, 16*, 27-54.
- Butler Y., & Jiyoon, L. (2010). The effects of self-assessment among young learners of English. *Language Testing, 27*, 5-31.
- Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G., Barbaranelli, C., & Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology, 100*, 525-534. doi:10.1037/0022-0663.100.3.525
- Chatterji, H. (2003). *Designing and using tools for educational assessment*. Boston, MA: Allyn and Bacon.
- Chickering, A. W. & Gamson Z. F.(Eds.). (1991). *Applying the seven principles for good practice in undergraduate education*. San Fransisco, CA: Jossey-Bass.
- Clariana, R. (1993). A review of multiple-try feedback in traditional and computer-based instruction. *Journal of Computer-Based Instruction, 20*, 67-74.
- Choen, V. (1985, January). A reexamination of feedback in computer-based instruction: Implications for instructional design. *Educational Technology, 25*, 33-37.
- Colby-Kelly, C., & Turner, C.E. (2007). AFL research in the L2 classroom and evidence of usefulness: Taking formative assessment to the next level. *Canadian Modern Language Review, 64*, 9-37.
- Covington, M. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. Cambridge, England: Cambridge University Press.

- Crooks, T. (1988). The impact of classroom evaluation on students. *Review of Educational Research, 58*, 438-481.
- Cubukcu, F., (2008). A study on the correlation between self efficacy and foreign language learning anxiety. *Journal of Theory and Practice in Education, 4*, 148-158.
- Davidson, F. & Lynch, B. (2002). *Testcraft*. New Haven: Yale.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL assessment practices in Australian and Hong Kong secondary schools. *Language Testing, 21*, 305-334.
- Dickinson, L. (1987). *Self-instruction in language learning*. Cambridge: Cambridge University Press.
- Dickinson, L. & Carver, D. J. (1980). Learning how to learn: Steps towards self-direction in foreign language learning in schools, *ELT Journal, 35*, 1-7.
- Eriksson Gustavsson, A., & Samuelsson, J. (2011). The importance of self-assessed oral and written language ability for boys' and girls' grades in school. *Journal of Theory and Practice in Education, 7*, 174-191.
- Finch, A. E. (2001). *English Conversation: Self-assessment in the Language Class*. Retrieved November 25, 2011 from [http://www.finchpark.com/arts/English\\_Conversation.pdf](http://www.finchpark.com/arts/English_Conversation.pdf).
- Fontana, D. & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology, 64*, 404-417.
- Forbes, E.W. (2007). *Improving the knowledge and use of formative assessment: A case study of a model of formative assessment in a K-3 science curriculum*. Unpublished Ed.D.

- dissertation, University of Delaware, Delaware. Retrieved April 19, 2010, from Dissertations & Theses: Full Text.(Publication No. AAT 3267194).
- Gareis, C. (2006). Collaborative leadership: The forgotten art of formative assessment. *Training and Technical Assistance Center* newsletter, School of Education, College of William and Mary. Retrieved March 11, 2010 from <http://education.wm.edu/centers/ttac/documents/newsletters/2006febmar.pdf>.
- Geeslin, K. (2003). Student self-assessment in the foreign language classroom: The place of authentic assessment instruments in the Spanish language classroom. *Hispania*, 86, 857-868.
- Hancock, C. (1994). Alternative assessment and second-language study: What and why? *ERIC Digests*. Washington, DC: ERIC Publications.
- Harlen, W., & Winter (2004). The development of assessment for learning: Learning from the case of science and mathematics. *Language Testing*, 21, 390–408.
- Harris, M. (1997). *Self-assessment of language learning in formal settings*. In *ELT Journal* VOL. 51/1, pp. 12-20. Oxford, England: Oxford University Press.
- Joet, G., Usher, E. L., Bressoux, P. (2011). Sources of self-efficacy. *Journal of Educational Psychology*, 103, 649-663.
- Leung, C. & Mohan, B. (2004). Teacher formative assessment and talk in classroom contents: Assessment as discourse and assessment of discourse. *Language Testing*, 21, 335-359.
- Lynch, B. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18, 351-372.
- Lynch, B. (2003). *Language Assessment and Program Evaluation*. New Haven, CT: Yale.

- McDonald, B. & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education, 10*, 209-220.
- McNamara, T. (2001a). Editorial: Rethinking alternative assessment. *Language Testing, 18*, 329–332.
- McNamara, T. (2001b). Language assessment as social practice: Challenges for research. *Language Testing, 18*, 333–349.
- Miller, L. & Ng, R. (1996). Autonomy in the classroom: Peer assessment. In R. Pemberton, S.L. Edward, W.W.F. Or, and H.D. Piersno (Eds.), *Taking control: Autonomy in language learning*. (pp. 133-146). Hong Kong: Hong Kong University Press.
- Moss, P. (1994). Can there be validity without reliability. *Educational Researcher, 23*, 5-12.
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*, 30-38.
- Olson, L. (2005). Classroom assessments stir glowing global interest. *Education Week, 25*, 8.
- Oscarson, M. (1989). Self-assessment of language proficiency: rationale and applications. *Language Testing, 6*, 1-13.
- Oscarson, M. (1997). Self-assessment of foreign and second language proficiency. In C. Clapham and D. Corson (Eds.), *Encyclopedia of language and education: Language testing and assessment* (vol. 7, pp. 175-187). Dordrecht, Netherlands: Kluwer Academic.
- Pajares, F., Miller, M. D. & Johnson, M. J. (1999). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203.

- Pajares, F., & Schunk, D. H. (2005). Self-efficacy and self-concept beliefs: Jointly contributing to the quality of human life. In H. Marsh, R. Craven & D. McInerney (Eds.), *International advances in self-research*, (vol. 2, pp. 95-121). Greenwich, CT: Information Age.
- Poehner, M., & Lantolf, J. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9, 233–265.
- Radford, B. W. (2010). *The effect of formative assessments on teaching and learning* (Unpublished master's thesis). Brigham Young University, Provo, UT.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18, 429-462.
- Rea-Dickins, P. (2004). Editorial: Understanding teachers as agents of assessment. *Language Testing*, 21, 249-258.
- Rea-Dickins, P. & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, 17, 215-243.
- Ross, S. (1998). Self-assessment in language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, 15, 1-20.
- Ross, S. (2005). The impact of assessment method on foreign language proficiency growth. *Applied Linguistics*, 26, 317-342.
- Sadler, D. (1998). Formative assessment: revisiting the territory. *Assessment in Education*, 5, 77-84.
- Schmoker, M. (2001). *The results fieldbook: Practical strategies from dramatically improved schools*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Shohamy, E. (2001). *The power of tests*. London: Longman.

- Shohamy, E. (2004). Democratic assessment as an alternative. *Language Testing*, 18, 373–391.
- Stiggins, R. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Stiggins, R. (2002). The assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83, 758-765.
- Stiggins, R. (2005a). *Student-involved assessment for learning*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Stiggins, R. (2005b). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, 87, December, 2005.
- Stiggins, R. (2007). Assessment through the student's eyes. *Educational Leadership*, 64(8), 22-26.
- Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2004). *Classroom assessment for student learning: doing it right-using it well*. Portland, OR: Assessment Training Institute.
- Teasdale, A., & Leung, C. (2000). Teacher assessment and psychometric theory: A case of paradigm crossing. *Language Testing*, 17, 163-184.
- Triggs, P., Weeden, P., Winter, J., & Broadfoot, P. (2000). The LEARN project Phase 2: Developing assessment in the primary school project report (University of Bristol Graduate School of Education CLIO Centre for Assessment Studies). Report for the Qualifications and Curriculum Authority. Retrieved February 10, 2010, from [http://www.qca.org.uk/qca\\_4490.aspx](http://www.qca.org.uk/qca_4490.aspx).
- Usher, E. L., & Pajares, F. (2008a). Self-efficacy for self-regulated learning: A validation study. *Educational and Psychological Measurement*, 68, 443-463.

- Usher, E. L., & Pajares, F. (2008b). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*, 751-796.
- Van de Watering, G., & Van der Rijt (2006). Teachers' and students' perceptions of assessments: A review and a study into the ability and accuracy of estimating the difficulty levels of assessment items. *Educational Research Review, 1*, 133-147.
- Vygotsky, L (1998). The problem of age. In R.W. Reiber (Ed.), *The collected works of L.S. Vygotsky*, (vol. 5, pp. 187-205). *Child psychology*. New York, NY: Plenum.
- Weeden, P., & Winter, J. (1999). The LEARN project: Learners' expectations of assessment for learning nationally. (University of Bristol Graduate School of Education CLIO Centre for Assessment Studies). Report for the Qualifications and Curriculum Authority. Retrieved February 10, 2010, from [http://www.qca.org.uk/qca\\_4490.aspx](http://www.qca.org.uk/qca_4490.aspx).
- Whiting, B., Van Burgh, J.W., & Render, G.F. (1995). Mastery learning in the classroom. Retrieved February 10, 2010, from <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED382688>
- Yang, H. (2007). Impact of language, anxiety and self-efficacy on accessing internet sites. *Cyberpsychology & Behavior, 10*, 226-233.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*, 64-70.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*, 166-183.

## Appendix A: Handouts for Missionaries

Group Instructions (Spanish)		Group 1
<b>Overview</b>		
Your district has been selected to participate in a project involving Language assessments administered the computer labs. It is very important that your missionaries take the assessments according to the following schedule.		
<b>Schedule</b>		
Week 1	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Entrance Questionnaire</li> <li>Elicited Imitation</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>The Entrance Questionnaire can be taken through the AssessmentTools application in the labs.</li> <li>The Elicited Imitation can be taken at <a href="http://ei.mtc.byu.edu">http://ei.mtc.byu.edu</a>, selecting MTC and Spanish in the drop downs to the right.</li> </ol>
	<b>Researcher Action</b>	Spend 5-10 minutes with the missionaries describing the intent of this study. Ask the missionaries if they are willing to participate in the study and if they are, please have them sign a consent form. Take the missionaries into a lab to complete the two assessments above.
Weeks 2-7	<b>Assessment Title</b>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>The missionaries in this group will not be taking the Language Speaking Assessment until their last full week.</li> </ol>
Week 8-9	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Language Speaking Assessment</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Take Assessment to measure how they have improved in the language.</li> </ol>
Week 8-9	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Exit Questionnaire</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Fill out online questionnaire</li> </ol>
	<b>Research Action</b>	Ensure that the missionaries to make sure that they have taken both the LSA and Exit Questionnaire.



## Group Instructions (Spanish)

## Group 2

## Overview

Your district has been selected to participate in a project involving Language assessments administered the computer labs. It is very important that your missionaries take the assessments according to the following schedule.

## Schedule

Week 1	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Entrance Questionnaire</li> <li>Elicited Imitation</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>The Entrance Questionnaire can be taken through the AssessmentTools application in the labs.</li> <li>The Elicited Imitation can be taken at <a href="http://ei.mtc.byu.edu">http://ei.mtc.byu.edu</a>, selecting MTC and Spanish in the drop downs to the right.</li> </ol>
	<b>Researcher Action</b>	Spend 5-10 minutes with the missionaries describing the intent of this study. Ask the missionaries if they are willing to participate in the study and if they are, please have them sign a consent form. Take the missionaries into a lab to complete the two assessments above.
Weeks 2-7	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Language Speaking Assessment</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Take Assessments</li> <li>Missionaries should take the Language Speaking Assessment as often as they would like, practice, and compare their responses to others.</li> </ol>
Week 8-9	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Language Speaking Assessment</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Take Assessment to measure how they have improved in the language.</li> </ol>
Week 8-9	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Exit Questionnaire</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Fill out online questionnaire</li> </ol>
	<b>Research Action</b>	Ensure that the missionaries to make sure that they have taken both the LSA and Exit Questionnaire.

## Group Instructions (Spanish)

## Group 3

## Overview

Your district has been selected to participate in a project involving Language assessments administered the computer labs. You will also participate in Computer Assisted Rater Instruction (CARI) to learn the criteria by which your language speaking ability will be rated. It is very important that your missionaries take the assessments according to the following schedule.

## Schedule

Week 1	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Entrance Questionnaire</li> <li>Elicited Imitation</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>The Entrance Questionnaire can be taken through the AssessmentTools application in the labs.</li> <li>The Elicited Imitation can be taken at <a href="http://ei.mtc.byu.edu">http://ei.mtc.byu.edu</a>, selecting MTC and Spanish in the drop downs to the right.</li> </ol>
	<b>Researcher Action</b>	Spend 5-10 minutes with the missionaries describing the intent of this study. Ask the missionaries if they are willing to participate in the study and if they are, please have them sign a consent form. Take the missionaries into a lab to complete the two assessments above.
Week 2	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>CARI</li> <li>CARI Questionnaire</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Missionaries should use the CARI product available in the labs</li> <li>During CARI training, the missionaries should complete the CARI questionnaire</li> </ol>
	<b>Researcher Action</b>	Take the missionaries to the lab and help them access the CARI product. Spend time with the missionaries to make sure that there are no questions regarding the CARI training. After each missionary completes the CARI training, have them fill out the CARI questionnaire.
Week 6	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Spanish CARI</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>The missionaries should return to the labs and complete the CARI certification in the Spanish language.</li> </ol>
Week 8-9	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Language Speaking Assessment</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Take Assessment to measure how they have improved in the language.</li> </ol>
Week 8-9	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Exit Questionnaire</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Fill out online questionnaire</li> </ol>
	<b>Research Action</b>	Ensure that the missionaries to make sure that they have taken both the LSA and Exit Questionnaire.

## Group Instructions (Spanish)

## Group 4

## Overview

Your district has been selected to participate in a project involving Language assessments administered the computer labs. You will also participate in Computer Assisted Rater Instruction (CARI) to learn the criteria by which your language speaking ability will be rated. It is very important that your missionaries take the assessments according to the following schedule.

## Schedule

Week 1	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Entrance Questionnaire</li> <li>Elicited Imitation</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>The Entrance Questionnaire can be taken through the AssessmentTools application in the labs.</li> <li>The Elicited Imitation can be taken at <a href="http://ei.mtc.byu.edu">http://ei.mtc.byu.edu</a>, selecting MTC and Spanish in the drop downs to the right.</li> </ol>
	<b>Researcher Action</b>	Spend 5-10 minutes with the missionaries describing the intent of this study. Ask the missionaries if they are willing to participate in the study and if they are, please have them sign a consent form. Take the missionaries into a lab to complete the two assessments above.
Week 2	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>CARI</li> <li>CARI Questionnaire</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Missionaries should use the CARI product available in the labs</li> <li>During CARI training, the missionaries should complete the CARI questionnaire</li> </ol>
	<b>Researcher Action</b>	Take the missionaries to the lab and help them access the CARI product. Spend time with the missionaries to make sure that there are no questions regarding the CARI training. After each missionary completes the CARI training, have them fill out the CARI questionnaire.
Weeks 2-7	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Language Speaking Assessment</li> <li>Spanish CARI (week 6)</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Missionaries should take the Language Speaking Assessment as often as they would like, practice, and compare their responses to others.</li> <li>Missionaries should complete the CARI certification in Spanish.</li> </ol>
	<b>Researcher Action</b>	
Week 8-9	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Language Speaking Assessment</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Take Assessment to measure how they have improved in the language.</li> </ol>
	<b>Researcher Action</b>	
Week 8-9	<b>Assessment Title</b> <ul style="list-style-type: none"> <li>Exit Questionnaire</li> </ul>	<b>Missionary Action</b> <ol style="list-style-type: none"> <li>Fill out online questionnaire</li> </ol>
	<b>Researcher Action</b>	Ensure that the missionaries to make sure that they have taken both the LSA and Exit Questionnaire.

### Appendix B: Language Criteria Rubric

Category	Rating/Score						
	1	2	3	4	5	6	7
Pronunciation	Pronunciation is usually unintelligible	Difficult to understand at times; frequent errors in sounds, stress, or intonation; accent inhibits communication or is irritating	Fairly easy to understand; some errors in sounds, stress or intonation; accent attracts attention but does not inhibit communication	Can be understood without difficulty; accent is not pronounced			
Grammar	Does not use language rules; speech consists mainly of individual words strung together, with no regard for correct forms or sentence structure	Uses a limited range of language rules; as many errors as correct forms	Correctly uses a fairly broad range of language rules most of the time; uses grammar that is clearly required by the situation	Consistently and correctly uses an extensive range of language rules as required by the situation; few if any errors, even in less common or complex forms			
Vocabulary	Vocabulary is inadequate to communicate intended ideas; often lacks even common, basic words and expressions	Uses some situation-specific vocabulary, but often lacks words and expressions needed to convey complete ideas; sometimes uses the wrong words or uses the same words repeatedly	Uses an adequate range of situation-specific vocabulary; words and expressions are sometimes imprecise, but speaker finds a way to convey intended meaning	Uses a broad range of appropriate and precise words and expressions needed to convey intended ideas; no searching for words			
Fluency	Speech is so slow or so fast that communication does not occur	Speech is slow enough (i.e. frequent or long pauses and fillers) or fast enough to cause discomfort to the listener	Rate of speech does not impede communication; occasional unnatural pauses and fillers do not distract significantly from the message	Rate and flow of speech are usually natural and facilitate communication			

\*Note: A “No Rating” or 0 score signifies that the speech sample is not ratable due to a non-language related cause (i.e. technological problems.)

### Appendix C: Entrance Questionnaire

1. What is your current age?
2. What is the highest level of school you have completed?
  - high school
  - less than one year of college
  - 1-2 years of college
  - 3 or more years of college
  - college graduate
3. How much did you like school?
  - I didn't like it.
  - I liked it okay.
  - I liked it a lot.
4. Overall, how well did you do in school?
  - I didn't do well.
  - I did okay.
  - I did well.
5. Did you study or speak your MISSION LANGUAGE (the language you will use in the mission field) before you came to the MTC?
  - No
  - Yes
6. Did you use your mission language as a common means of communication with your family or friends while you were growing up?
  - No
  - Yes
7. Have you lived in a foreign country where your mission language is spoken?
  - No
  - Yes

8. How long did you live in the country where your mission language is spoken?
- less than one month
  - 1-2 months
  - 3-6 months
  - 7-11 months
  - 1-2 years
  - more than two years
9. Have you done any formal study of your mission language (e.g. classes, self-study)?
- No
  - Yes
10. What kind of formal study of your mission language did you do? (Check all that apply.)
- elementary school class
  - jr. high school class
  - high school class
  - college or university class
  - school immersion program
  - private tutor
  - self-study (read books, listened to tapes, etc.)
  - other
11. How long was your total formal study of your mission language?
- 1 year or less
  - 2 years
  - 3 years
  - 4 years
  - 5 years or more
12. Overall, how well did you do in learning your mission language before your mission?
- I didn't do well.
  - I did okay.
  - I did well.
13. Did you study or speak a foreign language OTHER THAN YOUR MISSION LANGUAGE before you came to the MTC?
- No
  - Yes

14. What foreign language OTHER THAN YOUR MISSION LANGUAGE have you had the most experience with?
- Spanish
  - German
  - French
  - Russian
  - Japanese
  - Portuguese
  - Chinese
  - Italian
  - Korean
  - Other
15. Did you use this other foreign language as a common means of communication with your family or friends while you were growing up?
- No
  - Yes
16. Have you lived in a foreign country where this other foreign language is spoken?
- No
  - Yes
17. How long did you live in the country where this other foreign language is spoken?
- less than one month
  - 1-2 months
  - 3-6 months
  - 7-11 months
  - 1-2 years
  - more than two years
18. Have you done any formal study of this other foreign language (e.g. classes, self-study)?
- No
  - Yes

19. What kind of formal study of this other foreign language did you do? (Check all that apply.)

- elementary school class
- jr. high school class
- high school class
- college or university class
- school immersion program
- private tutor
- self-study (read books, listened to tapes, etc.)
- other

20. How long was your total formal study of this other foreign language?

- 1 year or less
- 2 years
- 3 years
- 4 years
- 5 years or more

21. Overall, how well did you do in learning this other foreign language?

- I didn't do well.
- I did okay.
- I did well.

22. Are you a(n):

- Elder
- Sister



### Appendix D: Exit Questionnaire

1. While at the MTC, I made \_\_\_\_\_ in developing my study skills.
  - no progress
  - a little progress
  - some progress
  - quite a bit of progress
  - a lot of progress
  
2. While at the MTC, I made \_\_\_\_\_ in my ability to plan and set goals.
  - no progress
  - a little progress
  - some progress
  - quite a bit of progress
  - a lot of progress
  
3. While at the MTC, I made \_\_\_\_\_ in my ability to work hard.
  - no progress
  - a little progress
  - some progress
  - quite a bit of progress
  - a lot of progress
  
4. While at the MTC, I made \_\_\_\_\_ in speaking and understanding my mission language.
  - no progress
  - a little progress
  - some progress
  - quite a bit of progress
  - a lot of progress
  
5. While at the MTC, I learned to teach the doctrine from the following lessons in my mission language: (Mark all that apply)
  - The Restoration
  - The Plan of Salvation
  - The Gospel of Jesus Christ
  - The Commandments (three or more commandments)
  - Laws and Ordinances
  - None of the above

6. Personal study time at the MTC was \_\_\_\_\_ in preparing me to be a successful missionary.

- not helpful
- somewhat helpful
- quite helpful
- very helpful
- extremely helpful

7. During the last seven days, about \_\_\_\_\_ percent of my daily speech was done in my mission language. (Please enter a number from 0-100.)

8. During the last two weeks, my teachers \_\_\_\_\_ spoke the mission language while teaching or working with the missionaries.

- never
- rarely
- sometimes
- often
- always

9. On average, I spent about \_\_\_\_\_ minutes a day studying my mission language. (*Please enter a number, not a word*)

10. How often did you study your mission language with your companion outside of class?

- Never
- Less than once a week
- Once or twice a week
- Several times a week
- Daily

11. I was \_\_\_\_\_ discouraged because I could not learn my mission language as fast as I wanted to.

- almost always
- often
- sometimes
- rarely
- never

12. The language learning activity that was most helpful to me at the MTC was: (*Write "no comment" if you do not wish to respond.*)

13. Please give one or two suggestions that would have helped you learn your mission language better at the MTC. (*Write "no comment" if you do not wish to respond.*)

14. Teaching progressing investigators was \_\_\_\_\_ in learning my mission language.
- not helpful
  - somewhat helpful
  - quite helpful
  - very helpful
  - extremely helpful
15. Chapter 7 of *Preach My Gospel* was \_\_\_\_\_ in learning my mission language.
- not helpful
  - somewhat helpful
  - quite helpful
  - very helpful
  - extremely helpful
  - I did not read *Preach My Gospel*, Chapter 7
16. Using a language dictionary was \_\_\_\_\_ in learning my mission language.
- not helpful
  - somewhat helpful
  - quite helpful
  - very helpful
  - extremely helpful
  - I did not use a language dictionary
17. My language textbook (e.g. Spanish for Missionaries) was \_\_\_\_\_ in learning my mission language.
- not helpful
  - somewhat helpful
  - quite helpful
  - very helpful
  - extremely helpful
  - I did not use a language textbook
18. The Speak Your Language program was \_\_\_\_\_ in learning my mission language.
- not very helpful
  - somewhat helpful
  - quite helpful
  - very helpful
  - extremely helpful

19. TALL (Technology Assisted Language Learning) was \_\_\_\_\_ in learning my mission language.

- not helpful
- somewhat helpful
- quite helpful
- very helpful
- extremely helpful
- I did not use the TALL program

20. Have you used TALL (Technology Assisted Language Learning) to help you learn your mission language?

- Yes
- No
- TALL is not available in my mission language

21. How often did you use TALL at the MTC?

- Less than once a week
- Once or twice a week
- Several times a week
- Daily
- Several times a day

22. Which TALL features were most helpful to you? (Check all that apply.)

- Preach My Gospel lessons
- Language tasks
- Vocabulary and phrases
- Grammar
- Listening comprehension

23. List one or two suggestions that could improve the effectiveness of TALL. (*Write "no comment" if you do not wish to respond.*)

24. Do you have some personal, written goals for your work at the MTC?

- Yes
- No

25. Please list one or two of your goals.