Olivet Nazarene University
Digital Commons @ Olivet

Ed.D. Dissertations | School of Graduate and Continuing Studies

5-2015

Alternative Assessments and Student Perceptions in the Foreign Language Classroom

Robin Barnard Bachelor
Olivet Nazarene University, Rbrb@mtco.com

Follow this and additional works at: https://digitalcommons.olivet.edu/edd_diss

Part of the Bilingual, Multilingual, and Multicultural Education Commons, Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, and the Secondary Education and Teaching Commons

Recommended Citation
Bachelor, Robin Barnard, "Alternative Assessments and Student Perceptions in the Foreign Language Classroom" (2015). *Ed.D. Dissertations*. 74.
https://digitalcommons.olivet.edu/edd_diss/74

This Dissertation is brought to you for free and open access by the School of Graduate and Continuing Studies at Digital Commons @ Olivet. It has been accepted for inclusion in Ed.D. Dissertations by an authorized administrator of Digital Commons @ Olivet. For more information, please contact digitalcommons@olivet.edu.

ALTERNATIVE ASSESSMENTS AND STUDENT PERCEPTIONS IN THE FOREIGN LANGUAGE CLASSROOM

by

Robin Barnard Bachelor

Dissertation

Submitted to the Faculty of

Olivet Nazarene University

School of Graduate and Continuing Studies

in Partial Fulfillment of the Requirements for

the Degree of

Doctor of Education

in

Ethical Leadership

May 2015

218148

ALTERNATIVE ASSESSMENTS AND STUDENT PERCEPTIONS

IN THE FOREIGN LANGUAGE CLASSROOM

By

Robin Barnard Bachelor

Dissertation

Dissertation Adviser	4-18-15 Date
Dissertation Reader	4-18-15 Date
Jeffrey G. Williamson Dissertation Coordinator	4-18-15 Date
Program Director	4-18-15 Date
Vice President for Academic Affairs	4-18-15 Date

©2015

Robin Barnard Bachelor

All Rights Reserved

ACKNOWLEDGEMENTS

Five years ago, I never would have dreamed that I would be completing my dissertation and doctoral studies at Olivet Nazarene University. This experience has impacted me in more ways than I could expound on in one page. My view of life, relationships, and leadership has positively changed forever. Through this uphill journey, I would be remiss if I did not credit my support team. I would like to thank my research-expert and diligent advisor, Dr. Kashama Mulamba (or endearingly, Kash) for all of his guidance, clarity, and expertise. My dissertation reader, Dr. Charles Perabeau, was a source of encouragement and kept my heels firmly planted. All of my professors steadily progressed me toward my goals and facilitated change and growth in my academic, professional, spiritual, and social realms. I am also grateful to School X and the Spanish teacher there for such an agreeable, cooperative experience during the data collection process. I offer special thanks to my family: my devoted husband, Ken, and my uplifting children, Heather, Joe, Jeremy, and Jessica, as well as my loving parents. I am also in debt to my Cohort 7 members and all of my friends and extended family for their love, patience, and sustenance during these grueling yet rewarding years. Finally, but not last, I send praise, thanks, and adoration to my Heavenly Father who led me into this process, unconditionally loved and sustained me throughout it, and will move me into whatever future He holds.

DEDICATION

I dedicate this dissertation to my family without whom success would have been difficult, if not impossible. No matter what happens in life, we have the power to respond and act in a manner that is pleasing to our Lord Jesus Christ. With this dedication, I implore my loved ones to commit all their ways to the Lord, always to set high goals and work diligently toward their completion, and to maintain hope for a bright future. I love you, Ken, Heather, Joe, Jeremy, Jessica, Scarlette, Berenger, Mom, and Dad!

ABSTRACT

The paradigm shift in education with the implementation of the Common Core State Standards has created the opportunity for foreign language educators to evaluate appropriate and beneficial assessments for their students. This study investigated how first-year students in a Midwestern high school perceived three different alternative assessments in the foreign language classroom: Dynamic Assessment, Task-based Assessment, and Formative Assessment using self- and peer-evaluation. The researcher correlated the perceptions to the students' assessment scores. Additionally, the researcher compared final exam scores of the control group to those of the experimental group. The results indicated that the experimental group students favorably perceived the alternative assessments types, chose Formative Assessment as their most preferred and Task-based Assessment as the least desired, and scored as well as the control group students on the final exam.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
	Statement of the Problem	2
	Background	4
	Research Questions	9
	Description of Terms	9
	Significance of the Study	11
	Process to Accomplish	12
	Summary	19
II.	REVIEW OF LITERATURE	20
	Introduction	20
	Dynamic Assessment	20
	Task-based Assessment	30
	Formative Assessment using Self- and Peer-Evaluations	44
	Student Perceived Value	56
	Conclusion	61
	Summary	62
III.	METHODOLOGY	63
	Introduction	63

Chapter Page

Research Design ……..…..……..…….….…..... 63

Population……..…….…..... 67

Data Collection……..……... 67

Analytical Methods…….…….…… 69

Limitations…….…….……... 71

Summary……..……..…... 72

IV. FINDINGS AND CONCLUSIONS……..…..……..... 73

Introduction……..……..... 73

Findings……..……... 74

Conclusions……..... 87

Implications and Recommendations…… 92

REFERENCES……..... 97

APPENDIX…….……117

A. Student Perception Surveys…….…….…. 118

LIST OF TABLES

Table Page

1. Best-Perceived Assessment: Themes…….……..81

2. Worst-Perceived Assessment: Themes……82

LIST OF FIGURES

Figures Page

1. Dynamic Assessment perceptions: Blue……..……...……..... 75

2. Task-based Assessment perceptions: Red……...…….....76

3. Formative Assessment perceptions: Green..……...…….....77

4. Best-Perceived Assessment type.……..79

5. Worst-Perceived Assessment type.……80

6. Final Exam scores…….…84

7. Dynamic Assessment (Blue) Correlation……..85

8. Task-based Assessment (Red) Correlation……86

9. Formative Assessment (Green) Correlation……..87

CHAPTER 1

INTRODUCTION

In the ever-evolving world of education, practitioners are faced with the challenge of meeting the needs of all students as well as the demands of state performance evaluations. In an effort to address this quest, educators must research best practices in the areas of both classroom methodology and assessment and must create a perceived value to students. To date, 45 states have agreed to use the Common Core State Standards (CCSS) as their benchmark for achievement in public schools (National Governors Association Center for Best Practices, & Council of Chief State School Officers, 2010). According to Morrow, Shanahan, and Wixson (2012), the CCSS will require appropriate assessment choices which, in turn, will drive curriculum and instruction. They go on and argue that assessment choices are crucial for successful student preparation; the assessments should emphasize critical reading, writing, and higher-order thinking skills. These assessments are not remotely similar to the former, traditional state assessments but will be used to measure teacher success, evaluation, and retention. Teachers now find themselves in the position of designing alternative assessments that measure reading, writing, and higher-order thinking so that students are better prepared for Common Core Assessments. This task is further complicated for foreign language teachers who are instructed to uphold all of the following standards: National Standards for Foreign Language Learning, existing state standards, and the Common Core State Standards (American Council on the Teaching of Foreign Languages [ACTFL], 1998; Illinois State

Board of Education, 2012; National Governors Association Center for Best Practices, & Council of Chief State School Officers, 2010). As a result, this researcher developed a study to measure the effectiveness of certain alternative assessments that would meet some of the current demands.

The results of this study may be of interest to foreign language teachers at the high school level as they restructure their curriculum and assessments to meet current demands and state testing measures.

Statement of the Problem

A paradigm shift has occurred with assessment expectations for foreign language students at the secondary level (Mandell, 1999; Poehner & van Compernolle, 2011; Sidek, 2012; Tamjid & Birjandi, 2011). Alliance of Excellent Education (2011) explained the shift:

> Our increasingly complex world demands much of its students. In almost every aspect of their lives, young people are being asked to learn more, process more, and produce more. These increasing demands mirror the world around them. Now more than ever, the nation's education system is being challenged by a technology-driven global economy that requires a skilled and deeply literate workforce. (p. 1)

Foreign language teachers become an integral component of this paradigm shift as they prepare students to meet the CCSS testing. Porter, McMaken, Hwang, and Yang (2011) completed a study comparing traditional state standards and assessments to what will be required with the Common Core State Standards. They found that more than half of all questions on state assessments are memory or recall in nature. Conversely,

assessments for the Common Core will employ questions that require higher order thinking, specifically those involving writing. Test-takers will need to explain, create, and analyze more than simply choose a possible response.

Using traditional, summative testing in the classroom as a way for preparing for state testing is no longer sufficient. Poehner and van Compernolle (2011) discussed the need for teaching that promoted development and was a process, not just an end-of-lesson assessment. Sidek (2012) explained that traditional testing, which typically required students to comprehend and process specific data, had to be modified to include assessments with meaningful tasks that were more communicative in nature. Tamjid and Birjandi (2011) supposed that teachers needed to move away from traditional, one-answer assessments toward performance-based tests and assessments in which the students' personal accountability was raised. All three research groups posited that best practices included designing some form of alternative assessment in the L2 classroom. With the implementation of the Common Core State Standards, foreign language teachers will need to assess their students in a more communicative way and detach from traditional assessment methods. Students must also adapt to alternative ways of being assessed. Instead of being grammar-centered, teachers will need to design effective assessments with a communicative focus while continuing to cover essential grammatical concepts and typical vocabulary. Possible assessment types that would facilitate this change include Dynamic Assessment (Antón, 2009; Poehner & van Compernolle, 2011; Wei, 2011), Task-based Assessment (Byrnes, 2002; Carless, 2007; Sidek, 2012; Skehan & Foster, 1997), and Formative Assessment using peer- and self-evaluations (Bryant & Carless, 2009; Colby-Kelly & Turner, 2007; Hanrahan & Isaacs,

2001; Tamjid & Birjandi, 2011). The purpose of this research was to examine the perceptions of first-year, high school foreign language students concerning the impact of alternative assessment types; in order to achieve this purpose, the experimental group's summative assessment scores were compared to the scores of the control group who did not receive interventions in order to guide the development of future assessments.

Background

Common Core State Standards

According to Education Northwest (n.d.), in 2009 the Council of Chief State School Officers (CCSSO) and the National Governors Association (NGAC) directed an all-encompassing educational reform, the Common Core State Standards. These two organizations formed discussion groups that were representative of all aspects of education, from kindergarten through higher education. Concurrently, they submitted a College and Career Ready Standards document for feedback from educators, parents, communities, businesses, researchers, and from the state educational authorities. After multiple drafts containing educational standards were reviewed and revised, a panel of 25 members released the first public draft of the Common Core State Standards in 2010. During this year, the panel released a final version with the following statement:

> Unlike past standards setting efforts, the Common Core State Standards are based on best practices in national and international education, as well as research and input from numerous sources … The Common Core State Standards represent what American students need to know and do to be successful in college and careers. (Education Northwest, para. 7)

Each state, as well as the District of Columbia, Puerto Rico, Guam, American Samoan Islands, U. S. Virgin Islands, and Northern Mariana Islands, was asked to adopt the Common Core State Standards in place of or in addition to its existing standards; 48 states had agreed to date. National Governors Association Center for Best Practices and Council of Chief State School Officers (2010) touted the new standards as a way to provide consistency throughout the United States and as a clear set of learning guidelines for all teachers, parents, and students. These organizations added that the adoption of consistent, rigorous standards would position students better in this global economy. The common core initiative would help ensure more continuity for students from kindergarten to college or the work place. Concurring with National Governors Association Center for Best Practices and Council of Chief State School Officers, Illinois State Board of Education (2012) believes that the Common Core State Standards

> are designed to be relevant to the real world, reflecting the knowledge and skills that our young people need for success in both college and work… As students advance through the grades and master the standards in reading, writing, speaking, listening, and language, they are able to exhibit with increasing fullness and regularity [the] capacities of the literate individual. (pp. 1-4)

Foreign language teachers are charged with following the grades 6-12 literacy standards as well as the current Illinois state standards for the teaching of foreign language. With this charge, new teaching practices must be evaluated. Teachers will be designing new assessments and discerning best classroom practices so that students are more literate and better positioned for life.

Alternative Assessments

While creating valuations that prepare students for the Common Core State Assessments and continuing to follow the Illinois State Standards may be challenging, plausible alternatives can be developed by foreign language teachers as they adapt to the paradigm shift.

One type of alternative assessment is Dynamic Assessment, based on Vygotsky's (1978) theory that teacher interventions with clear examples and instruction, along with individualized prompts, would help students move into self-reliance and mastery. Antón (2009) used a pre/posttest method to go along with Vygotsky's theory. After the initial assessment, the researcher documented specific points of remediation to address during the unit. She conferenced with her students to discuss their individual needs and then retested them. Using a similar approach, Wei (2011), who sought to use Vygotsky's theory, developed dynamic assessments for her classroom. She discovered her students' needs through an initial analysis followed by goal setting. Wei then planned assessments and made her students comfortable with the testing process. She continuously re-evaluated and revised her plans to meet the needs of her students and gave constant feedback both verbally and through multiple, short assessments during the unit. The format used for the current study was Dynamic Assessment focused on conferencing and group strategy sessions as well as a pre/posttest, similar to the unit designs of Antón and Wei.

Another potential alternative assessment is Task-based Assessment. According to Byrnes (2002), Task-based Assessment focused on language use and meaning that is contextualized in a communicative manner, typically through writing. The purpose of

Byrnes' research was to promote a shift from grammar-based instruction through a different assessment approach. Byrnes created rubrics to evaluate students' work to ensure consistency and use as a guideline for students while they worked. Assessments were created with communication in a real-world context, or as close to real-world as possible, at the forefront while still addressing the content needs of the textbook. The assessments were all writing assignments. The current study sought to assess student perceptions of Task-based Assessment using a real-world context. Similar to Byrnes, Ke (2006) also researched Task-based Assessment in order to promote communicative competency. Ke used a combination of reading, writing, listening, and speaking skills with real-world task assessments. The current research for this study incorporated all four skills, just as Ke had done. Carless (2007) continued the Task-based Assessment research by interviewing secondary teachers who used this assessment in their classrooms. He concluded that some traditional teaching methods incorporating grammar should preclude any Task-based Assessment. For the purpose of this study, teaching grammar, such as Carless suggested, was incorporated, along with the creation of an assessment combining reading, writing, listening, and speaking skills. Byrnes, Ke, and Carless all found that communicative competency increased with their students by implementing Task-based Assessments.

Formative Assessment using self- and peer-evaluation forms is another viable alternative assessment. Formative Assessment grounds itself in evaluating students throughout a unit so that students are more aware of the end of unit goals and can address their own needs. Tamjid and Birjandi (2011) realized a need for altering traditional assessments in order to stay current with the movement toward learner-centered

classrooms. In their study, all students were assigned writing prompts that the instructors graded; however, the experimental group also completed self-assessment rubrics and were randomly given another student's assignment to assess using the same rubric. The students all revised their assignments before submitting to the instructors. Tamjid and Birjandi found that the experimental group improved their metacognition which led to better thinking and learning skills that could be used on future assignments. Bryant and Carless (2009), also used peer-assessment methods; they hypothesized that self- and peer-assessments would stimulate more learner independence and create an atmosphere in which students desired improvement through reflective thinking. Their students viewed the peer-assessment rubric as a tool to help them earn better grades instead of a waste of time. The current study made use of Tamjid and Birjandi's process of self- and peer-assessment as a reflective measure to improve student writing while incorporating unit grammar and vocabulary into the prompt.

Student Perceived Value

To continue the thoughts of a learner-centered classroom with student reflective thinking, the critical part of this research was student perceptions. Students must perceive value for any incentive; this is critical to success in the classroom. According to Palloff and Pratt (2007), student perceived value is essential. They suggested that educators design activities that interest students by relating to their life activities, communicate clear expectations, create positive classroom atmospheres, and make use of alternative assessments. According to Riedinger (2006), students must develop a sense of ownership, or buy-in, in order to develop higher-order thinking skills in which they can intensely reflect and critically analyze information with improved performance as the end goal. An

essential component of the Common Core State Standards is promotion of higher-order thinking (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). In this research study, student perceived value was evaluated in an effort to find best assessment practices and to encourage higher-order thinking.

Research Questions

For this study the following questions were developed in order to focus the study and determine appropriate methodology:

1. What perceptions do first-year foreign language students have of Dynamic Assessment, specifically focused on conferencing and group strategy sessions?
2. What perceptions do first-year foreign language students have of Task-based Assessment, specifically focused on real world communication?
3. What perceptions do first-year foreign language students have of Formative Assessment, specifically using peer- and self-evaluations?
4. How do the summative assessment scores of the students who received interventions differ from those who did not?
5. How do student perceptions of the alternative assessments correlate with their test scores?

Description of Terms

Alternative assessment. Alternative assessment is a way of measuring performance through a variety of open-ended, creative, or communicative means (Brown, 2004).

Common Core State Standards. This is an educational initiative to provide rigorous, consistent standards for kindergarten through twelfth grade. These standards encourage higher-order thinking, real-world assessments, and the use of critical reading with writing (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

Contextualized. *Any material or assessment that is contextualized is written with a real-world content or application* (Shrum & Glisan, 2010).

Dynamic Assessment (DA). Dynamic Assessment is the promotion of student growth and discernment of student abilities during the assessment process through a teacher's hints, prompts, and leading questions in order to promote cognitive functioning (Haywood & Tzuriel, 2002; Lantolf & Poehner, 2004).

L1. This term indicates the first language of the learner or language one, the student's native language (VanPatten & Williams, 2007).

L2. This term indicates the second language that the student is learning, language two (Ellis, 1997).

Peer-assessment. Peer-assessment is a process in which students assess other students' work in a structured manner (Hanrahan & Isaacs, 2001).

Self-assessment. Self-assessment is a process in which students assess their own work in a structured manner (Hanrahan & Isaacs, 2001).

Target Language. The target language is "the language the learner is trying to learn" (Ellis, 1997, p. 144).

Task-based Assessment. Task-based Assessments have activities that require students to "use language, with an emphasis on meaning, to obtain an objective" (Bygate, Skehan, & Swain, 2001, 245).

Traditional assessment. Traditional assessment is a way of measuring achievement through objective questioning that contain one correct or best answer, typically as a summative test (Brown, 2004).

Zone of Proximal Development (ZPD). The ZPD is the range of what students are able to do without any assistance to what they are able to do with interventions; the teacher gains understanding and promotes individual development to increase the zone (Vygotsky, 1978).

Significance of the Study

While many studies have been completed that show the benefits of certain alternative assessments in the foreign language classroom, very few included high school students as participants. Most included university-aged students as subjects (Antón, 2009; Byrnes, 2002; Carless, 2007; Ke, 2006; Tamjid & Birjandi, 2011; Wei, 2011). Hamidi (2010) addressed the necessity for more studies: "L2 practitioners are starved of much of the latest findings in L2 assessments as well as the assessment literature, excluding the fact that most of them are not of use to the daily assessment practices of teachers" (para. 69). In a quest to help meet the current need for additional L2 assessment research, this study specifically focused on students taking high school courses, typically taught in the initial year. Additionally, finding existing studies centered on multiple alternative assessments was difficult; the researcher chose to use three different types, Dynamic Assessment, Task-based Assessment, and Formative Assessment, based on current

research that touted these assessments as best practice choices in the L2 classroom (Poehner & van Compernolle, (2011; Sidek, 2012; Tamjid & Birjandi, 2011). The primary reason for providing this variety of alternative assessments was to discern student perceptions for comparison purposes and perceived interest. According to Bakar, Sulaiman, Akhtar, and Rafaai (2010), student motivation, based on their perceptions of methods and assessments, assisted teachers in organizing their learning goals and in choosing assessments. This research sought to find alternative assessments that reflected learning as well as motivated students. Foreign language teachers will need to develop new assessments by 2015, and this data provided information to spur discussion and encourage collaborative efforts.

Process to Accomplish

Population

According to the Illinois School Report Cards (2012), School X was a kindergarten through 12th grade district with 920 students, 300 of whom attended the high school. The ethnicity of the district was primarily White, at approximately 727 students. Black students encompassed 74 out of the total population. There were 55 Asian/Pacific Islanders, as well as 55 Multi-ethnic students. Nine students were Hispanic or Native American. One hundred and one low-income students attended. The district employed 66 full time teachers, with a 14:1 student to teacher ratio, and four administrators. Ninety-one members of the student body had an Individual Education Plan (IEP), according to the records from School X, and the average ACT score of the previous junior class was 24. The graduation rate was just over 95%. School X's location was metropolitan, lying on a major river, and was approximately 165 miles from two megacities in the

Midwestern section of the United States. Several connecting small towns, along with this mid-size city, composed the student population.

School X typically offered three sections of Spanish One each fall. This is a high school, freshman-level course. Generally, the students in these classes were eighth graders or high school freshmen; however, sophomores, juniors, and seniors are allowed to take Spanish One.

Sampling

The school's computerized scheduling tool randomly assigned students into each section. The school counselors made some adjustments to balance class sizes and to meet individual needs relating to special needs accommodations. The classroom teacher taught three sections of Spanish One during the fall of 2013. Students ranged from 13 to 15 years of age at the onset of the study. To choose the sample, the researcher included the first section of the day as the control group, 27 students, in which 14 agreed to participate in the study. This group had nine females and five males. The second and third sections were the experimental group, 54 students, in which 34 agreed to participate in the study. This group had 15 females and 19 males. The researcher designated the experimental and control groups before seeing the lists of students in each class to reduce any possibility of bias. This study was conducted with 48 total participants, students in their first year of foreign language study earning high school credit.

Methodology

The research was completed during the fall of 2013 over a 15 week time frame. The researcher chose a concurrent nested design study that fell within the mixed-method realm (Robson, 2011). The primary method was quantitative through the analysis of the

summative test scores of the control and experimental groups as well as data analysis correlating the perceptions of the experimental group with the test scores. Teacher journal observations stood as a secondary, qualitative measure. The teacher recorded naturally occurring data, notations of attitudes, behaviors, and comments relating to the alternative assessments used with the experimental group during the study in the journal. Additional qualitative measures required the researcher to note the students' overall preferences of alternative assessment type through the post survey open-ended questions by categorizing the responses into themes.

The School X teacher assessed the 14 members of the control group using fairly traditional, previously used quizzes and tests throughout the semester. The remaining 34 participants, the experimental group, took alternative assessments instead. The researcher sought the perceptions of these students concerning three different alternative assessments types. Through researcher-developed surveys, the researcher acquired information relating the students' interests in each assessment as well as the assessment's perceived accuracy in revealing what they learned. The classroom teacher tested all students at the completion of the study with a more traditional, summative exam in order to show the impact of alternative assessments as replacements for more traditional ones. These summative exam scores served to compare the performance of the control group with that of the experimental group. The classroom teacher also completed a journal of the experience, especially noting student reactions, behaviors, and comments during the alternative assessment process. The teacher journal with notations of student perceptions, the student perception surveys, and the summative assessment scores worked together to

show best assessment choices. This mixed-method study, within the concurrent-nested design (Robson, 2011), assessed students' perceptions.

The researcher began by collecting data from the Educational Planning and Assessment System (EPAS) in School X's confidential computer server space. It was not possible for the researcher to give a pretest to the beginning Spanish learners, so the researcher used the most recent reading score for each student to determine the level of each student in the control and experimental groups. The researcher ran an independent *t* test on the two groups to show equivalency. The two groups were similar enough to have confidence that any difference in summative assessment data was due to the intervention.

During the 15 weeks of the study, the classroom teacher taught three units to the students while keeping a journal about the process. The teacher used the same teaching methods for covering the required vocabulary and grammar goals for the control and experimental groups. Departmental goals are established at School X and were followed for all classes in the control and experimental groups. The primary difference was in the assessment choices. The teacher administered the department's traditional, summative final exam to both the control and the experimental groups. The control group took traditional assessments through homework, quizzes, and traditional summative exams that included short answer, multiple choice, and true/false responses. In contrast, the experimental group took a Dynamic Assessment for unit one, a Task-based Assessment for unit two, and a Formative Assessment using peer- and self-evaluations involving writing assignments for unit three. The researcher gave a code word to each alternative assessment type: Blue for Dynamic Assessment, Red for Task-based Assessment, and Green for Formative Assessment using self- and peer-evaluations. At the end of each

unit, the students in the experimental group completed a researcher-developed survey with a 4-point Likert scale.

According to Leedy and Ormrod (2013), Likert-type scales should be used when a researcher is attempting to evaluate an attitude so that the attitudes can be simplified and then quantified. The survey allowed the researcher to discover how well the unit assessment reflected the students' knowledge and whether or not the students would like to be assessed in that way again. The response choices included strongly agree, agree, disagree, and strongly disagree. The end-of-research survey gave students two open-ended questions, asking for the students' highest and lowest preferences among the three alternative assessment types.

The same procedures were used to collect and analyze data for the following three research questions:

Research Question One: What perceptions do first-year foreign language students have of Dynamic Assessment, specifically focused on conferencing and group strategy sessions?

Research Question Two: What perceptions do first-year foreign language students have of Task-based Assessment, specifically focused on real world communication?

Research Question Three: What perceptions do first-year foreign language students have of Formative Assessment, specifically using peer- and self-evaluations?

The Likert-type surveys given to the experimental group at the end of the three units were analyzed. For each survey, the researcher categorized the students' responses and used a bar chart to express the experimental group's agreement level, how the particular assessment reflected learning. Since four-point, Likert-type data is measured on an interval scale, the researcher gave descriptive analysis of the tendencies. A one-way

ANOVA measured the associations between the student choices to show statistically significant differences by assessment type (Leedy & Ormrod, 2013). To follow up the one-way ANOVA, a Bonferroni post hoc determined where the differences existed. The researcher continued by writing descriptive, narrative accounts to relate the teacher journal observations after transcribing the journal into common themes. The themes revealed student perceptions of the alternative assessments from the teacher's perspective.

The next research question required comparative analysis to determine any differences between the final exam scores of the experimental and control groups:

Research Question Four: How do the summative assessment scores of the students who received interventions differ from those who did not?

An independent *t* test was used to "determine whether two groups of scores are significantly different" (Gay, Mills, & Airasian, 2012, p. 351). The means of the control and the experimental groups were compared to show any statistical significance using a box plot.

The final research question required analysis to determine any relationship between the experimental students' perceived value for each assessment and their test scores on each assessment:

Research Question Five: How do student perceptions of the alternative assessments correlate with their test scores?

Initially, a spreadsheet was composed to display each student's alternative assessment scores. The researcher continued by categorizing the student perceptions of the three alternative assessments. The researcher used a correlation coefficient formula to

determine a positive, negative, or nonexistent correlation between the students' perceptions and their alternative test scores. Correlational research "involves collecting data to determine whether, and to what degree, a relation exists between two or more quantifiable variables…a decimal number between -1.00 and +1.00" (Gay, et al., 2012, p. 624). For each assessment, the researcher correlated the student's alternative assessment score to his/her combined survey answers and displayed the results with scatter plots. The researcher used the student data to rank the three alternative assessments in order of preference, based on information from the post-study survey and then described these results. Narrative descriptions of the open-ended survey questions completed the analysis. The researcher took the students' words, found patterns, and developed themes to reveal correlative information.

Viability of the Study

The researcher requested and received approval from a School X administrator to complete this study. After meeting the students in August, 2013, the researcher explained the study and asked for the students to sign a research study assent form. The researcher sent home letters to the parent/guardian and acquired consent signatures. The researcher, in cooperation with the classroom teacher from School X, personally controlled all aspects of the study, including alternative assessment design for each unit, classroom procedures during tests, and data collection and storage procedures. The classroom teacher and the researcher also ensured that all students took the departmentally required summative assessment. Student data was accessed through the school secured server and self-reporting on the surveys.

Ethics

Minimal risks existed for the participants of this study. The students were treated the same as in any other school year or with any other teacher, using similar methodology, using the same text book, and following all rules and procedures at School X. The experimental group received the same vocabulary and grammar instruction as the control group, along with the required department summative exam. None of the department goals or objectives for Spanish One was altered for either group.

Summary

Educators today are transitioning from traditional standardized assessments toward ones that promote higher-order thinking and real-world applications. In order to prepare students to meet the expectations of new assessment measures at the state level, local teachers seek data from which they can gain ideas for their own, unique situations (Alliance for Excellent Education, 2011). This study explored three different alternative assessment types: dynamic, task-based, and formative, in which student perceptions of each assessment were correlated to their test scores. Further, summative test scores of both the control and experimental groups were compared. To accomplish this study, the researcher completed a comprehensive literature review of the three alternative assessment types, as well as the value of student perceptions.

CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

After making the decision to design alternative assessments for the L2 classroom, the researcher investigated the possible alternatives. Studies in the field of foreign language pedagogy and second language acquisition have indicated that Dynamic Assessment, Task-based Assessment, and Formative Assessment using self- and peer-evaluations show potential effectiveness (Asghar, 2010; Byrnes, 2002; Geeslin, 2012; Lantolf & Poehner, 2010). This chapter will discuss the history of relevant studies to show the need and significance of these alternative assessments in the L2 classroom. The researcher will incorporate literature presenting the importance of student perceptions as rationale for surveying the experimental group.

Dynamic Assessment

Dynamic Assessments, according to Poehner (2008), are "L2 pedagogical interactions, including the correction of over- and underestimates of learners' abilities, and the identification of problem areas…and the possibility of supporting learners' efforts to stretch beyond their current capabilities" (p. 176). Dynamic Assessment must involve mediation from a mentor who assists students with their learning. Through the mediation process, the teacher would gain a clearer understanding of each student's abilities through a pretest and then through interaction with the students. The teacher

would become more effective, then, in facilitating growth (Feuerstein, Rand, Jensen, Kaniel, & Tzuriel, 1987). Using a pretest to assess prior knowledge and determine areas of concern could be an effective way for teachers to meet the needs of their students. Garb (1997) describes the process very simply:

> The role of the evaluator is to identify the pupils' problems during the pretest and to provide the necessary mediation during the learning phase (mediation). Items on the posttest are identical to those of the pretest in level, background knowledge, grammatical structures, new terminology, and required strategies, but differ in content. (para. 17)

L2 educators may benefit from a needs analysis before beginning a unit. Teachers could be better equipped to customize their lesson planning with this knowledge and would also save time not covering material that students may already know. Dynamic Assessment, in essence, is a Formative Assessment method in which instruction and assessment can be integrated for the learners by promoting their development (Lantolf & Poehner, 2004). Wei (2011) touted Formative Assessment measures in the L2 classroom as a positive tool for improving learning, motivating students, and increasing self-confidence. Today's L2 teachers may find Dynamic Assessment useful as they implement various Formative Assessments into their routines.

The idea for Dynamic Assessment originated with the social development theory of Vygotsky (1978) who coined the term zone of proximal development (ZPD), according to Poehner (2008), ZPD refers to optimal learning conditions which are achieved when a student interacts with a teacher or more knowledgeable individual. The student progresses to his/her potential more quickly when collaborating with someone

else than when performing independently. While Vygotsky never used the term Dynamic Assessment or realized its potential as an assessment tool, his foundational work remains as groundbreaking for the theory (Vygotsky). Using Vygotsky's ideas, Campione and Brown (1985) found that

> Children experience cognitive activities in social situations and come to internalize them gradually over time. At the outset, the child and an adult work together, with the adult doing most of the work and serving as an expert model. As the child acquires some degree of skill, the adult cedes the child responsibility for part of the job and does less of the work. Gradually, the child takes more of the initiative, and the adult serves primarily to provide support and help when the child experiences problems. (p.5)

Modern day educators could make use of Vygotsky's theory as they interact with students in order to facilitate the best possible academic growth. Contrary to Piaget's theory that cognitive development precedes learning, Vygotsky believed that social learning actually occurs prior to cognitive development. Vygotsky posited that every child has a zone, or distance, between his/her ability to independently perform and the ability to perform with peer or adult collaboration. To this end, Vygotsky promoted active learning in the classroom, instead of traditional lecture methods, in which reciprocal exchanges of information occurred. Teachers facilitated student thinking during the assessment process, thus enabling students to construct meaning for themselves.

Feuerstein, Feuerstein, Falik, and Rand (1979), who began working with low-achieving students in the 1950's in Israel, researched and developed a technique called

the Learning Potential Assessment Device, similar to Vygotsky's theory. In their study, certain learner performance behaviors were identified and labeled as deficiencies. The classroom teacher in the study mediated the learning experience through addressing the labeled deficiencies. Feuerstein et al. named this mediation the elaboration phase in which data was used to promote growth. Finally, Feuerstein et al. staged an output phase in the classroom in which students were encouraged to problem-solve independently with teacher facilitation. The learners discovered test-taking strategies and logical ways to solve and remember test problems. Feuerstein et al. referred to this as a mediated learning experience, intended to maximize cognitive change. Vygotsky's (1978) theory and Feuerstein et al.'s research established what is known today as Dynamic Assessment and became models for many special education classrooms and future researchers, including the current study.

Continuing Dynamic Assessment studies in the United States during the 1960's, Budoff (1968) exchanged traditional IQ testing for the Dynamic Assessment alternative. He developed a system of test-train-retest and a formal system for interviewing learners. By using this method, Budoff claimed that students can profit from their experiences and assessments can show actual growth when students are more comfortable with the testing process. Lunt (1993) referred to Budoff's idea when he promoted "intelligence as the ability to profit from experience…to minimize the artificiality of the test (p. 163). Teachers today should be encouraged by Budoff's idea that familiarity with the assessment process can assist students in showing their intelligence and acquisition of information. As Budoff continued his work through the 1970's and 1980's, Campione and Brown (1985) began researching in a similar fashion. The researchers gave children

tasks or tests and observed their performance as well as their strategies for problem solving. Campione and Brown developed their own standardized interviewing system. Campione and Brown, along with Budoff, refined Vygotsky's (1978) theory and gave it practical use in the United States educational system. Foreign language teachers could consider these researchers' ideas as teachers design assessments that allow students to strategically determine solutions for test problems. Teachers could customize their own pretest, intervention, and posttest strategies.

Speece, Cooper, and Kibler (1990) began using Dynamic Assessment to determine individual differences. These researchers defined Dynamic Assessment as, "a training paradigm in which the examiner takes an active role in teaching…and then measures the degree to which this training resulted in learning" (p. 113). Speece et al. concluded that using Dynamic Assessment better identified candidates for special education than did traditional, static assessments while establishing that Dynamic Assessment benefited students and should play a greater role in educational practices. Furthering the body of research, Peña, Quinn, and Iglesias (1992), decided to test native Spanish speakers in the United States public education system for learning disorders by using Dynamic Assessment. They determined a problem with non-native English speakers who were being over-identified as special education candidates due to their lack of English language knowledge and United States testing methods. Allowing the Spanish-speaking students to practice and then work with a mentor significantly increased scores and eliminated many students from being labeled as special education. Studies such as those of Speece et al. and Peña et al. set the stage for special education teachers to

attempt Dynamic Assessment practices in their classrooms. However, Dynamic Assessment has use outside of the special education realm.

In recent years and with greater frequency, teachers have begun using Dynamic Assessment with foreign language learners. This researcher sought to incorporate Dynamic Assessment into the L2 classroom, as an alternative to traditional methods, in an effort to increase student learning potential. As a basis, this researcher investigated Kozulin and Garb (2002), who began their efforts with Dynamic Assessment due to a belief that effective learners "monitor their reading, plan strategies, adjust effort appropriately, and evaluate the success of their on-going efforts to understand" (p. 114). They recognized that learners needed assistance with monitoring their progress and planning strategies. Kozulin and Garb placed the burden of responsibility on the teacher, based on research from Palinscar and Brown (1984): strategy instruction is a best practice. This practice demands that teachers spend more time developing sound processes for learning and facilitate strategic problem solving activities, while spending less time concerned with the final product. Kozulin and Garb (2004) chose to leave the special education classroom and focus on second language learners. They used Dynamic Assessment to inform the instructor of learning potential for English reading comprehension activities. The students completed a pretest to acquire a reading comprehension level. The researchers analyzed the results and imposed unique learning strategies designed to meet the needs of the students. Kozulin and Garb (2004) found that their Dynamic Assessment procedures significantly impacted posttest scores. The researchers concluded with a claim: Dynamic Assessment is effective for assisting

teachers to develop intervention strategies and is more reliable than traditional methods in helping students meet their potential in the L2 classroom (p. 75).

Dynamic research continued in the L2 classroom with two researchers who have had more impact on Dynamic Assessment knowledge than any others, individually and as a team, in modern time: Lantolf and Poehner. They produced an extensive set of publications, including but not limited to the following: Lantolf (2009), Lantolf and Poehner (2004), Lantolf and Poehner (2006), Lantolf and Poehner (2010), Lantolf and Poehner (2011), Poehner (2005), Poehner (2007), Poehner (2008), Poehner (2009), Poehner and Lantolf (2010), and Poehner and van Compernolle (2011). They prescribe to the theory that mediation is an integral part of the process of assessment in the foreign language classroom and assists with learner development. Their most recent endeavor is Project 2010-2014: Dynamic Assessment in Foreign Language Education. Through Dynamic Assessment research in the L2 classroom, Poehner and Lantolf (2010) compiled data to compose a comprehensive casebook. They have also written a guide for foreign language teachers on best practices, focusing on the use of Dynamic Assessment in the classroom. One of their main components is mediation, from which emerges "cooperative dialoguing between the mediator and the learner" (Lantolf & Poehner, 2006, p. 46). L2 teachers could make use of mediation techniques in their classrooms so that the students could formulate test-taking strategies specific to foreign language testing and so that the classroom teacher could discern learning problems.

With the purpose of many traditional assessments being questioned in education circles, Poehner and van Compernolle (2011) argued that Dynamic Assessment methods provided teachers with the tools to help their students' progress. To give credence to their

premise, the researchers video-recorded Dynamic Assessment sessions in which the mediator provided differing levels of support based on the needs of the student taking the test at the time. The current study made use of a pre- and posttest method but did not compare them, as many other Dynamic Assessment studies had, after taking under consideration Poehner and van Compernolle's argument that Dynamic Assessment should be a starting point and not necessarily a posttest intervention. Poehner and van Compernolle concluded that Dynamic Assessment would meet learners' needs and would provide a collaborative and cooperative framework. Poehner and van Compernolle's work provides an option for foreign language teachers who choose not to use pretests but would still like to use Dynamic Assessment techniques as part of their curriculum.

Numerous other researchers have added to the body of knowledge concerning Dynamic Assessment. Adair-Hauck, Glisan, Koda, Swender, and Sandrock (2006) noted recent dramatic changes in the foreign language classroom, including a focus on communication and performance instead of pure grammar drill. However, they also pointed out that while instruction and methods have rapidly evolved, assessments have remained fairly traditional, causing disconnect between instruction and assessment. Adair-Hauck et al. named Dynamic Assessment as one method available to teachers that could potentially close the gap between the new standards and practices in foreign language education and assessments. The lynch-pin in Dynamic Assessment, though, is the teacher. Haywood and Lidz (2007) elaborated: "much of the interpretation of DA data depends on the skill and experience of the examiner" (p. 3). In their study, Haywood and Lidz emphasized the role of the teacher in mediating student growth through purposeful interventions. Today's L2 educators may find that mediating growth is an expected role

with the educational trend toward growth model, so Dynamic Assessment could lend itself to more use.

Ableeva (2008) continued to promote the use of Dynamic Assessment in the French classroom. Ableeva's data showed that listening ability increased over time due to mediation. Additionally, learners were more likely to be able to transfer their knowledge to complex activities after experiencing Dynamic Assessment interventions. Hill and Sabet (2009) switched their focus from listening comprehension to speaking activities, as they related to Dynamic Assessment. Hill and Sabet emphasized the developing cognitive process in their study, as well as the positive relationship formed between the teacher-mediator and the student-learner. These researchers validated the use of Dynamic Assessment for L2 speaking activities as a "development-oriented process of collaborative engagement…[that] reveals the underlying causes of performance problems and helps learners overcome them" (p. 537). In the same year, Antón (2009) chose Dynamic Assessment as a focus, instead of narrowly concentrating on listening or speaking. Antón echoed Ableeva and Hill and Sabet's urge for teachers to take an active role in mediating strategies with students. Antón stressed that teachers need to "create a symbiotic relationship between teaching and assessment as learners' progress is assessed through the program in combination with multiple opportunities for learning" (p. 578). This research also indicated that Dynamic Assessment is not a replacement for assessments that may already be in place; rather, it is a complement. Antón concluded that Dynamic Assessment led to greater understanding of each student by the instructor that resulted in more effective interventions and a better chance for student success.

Following the same line of thought, Mardani and Tavakoli (2011) questioned why teachers assess their students; they determined a problem with what they discovered. Many teachers assess only as a tool for gathering information and giving a grade instead of for the development of the students. Mardani and Tavakoli divided the students into control and experimental groups; the experimental group students received interventions based on their pretest scores. These sessions allowed for discussion and strategizing so that students felt better prepared. The researchers emphasized the relaxed and cooperative atmosphere of the Dynamic Assessments. Following the posttest, Mardani and Tavakoli found an increase for both groups; however, the experimental group significantly outperformed the control group. Mardani and Tavakoli concluded that Dynamic Assessment could give learners confidence and should be a part of foreign language assessment processes.

Wei (2011) noted a world-wide focus on assessment models and decided to gain the perceptions of her students as she implemented Formative Assessment into her English classroom. Wei observed, interviewed, and surveyed her 227 students over the course of three semesters at Hechi University in China. The researcher described a process for Formative Assessment implementation: discovering student needs, setting goals, finalizing a plan for assessment and explaining it to the students, implementing the plan, and evaluating and revising plans during the semester. Wei emphasized quality planning and excellent student-teacher communication while feedback should be specific, positive and critical, descriptive, but never evaluative because the purpose is student improvement. Wei concluded that Formative Assessment is motivating, encourages independent learning, and positively affects learning. Teachers of foreign language could

strongly consider Wei's research that emphasized planning and positive, productive student interaction.

Sadeghi and Khanahmadi (2011) determined a need to supplement the traditional assessment methods found in English as a Foreign Language courses in Khoy, Iran, with Dynamic Assessment methods. The researchers chose 60 students from a larger pool after first assessing minimal competency requirements with the Preliminary English Test. The researchers administered a pretest and followed with typical classroom instruction over the course of one semester. Sadeghi and Khanahmadi randomly divided the students into a control group and an experimental group. After each quiz during the course, the researchers brainstormed with the experimental group various strategies for conquering different question types and for ways in which to apply information from one situation to another. Students took a posttest to complete this quasi-experimental study. Sadeghi and Khanahmadi found significantly higher scores in the experimental group. The researchers concluded that Dynamic Assessment is a meaningful assessment method that produces benefits for the students. As L2 teachers employ diverse methods to meet the needs of their students, Sadeghi and Khanahmadi's suggestions for providing strategy sessions could be quite applicable.

Given the body of research that promotes Dynamic Assessment as effective in promoting student growth, this research would not have been complete without Dynamic Assessment as a choice.

Task-based Assessment

A Task-based Assessment, according to Adair-Hauck et al. (2006), is effective for "assessing students' progress in meeting the standards since they require goal-directed

use of language, use of multiple skills or modes of communication, and integration of content" (p. 361). Disagreement exists among the researchers in defining the word *task*. While it seems a simple word, a function to be performed, some researchers have modified the definition slightly to fit their own purposes. Long (1985) defined the word in a basic sense as something done for oneself or someone else, whether it be for a reward or not. Richards, Platt, and Weber (1985) added a twist when they indicated that a task is some activity or action through the processing of language. This definition is fairly broad and could apply to many situations in the L2 classroom. Most L2 teachers instruct their students through activities in which they are processing the language but would not consider their instruction as task-based.

Prabhu (1987) changed this definition slightly to fit the educational setting by adding that the activity must lead students to arrive at an outcome, and the process must be controlled by a teacher. L2 teachers could benefit from this definition since most teachers are writing outcomes for their classes. The task-based activities in the classroom are the process through which the outcome is accomplished. Nunan (1989) believed that a task is focused on meaning, instead of on grammar and form, as students interact in the target language. Teachers using the communicative language method in which communication trumps grammatical accuracy, would possibly choose Nunan's view of Task-based Assessment. Skehan (1996) added to the definition that the task must have real world application or use and be communicative in nature while focusing on the outcome. Skehan also divided task activities as being either weak, used as an integral part of the L2 classroom, or strong, used as the primary teaching and assessment method.

With the implementation of the CCSS (National Governors Association Center for Best Practices, & Council of Chief State School Officers, 2010), task-based activities may align well for L2 educators, using Skehan's interpretation. The CCSS encourages authentic activities and is outcome based. Lee (2000) came up with a complex list of criteria for what defines a task: it must be in a classroom, have interaction that is structured by the teacher, be focused on communication or meaning, and must eventually produce work in the target language. A more usable and agreeable definition for use by L2 teachers comes from Bygate, et al. (2001): "A task is an activity which requires learners to use language, with emphasis on meaning, to attain an objective, and which is chosen so that it is most likely to provide information for learners which will help them evaluate their own learning" (p. 11). Ellis (2003) echoed Bygate, et al. and added that tasks should resemble the real world and may be oral or written. The most common thread found in definitions of task for the L2 classroom is that the task must focus on meaning and be tied to an outcome. The CCSS provides teachers with criteria that promote literacy through comprehension and collaboration. Task-based learning may assist the L2 teacher in making assessment choices that support the CCSS.

Task-based Assessment had its inception in the 1980's with the research efforts of Krashen (1981; 1982; 1985), Long (1985), Prabhu (1987), and Nunan (1989). Considered the front-runner for Task-based Assessment, Krashen (1982), for example, created a list of five hypotheses on which his theory of language acquisition is based. The Acquisition-Learning hypothesis lays the groundwork for all of the other hypotheses. L2 performance is acquired from two, independent systems, one that is learned and one that is acquired. Acquisition, according to Schütz (2007), is "the product of a subconscious process very

similar to the process children undergo when they acquire their first language. It requires meaningful interaction in the target language…in the communicative act" (para. 4). On the other hand, learning comes from formal instruction and is a conscious, cognitive process. Krashen (1982) valued the subconscious acquisition process over the conscious learning process. The Monitor hypothesis shows the relationship between the two independent systems and delineates the value of acquisition over learning. Krashen (1982) called the conscious process the monitor. Students must think about grammar rules and focus on form to address language problems. Monitoring is the practical part of language acquisition. However, without initial, meaningful interactions in the target language, learners would not have the building blocks already in place to monitor. Learners use the monitoring system in different ways. Some learners overuse it and focus on grammar rules; others underuse it, while others optimize their use of conscious knowledge. Krashen (1982) emphasized that the monitor's function is minor; its role is to polish communication and make it appear more native. While Krashen's (1982) other hypotheses, Natural Order, Input, and Affective Filter, play crucial roles in his theory and for language study, the Acquisition-learning hypothesis and the Monitor hypothesis relate to the current study.

Other researchers have concurred with Krashen's (1982) theory that meaningful communication took precedence over conscious grammar instruction. For example, Long (1985) purported that using real-world tasks in an educational setting would create a more productive learning environment and increase L2 acquisition. Long's (1996) major contribution to the field of language teaching was the Interaction hypothesis, in which face-to-face communication is promoted as one of the crucial components for second

language acquisition. The current study promoted student face-to-face conferencing, as suggested by Long. Additionally, Long (1997) offered his concept of forms, meaning, and form. A teacher focused on forms would be dedicated to vocabulary and grammatical structures with verb drilling. One who emphasized meaning would almost completely rely on communication for L2 acquisition. Long determined that a focus on form was best, in that students could be directed to linguistic elements when necessary with the attention on meaning. Long stated that "Focus on meaning alone (a) is insufficient to achieve full native-like competence, and (b) can be improved upon, in terms of both rate and ultimate attainment, by periodic attention to language as object" (para. 1). Focus on form is a major tenet of task-based learning. L2 teachers should consider Long's view that effective communication is hindered when sufficient grammatical and vocabulary structures are not in place. To complement his focus on form, Long introduced the Interaction hypothesis: face-to-face classroom interaction is not only valued but also crucial for paramount L2 acquisition. Long balanced form with meaning so that students can communicate effectively.

Similar to Long's (1985) studies, Prabhu (1987) published findings from the Bangalore Project out of southern India. Büyükkarci (2009) credited Prabhu with being the developer of Task-based Assessment in the modern educational realm due to this research. The premise behind task-based learning was that students would more readily acquire a second language when they were immersed in a task instead of on the language itself. Prabhu believed that the best way to assess a task was to complete a pre-task first. He detailed a system for this as follows: The teacher would complete a whole-class task while including the students in the process. The task should be similar to, but not the

same as, the final assessment task. This would serve as preparation for the individual students in a comfortable setting. The teacher talked through ever step and emphasized the outcome. Students asked questions and gave other reasonable ways in which to complete the task to assist different levels of learners or learners with different learning styles. The pre-task was not just a demonstration but served as a way to mediate for the students through conversation. Vygotsky's (1978) sociocultural theory again surfaces as the rationale for the teacher-student interaction. The teacher is the expert without whom the students would not experience growth at the same rate or at all.

Later in the 80's, Nunan (1989) added to the body of information about task-based learning and assessments. He asserted that what is taught should be fused with how the subject is taught. Again, banking on Vygotsky's (1978) sociocultural theory, Nunan encouraged teachers to strategize with their students. He wanted students to be comfortable with uncertainty and have ways to cope with new situations. Nunan helped students make intelligent guesses based on past knowledge and to develop critical thinking skills. Nunan's goal was to empower students into becoming more effective through being "adaptable, creative, inventive, and above all independent" (p. 81). Nunan's ideas should be an inspiration for L2 educators.

It was not until the late 90's that task-based learning and assessment caught fire. As a team, Foster and Skehan (1996; 1997; 1999) and Skehan and Foster (1997; 1999) as well as individually (Foster, 1999; Skehan, 1996; 1998) served to assist teachers in implementing task-based instruction. Skehan and Foster (1997) realized that foreign language proficiency developed through completing meaningful language tasks instead of through drills and contrived practice. To support their realization, the researchers

conducted task-based activities research in the foreign language classroom. Skehan and Foster (1997) analyzed the fluency, accuracy, and complexity of language when students had time to prepare and in relation to the difficulty level of the tasks. Examples of tasks that L2 educators could possibly use required an exchange of personal information, composing a narrative about a cartoon strip, and giving advice to subjects writing personal problem letters. Skehan and Foster (1997) found that, in most cases, students needed sufficient planning time to fluently and accurately perform. The researchers encouraged teachers to choose tasks carefully, to allow adequate planning time, and to limit the areas of focus for assessment purposes.

Continuing with the trend toward communicative approaches in the L2 classroom, Ellis (1997) is known for tying second language acquisition research with Task-based Assessment, specifically the Input-Interactive-Output model (Block, 2004). Ellis (2000) related that "…what is important for acquisition is the opportunity for learners to engage in meaning negotiation" (p. 199). L2 teachers could appreciate Ellis' perspective. His resolve for authentic conversations in the classroom directly correlates to the Common Core State Standards that strive to prepare students for the real world and work force (Heining-Boynton & Redmond, 2013). Echoing Ellis, Johnson (2000) detailed task-designs in a practical manner for educator use while Kim (2002) promoted task-based methods as an increasingly popular tool for second language acquisition. Expanding on his initial research, Ellis (2003) related that any tasks must specify what has to be done; promote meaning over structure; be as authentic as possible; involve one or more of the four skills of listening, reading, speaking, and writing; incorporate a cognitive process; and have a clear objective or outcome. Ellis (2006) provided additional advice for

teachers with practical principles for creating opportunities to develop second language competency through the collaborative process. He directed teachers to find appropriately leveled tasks that meet clear goals, to communicate those goals with the students, to make students an active part and encourage them to take risks. Also, Ellis added that tasks must stay focused on meaning over structure while still providing some opportunities for instruction about grammar and form. Ellis provided great information for teachers interested in designing Task-based Assessments.

Byrnes (2002) completed a qualitative case study with the Georgetown University German Department over the course of three years. She determined a problem concerning the inadequacy of current foreign language assessments, revealing that their assessments did not accurately evaluate students. She added that inadequate evaluations actually deterred learning and student growth. Byrnes surmised that task-based writing assessments would create a common knowledge base for the students and would serve the faculty as a guide for curriculum and learning goals. The purpose of the research was to promote a shift from grammar-based instruction to language use and meaning instruction. Byrnes concluded that Task-based Assessments more accurately reflected student achievement, as perceived from instructor and student observations and interviews. Byrnes also determined that Task-based Assessments helped to bridge the gap between content knowledge and workable language use. While this assessment restructuring promoted unity within the Georgetown German Department, restructuring also created a learner-centered environment. Students reported greater enjoyment of the class and more creativity and choices within the class due to the contextualized, Task-based Assessments. All of Byrnes findings may be practical as L2 teachers are

redesigning their methodology and curriculum to align with the Common Core State Standards and changing global demands.

Opponents of task-based learning continued to criticize its apparent lack of teaching structure and its neglect of the cognitive process in language acquisition. Richards and Rogers (2001) stated that task-based learning was a great idea, as an idea only; they did not believe that this method was practical or applicable for the classroom and was simply an ideology. While Swan (2005) agreed to some of the merits of task-based learning, he claimed that using this method was too advanced for most teachers and too complex to consistently practice over an entire course. Swan also criticized the noticing aspect of task-based learning, in which learners are responsible for noticing vocabulary and structure, mainly on their own, to form the language. Nassaji and Fotos (2011) noted that most criticisms of task-based models are grounded in outdated techniques, mainly the grammar-translation method. These researchers provided evidence of a balanced and updated approach that has a "focus on grammar through interactional feedback, focus on grammar through structured grammar-focused tasks, and focus on grammar through collaborative output tasks" (para. 4). Nassaji and Fotos provided examples for today's educators of ways to explicitly teach grammar through a more communicative approach. Klapper (2003) acknowledged the criticisms, but provided solutions similar to those of Long (1997) in which a mixed approach may be more beneficial to the students. Klapper supported the task-based approach through research that conventional language approaches have not been found as effective. Klapper suggested the use of Ellis' (2006) focus on form with guided practice to complement the main method using communication and tasks.

Littlewood (2004) also continued Ellis' (2003) discussion on forms, meaning, and form. He suggested that the teacher does not have complete control over the communication levels in the classroom. Rather, the students do. He described that a teacher may execute a communicative teaching, task-based plan, but some students may choose to focus on grammar and structure and not really interact in the target language. Littlewood then criticized the use of the word *task* since this word has a negative connotation and would not stimulate enthusiasm toward learning. Littlewood would like to remove all labels but understands that labels help educators feel comfortable and allows for common language within the field. However, he would like to change the name of this method, task-based, to something more user-friendly. Littlewood suggested

> the main common denominator of communicative and task-based approaches in their various forms is that, even when they use form-focused procedures, they are always oriented towards communication, my own preferred working label (to cover both communicative and task-based language teaching) is communication-oriented language teaching. (p. 325-326)

Littlewood (2007) also desired to provide practical information for educators on how to implement task-based approaches more effectively. He compiled ideas and experiences from educators in order to assist those in the field in creating a learner centered environment. Littlewood concluded that teachers must move away from lecture-type settings toward active, independent learner settings.

As more teachers became aware of task-based methods, Jeon and Hahn (2006) realized that few actually used the method. Jeon and Hahn proposed three research questions to determine how well teachers understood the Task-based language teaching

(TBLT), teachers' perspectives on TBLT implementation, and the reasons teachers chose not to use TBLT in their classrooms. Jeon and Hahn found that most of the teachers understood TBLT and its purpose, agreeing in the method's relevance in relation to communicative teaching, yet a significant number did not implement TBLT anyway. Some reasons cited for not using TBLT included extensive preparation time, psychological burden on the teacher, lack of confidence as facilitators, and difficulty in assessing students. However, teachers who reported using TBLT revealed the following benefits: group work productivity and interactions, improved learner motivation, and better academic progress. Jeon and Hahn concluded that TBLT should be promoted, and the teachers who implement it should have specific training on TBLT methods and classroom management.

Carless (2007) found similar results to those of Jeon and Hahn (2006). Carless found that secondary school teachers held wide ranges of opinions about task-based learning while university instructors were more interested in task-based approaches. The interview process provided the following concerns about using task-based learning: loss of classroom control, more off-task time, inadequate time to prepare for this method, priority of competing text book, and weak grammar instruction. Carless allowed for discussion during the interviews; the educators came to a consensus that they would support a weak version of using task-based learning that was shared with continued traditional models. Carless concluded that task-based approaches could be beneficial but, for most educators, would need to be balanced with traditional methods. He confirmed that using a weakened version of task-based learning would be more advantageous than using traditional methods by themselves.

While popular among researchers, Task-based Assessment lacked mainstream support in the classroom, so additional research began with a caution that teachers take greater care in task choices (Robinson, 2001, 2007; Robinson & Gilabert, 2007). Robinson determined that task complexity, difficulty, and production levels all interrelated. In another study, Robinson (2003) desired to show the correlation between the cognitive processes and Task-based Assessments as they relate to learner development and performance. Robinson contrasted the past research with his ideas:

> In relating task-based pedagogy to acquisition processes some have argued that the meaningful language exposure that task work makes available to learners enables unconscious "acquisition" processes to operate successfully on the comprehensible input tasks can provide: language production, and attention to form, are of much less, if any, importance. In contrast, the proposal made here is that task-based learning, sequenced according to the criteria I describe and others like them, leads to progressively greater attention to, "noticing", and elaborative processing and retention of input. (p. 47)

Robinson seemed to strike a compromise between the traditional methods and the communicative teaching method here as he called for a blend of the modes. For the purpose of this study, both meaning and structure were considered. Agreeing with Robinson, Tulung (2008) added that "modifications [making note of grammar] during interaction are used primarily to make language comprehensible so that communication can take place" (p.110).

Willis and Willis (2007) published a procedural book for educators on task-based instruction. While the book is not full of research findings, it is complete with activities

and design ideas along with educator experiences with task-based methods. Norris (2009) provided the research with his suggestions for successful task-based teaching. Norris promoted learning by doing, through tasks. Norris noted, "Key is the idea that holistic activity structures, such as tasks, offer an ideal frame within which knowledge use can be experienced and understood, and from which learning opportunities should be developed" (p. 579). L2 teachers who desire to restructure their entire curriculum with task-based teaching and assessments would benefit from Norris' ideas. He posits that the assessments should mirror the teaching methods. Additionally, Norris offered rationale for using Task-based Assessments:

> Assessment…emphasizes the performance of target tasks (as opposed to the demonstration of knowledge about the language), primarily as a mechanism for providing meaningful feedback to learners and teachers, for determining students' abilities with target tasks, and for ensuring an overall focus on target-task learning throughout the program. (p. 582)

Task-based Assessment can be used in the L2 classroom. Norris' idea, that Task-based Assessment would provide better feedback so that the students' needs could be addressed and met, makes sense to educators. As today's teachers are struggling to address all learners' needs, assessments that can potentially identify those needs and track that learning targets are met are ideal and essential.

Sugita (2009) realized the shift from traditional, accuracy-focused writing to communicative, fluency-oriented writing; however the assessment methods for evaluating writing had not changed. Sugita proposed to establish a framework for and to develop assessment of task-based writing. He also sought to examine the reliability and

validity of his methods and rating scales. Sugita found that the teachers were individually consistent when using his 5-point scale for students' writing ability. Jabbarpoor (2011) addressed the case for task-based methods in the L2 classroom after its surge. L2 teachers should first realize the shift that has occurred in foreign language teaching that focuses on authentic communication and meaning. Both Sugita and Jabbarpoor emphasized that teachers must continue to update their assessments to meet the current needs.

Sidek (2012) believed that teaching foreign language with traditional approaches was not necessarily effective or predictive of competence in the language either. Sidek established that reading proficiency had more impact on continued foreign language study than anything else. Sidek purposed to connect Communicative Task-Based Language Teaching (CTBLT) with reading instruction in Malaysia with his historical research. Sidek found that the majority of reading tasks fell into the cognitive processing theory and were not socio-cultural. Sidek did find that most instructional approaches used a form of task-based instruction; however, he also found that schools lacked interaction, an essential part of true communicative approaches.

A host of other researchers have joined the task-based learning team and have produced their own versions of what may be effective for educators and students today, many including task-based learning in connection with technology (Ahmadian, 2012; Butler, 2011; Downing, 2012; East, 2012; Hashemia, Azizinexhadb, & Darvishi, 2012; Hismanoglu & Hismanoglu, 2011; Meurers, Ott, & Ziai, 2010; Tavakoli & Foster, 2011). However, while the term *task* may be debated and have slightly different definitions, one thing remains, the concept. Having students use communication to perform an authentic

task has been shown effective and beneficial for L2 acquisition and was a viable choice for this study.

Formative Assessment using Self- and Peer-Evaluations

> Formative Assessment's aim:
>
> is to monitor student learning to provide ongoing feedback that can be used by instructors to improve their teaching and by students to improve their learning. More specifically, Formative Assessments: help students identify their strengths and weaknesses and target areas that need work; help faculty recognize where students are struggling and address problems immediately. (Eberly Center for Teaching Excellence, n. d., para. 1).

As the L2 students progress in knowledge and language ability, the need to work on writing skills in the target language becomes apparent. As they begin to write, it is essential that they receive proper, positive feedback so that they will improve and remain encouraged to continue writing. Tamjid and Birjandi (2011) acknowledged that writing is especially difficult in the foreign language setting and is one of the most difficult tasks for the teacher to conquer. However, it is not simply the responsibility of the teacher to train writing skills; students need to take an active role to be successful. Using self- and peer-evaluations to monitor writing assisted the teacher who did not have time to grade several drafts of a paper. Beyond the grading aspect, students benefitted from the chance to try writing in Spanish, with the opportunity to correct their work, before being formally graded by the teacher. Hanrahan and Isaacs (2001) noted that self- and peer-assessments contributed to effective learning in the L2 classroom and that students reported the activity as a benefit. National Capital Language Resource Center (n.d.)

defined self-assessment as students "engage[d] in deliberate thought about what they are learning…In this kind of reflection, students step back from the learning process to think about their language learning strategies and their progress as language learners" (para. 5). Self-assessment has the benefit of being motivational to the students.

Oscarson (1978) and Holec (1979) completed groundbreaking studies for L2 teachers relating to self-assessment. Oscarson surveyed self-assessment practices used in multiple subject areas all around the world. He found that self-assessment, though not common practice, was being used by some institutions. The most typical type allowed students to check their work on a given set of objective questions by placing a grid over their work to see if their answers matched the key. The main purpose for self-assessment at this point in time was for students to discern their level or placement on a practice exam. From the practice results, students could estimate their score for an upcoming assessment and then study accordingly. Oscarson concluded that much more research was needed but also that students benefitted from self-assessment because of its practicality for goal-setting. Holec more boldly determined that self-assessment practices allowed more learner autonomy through defining individual objectives, monitoring self-progress, and adjusting methods throughout a course. Holec believed that self-assessment created a personal and subjective, almost customized, approach to education. Holec coined the term *learner autonomy*. Little (1991) described learner autonomy as learning in which students acknowledge responsibility for their own learning process. Schunk (2005) added to Little's description through his belief that learner autonomy also included student motivation and goal evaluation throughout the process.

Throughout the 80's and 90's, several researchers continued their studies about the benefits of self-assessment, including Blanche (1988), Blanche and Merino (1989), and Blue (1994), and Oscarson (1984; 1998). Oscarson (1984) provided practical cards that teachers could distribute to students in their quest for self-assessment tools. These cards simply listed abilities and then had initial boxes for the student and teacher with each. In this way, the learner knew the goals set by the teacher and was accountable to attaining those goals. Oscarson (1998) attributed six benefits to this system and others similar to it: promotion of learning; increased levels of learner awareness; motivation toward goals; broader experience with varied techniques; shared responsibility for learning; and post-educational benefits. Second language teachers could potentially use these ability cards to track their students' competencies, similar to a Formative Assessment tracking aid.

Blanche (1988) reviewed the available literature and compiled a list of conclusions that he had drawn. Blanche noted that students' abilities to self-assess are highly dependent upon their linguistic abilities. Additionally, students may not accurately self-assess due to language errors and gaps in learning. On the positive side, learners who use self-assessment tend to be more motivated and can assess their ability to purely communicate very well. Blanche and Merino (1989) contributed to the field of self-assessment information through their research about accurate self-reflection. A major part of learner autonomy is being able to accurately gauge one's own progress, abilities, and performance. Blanche and Merino formulated their thinking around Krashen's (1982) monitor and model theory. The student, as the monitor, highly affected the ability to self-regulate along with the quality of the self-evaluation; the student monitor had to accept

the teacher as his/her mentor or learning was negatively impacted. Blanche and Merino tied past theory with current practice to support self-assessment in the L2 classroom. Concurring with them, Blue (1994) stated that

> evidence was found to support the belief that students who do assess their language level realistically may persevere with language learning to a greater extent than those whose assessment is unrealistically high or low. Both performance and progress must be monitored by teachers, with constant feedback, for students to be able to realistically self-assess their progress in language learning. (p. 18)

Blue continued with the advantages of self-assessment for the students. He believed that students need to evaluate their level of effort so that effort may be increased in the future. Blue also stated that students, through self-evaluation, begin to value their own capabilities, consequently building self-confidence. Another asset for including self-evaluation activities with L2 writing assignments was student realization that competence in an area did not always equate with the performance given. This realization helped students to see their own strengths and weaknesses and compensate accordingly. Knowledge of self-evaluation techniques could guide L2 teachers as they strive to improve student performance.

While self-assessment became a useful tool in the L2 classroom, peer-assessment had very little appeal. Peer-assessment entails the following: “Students internalize the characteristics of quality work [through] the work of their peers. However, if they are to offer helpful feedback, students must have a clear understanding of what they are to look for in their peers' work” (National Capital Language Resource Center, n. d. para. 1). For

peer-assessment to have positive effects, students must trust their own abilities to appraise work, as well as the abilities of their classmates. Hedgcock and Lefkowitz (1992) completed a study in which certain students met in small groups to assess each other's French writing assignments. The experimental group who received this intervention completed the writing assignment with substantially higher scores than those in the control group who only received some teacher feedback during the process. Hedgcock and Lefkowitz concluded that collaborative efforts in the foreign language writing process assisted learners with self-awareness which led to better writing, as well as a lack of inhibition in the future when writing assignments were given.

Black and Wiliam (1998a) advanced that "improvement in classroom assessment will make a strong contribution to the improvement of learning (p. 7). In an overview of many different types of assessments for the classroom, Black and Wiliam encouraged Formative Assessment as a best practice. They acknowledged, though, that peer-assessment benefitted lower-achieving learners more than higher-achieving learners; however, higher achievers were more productive overall and did not require as much external assistance. Brown and Hudson (1998) agreed that Formative Assessments, specifically but not only self- and peer-evaluations, provided teachers with effective assessment choices in the foreign language classroom. Black and Wiliam (1998b) provided information to inform and assist classroom teachers. Primarily, they wanted to emphasize that Formative Assessment raised standards but that teachers need to re-evaluate their assessment methods to ensure that their assessments actually promoted learning, were not overly competitive, and provided positive feedback. One of the suggested assessment methods was self- and peer-evaluation in which "pupils can assess

themselves only when they have a sufficiently clear picture of the targets that their learning is meant to attain" (p. 145). Black and Wiliam emphasized clear communication of the learning target to the students. When students understood what teachers wanted, they were more likely to obtain the goal.

Black, Harrison, Lee, Marshall, and Wiliam (2004) provided updated findings in relation to self- and peer-evaluation:

> Students can achieve a learning goal only if they understand that goal and can assess what they need to do to reach it. So self-assessment is essential to learning…peer assessment turns out to be an important complement to self-assessment. Peer assessment is uniquely valuable because students may accept criticisms of their work from one another that they would not take seriously if the remarks were offered by a teacher. (p. 14)

Clearly, Black et al. added that teachers must take active roles in training low-achieving learners how to self- and peer-assess. This training process must begin early and become successful as students become more adept at what and how to evaluate themselves and their peers. By using self- and peer-evaluations properly, teachers have more time to facilitate instruction at the time and use feedback to inform teaching practices in the future. Lee (1998) echoed by stating that learners themselves must take responsibility for the known goals by monitoring and assessing actively, and Ross (1998), provided mixed-result data concerning the value of self-assessment in the L2 classroom. In a study in which learners evaluated themselves on recently completed course material and teachers assessed their students at the same time, Ross found statistically significant discrepancies. The researcher attributed the differences between student and teacher scores to the

learners' language experience and skills. Ross reported that self-assessment was a valuable alternative assessment tool, but that accuracy of the information from oneself was too dependent upon the individual learners.

Another pioneer in the field of self-assessment was Pintrich (2000) who correlated motivation with self-regulation. He believed that students who self-assess develop positive self-reflection because they know that their successes came primarily from their own efforts and attributed struggles with ineffectual efforts. Schunk (2005) noted Pintrich's view as a benefit to student growth: "Students who adopt mastery goals are more likely to report monitoring and attempting to control their cognition with various learning and cognitive strategies, and to seek ways to increase their awareness of their understanding and learning" (p. 89). Pintrich developed the conceptual framework that is used for self-assessment. He connected the area of cognition with phase of forethought, planning, and activation; motivation with monitoring; behavior with control; and context with reaction and reflection. Pintrich set the stage for future use of self-assessment by clarifying its benefits and providing a framework for the field.

Numerous additional researchers completed studies in the field of foreign language self- and peer-assessment, attributing its definitive use based on data and its merits for aiding the learner in meeting learning goals (Alonso Alonso & Palacios Martínez, 2005; Boud, 2007; Boud & Lee, 2005; Coronado-Aliegro, 2006; Deakin-Crick, Lawson, Sebba, Harlen, & Yu, 2005; Little, 2005; Sebba et al., 2008). This researcher supports the use of self- and peer-assessments as a beneficial component of the learning process; thus, incorporating it into the current study. However, one argument criticizing the use of peer-assessment, as discussed by Liu and Carless (2006), was resistance from

students and educators who opposed peer-assessment as part of a student's grade. Some feared that social influences or lack of grading experience would create an unjust system by which students were evaluated. Liu and Carless countered this argument by promoting the inclusion of very specific grading criteria and controls for the peer evaluator. The researchers also encouraged teachers to frequently include peer-assessment throughout the course so that students are comfortable with the process and more effective with their assessments. Various other important studies from Bryant and Carless (2009), Cartney (2010), Colby-Kelly and Turner (2007), Hanrahan and Isaacs (2001), Lindblom-Ylänne, Pihlajamäki, & Kotkas (2006), Ross (2005), Tamjid and Birjandi (2011), and Yang (2011) in recent years have added to the consensus that self- and peer-assessment should be essential components of an L2 classroom's alternative assessment repertoire. Each study allows for L2 teachers to gain insights from current research that may be applied to their own classrooms.

Hanrahan and Isaacs (2001) saw an increasing need for greater student participation in writing classes. The researchers completed a qualitative study with 233 third-year college students for the purpose of gaining student perceptions about peer- and self-assessment. The professors assigned a research essay; a portion of the grade would come from completing a peer- and self-assessment. Students also provided copies for their essays to tutors who submitted feedback. The instructors encouraged the students to revise their essays after internalizing the three sets of feedback. At the end of the course, the students completed an open-ended questionnaire about the positive and negative aspects of the peer- and self-assessment system. Hanrahan and Isaacs grouped the comments into eight different categories. Some of the criticisms included task difficulty

because of lack of experience, discomfort because of the critical aspect, and time consumption. Positive themes included motivation, productive activity, and better understanding of expectations. Overall, the researchers found that students perceived the assessment process as beneficial. Hanrahan and Isaacs concluded that many of their problems would have been avoided if they had provided clearer standards and training with modeling of assessment at the beginning of the course.

Ross (2005) indicated that an ideological trend toward considering alternative assessments in foreign language classrooms prompted him to complete a longitudinal study of eight cohorts of Japanese university students. Ross wanted to know if Formative Assessments were less reliable than conventional assessments. He also sought to determine if Formative Assessments increased the students' proficiency rate and end result. Formative Assessments included peer- and self-assessment, as well as projects and portfolios. Ross found that Formative Assessments were substantive in relation to achievement and growth; however, the effects were mainly noticed in the area of listening comprehension. Reading comprehension scores did not vary much from cohort to cohort, regardless of the assessment method. Ross concluded that the primary benefits of Formative Assessments were better student attention and higher participation levels in class, as well as improved listening skills in the target language.

Lindblom-Ylänne, et al. (2006) realized that education had shifted from a focus on factual knowledge to higher-order thinking and subject competency. In order to meet the needs of today's learners, the researchers decided to promote student engagement in the learning process by implementing a peer- and self-assessment system. Each student used a grading rubric to score his/her own essay. Then the instructor and another student

were given a copy of the essay, without his/her name to complete the rubric. After all grading was completed, the three rubrics were combined to find the mean, which became the student's score. Students were not informed of this factor to avoid inflated grading. Lindblom-Ylänne et al. found that the overall quality of all essays was good and that all participants positively reviewed this process. The researchers also found few discrepancies among the three scores when scoring technical merit, use of literature, and format. Lindblom-Ylänne et al. saw differences in rating critical thinking, with peers grading much easier and the instructors more harshly. Students felt that assessing their peers' essays helped them by seeing different writing styles and ways to approach subjects. This study contradicts previous ones in which peer- and self-assessments were much higher than those of the instructor.

Colby-Kelly and Turner (2007) saw a trend toward using Formative Assessments in Math and Science classrooms and were curious as to whether this trend would or would not be useful in foreign language classrooms. The researchers sought to determine student and teacher perceptions of Formative Assessments, the benefits of Formative Assessments, and the nature of Formative Assessments. The teachers readily agreed that using peer assessment as a formative measure promoted learning; however, the researchers discovered that teachers hesitantly used this measure and remained traditional with teacher-student feedback. The researchers desired continued use of Formative Assessments; however, they would desire expanded use of peer-assessments and peer-feedback.

Bryant and Carless (2009) determined a need for better foreign language classroom assessment. They hypothesized that self- and peer-assessments would

stimulate more learner independence and create an atmosphere in which students desired improvement through reflective thinking. Students were actively encouraged to consider the assessment rubric when initially completing assignments and during the revision process. Students viewed the rubric as a tool to help them earn better grades instead of a waste of time. The teachers reported that students were initially pleased with the process and actively engaged; however, once acclimated to the process, students complained of boredom. Many students trivialized their comments or simply checked the *good* box. Other students did not want to offend their friends, so they did not give honest evaluations. Some advanced students complained that the weaker students were not capable of providing accurate and useful feedback. Toward the end of the study, when students were preparing for their end-of-the-year assessment, teachers encouraged the learners to revisit their evaluation rubrics in order to anticipate areas that needed addressing. Students saw value in this activity and were freshly motivated. Overall, both teachers and students viewed the assessment process as valuable and felt that the process did encourage personal responsibility. Bryant and Carless determined that Formative Assessments, such as those in this study, were best combined with a summative assessment for maximum effectiveness and should be used to reduce the dominance of the summative assessment by itself.

Cartney (2010) reviewed literature relating learning with assessment being involved in curriculum restructuring that required the use of Formative Assessment at his university. The researcher conducted a case study using an action research design that incorporated peer-assessment for the purpose of improving learning. Cartney held a workshop with his students to accomplish the following: alleviate stress, explain the

process, provide copies of the marking sheets, and provide examples of marking using anonymous essays. Cartney found that student anxiety was reduced after the groups began to work together and built cooperative relationships. Students reported increased ability to focus on their own potential the longer the groups were together. The students also divulged differences concerning the feedback process, some claiming adequate and helpful comments while others received little feedback at all. Carney focused on the anxiety level of her students for this study and concluded that meeting the emotional needs of the students must occur so that learning is more effective.

Tamjid and Birjandi (2011) realized a need for altering traditional assessment in order to stay current with the movement toward learner-centered classrooms. The purpose of completing this study was to determine if self- and peer-assessment had any role in enhancing student autonomy and awareness. The experimental group completed self-assessment homework about their writing that students discussed in small groups during class. These students also assessed their peers' essays. The control group only received instructor feedback the following day in class. The instructor allowed both groups time to revise their essays before final assessment. Tamjid and Birjandi found their null hypothesis unfounded; self- and peer-assessment had a positive effect on learner responsibility and self-awareness. Tamjid and Birjandi concluded that reflective measures, such as self- and peer-assessment, assist teachers in moving toward a learner-centered environment in which learners have greater control over the learning process.

High school teachers typically discover similarities between their students and those of college students in their beginning years of language study. Yang (2011) realized that most college students exhibited weak writing skills and decided that peer assessment

may move students toward proficiency. Students uploaded written assignments to a website approximately every three weeks. The site helped students to see modeling of the assignment, scaffolding, articulations, as well as receive coaching with the ability to reflect and explore independently. Yang found that the students seriously engaged in this process. The researcher discovered that the students actively interacted, felt that the site positively impacted their writing, and held positive attitudes about the peer-review process. Yang concluded by advocating the implementation of a peer-review system when writing must be assessed.

More researchers continue providing evidence for the value of self-and peer-assessments in the L2 setting. While there are criticisms that some students do not have the skills to properly assess others' work, proponents argue that teachers can train their students to proficiently assess. Teachers can perform a demonstration of what peer-assessment would look like in front of the class. L2 teachers should also provide clear instructions with a rubric or assessment form to be completed by the peer editor. Additionally, during the peer-assessment process, teachers should act as facilitators as they circulate the room assisting students and prompting for explanations from the students. Most researchers noted the value-add of increased student motivation through the self- and peer-assessment process. Since this study also concerns itself with the students' perceived value of each alternative assessment, a review of student perceptions is necessary.

Student Perceived Value

While it is important to provide a wide range of literature on proven alternative assessment types for the L2 classroom, it is also vital to include literature concerning the

feelings students have about their assessments and how those assessments are of some benefit. Developing ownership for one's work is essential for continued success (Riedinger, 2006). Ownership and student value have been shown to be developed through the following: activities that relate to life, clearly communicated learning targets, positive classroom atmospheres, and alternative assessment creation (Palloff & Pratt, 2007). Additionally, Gardner (1985) defined motivation, specifically for L2 learners as "a complex of constructs, involving the combination of effort plus desire to achieve the goal of learning the language plus favorable attitudes toward learning the language" (p. 10). L2 educators should be prompted to develop assessments that the students perceive as valuable and that increased motivation toward second language acquisition.

Machemer and Crawford (2007) found, however, that students did not show strong preferences toward one assessment type or teaching method as much as they valued results. Students stated that they preferred higher activity levels in the classroom but that anything that helped them perform better was perceived as valuable. This could be problematic for teachers deciding methods and assessments since student opinions vary greatly. Most studies in the realm of student perceptions had a basis in the learner-centered classroom which values the opinions of students.

Burkšaitienė and Teresevičienė (2008), for instance, saw a change in education with a shift away from teacher-centered instruction toward learner-centered instruction; constructivist theory and higher-order thinking designs cemented their opinion. Burkšaitienė and Teresevičienė's purpose was to gain the students' perceptions of using alternative assessments. Students completed a project with a self-chosen group and a writing portfolio that the instructor designated time in class for progress to allow for

feedback and assistance. Burkšaitienė and Teresevičienė concluded that alternative techniques that used peer and instructor, in-class help was effective, useful, productive, and motivational for their students.

Asghar (2010) felt that cooperative learning techniques could be advantageous for her college Physiotherapy students. The purpose of the study was to attain the perceptions of these students about reciprocal peer coaching (RPC) and the assessment process. RPC required students to provide feedback to their classmates as they completed tasks throughout the course. Students reported that they appreciated the RPC approach because it confirmed their areas of strength and allowed them to immediately seek help for areas of weakness. Students also revealed that they enjoyed learning and working as a group since each person in the group had to meet basic competencies. Asghar found that RPC did help her students become more self-regulated, focused, process-oriented, and confident. Asghar concluded by touting RPC's merits for her students: higher motivation, better time management and goal setting, and increased emotional well-being.

Birjandi and Tamjid (2010) realized that educators were considering the role of the learner more than in past years. He also acknowledged how motivation is a primary consideration for learners. All students completed a motivational questionnaire with a Likert scale that considered attitude, self-confidence, and anxiety in relation to the course, as well as a language proficiency test. The researcher concluded that the use of journal writing and self-assessment practices effectively increased students' learning and motivation while improving self-confidence.

Little (2005) supported a learner-centered approach to teaching which maintained self-reflection. Hamidi (2010) agreed with Little as he described a shift from product-

based assessments toward process-oriented assessments in which students have more voice. Further, Hamidi claimed that today's assessments must promote learner autonomy, motivate, employ more learner strategies, actively engage students, and meet students' developmental needs. In order to accomplish these demands, teachers must understand their students and know what the students perceive as valuable in meeting their educational goals. One of the ways to acquire student perceptions is through involving students in the planning process. Students should be aware of all learning targets and know what they need to do to meet the educational outcomes.

Bakar, et al. (2010) further discussed motivation as it related to language learners. The researchers acknowledged its obvious importance and believed that motivation "initiates, directs, coordinates,…and evaluates cognitive and motor processes" (p. 72). MacIntyre, MacKinnon, and Clement (2009) named motivation "the engine that drives the system" (p. 44). If motivation drives students, then teachers need to understand their students' motivations through seeking out student perceptions.

Motivation comes from perceived value. Perceived value for students is, according to Kaufman and Dodge (2009), "how worthwhile or important a person perceives an activity to be. Individuals that find personal meaning or importance in engaging in a behavior will be more likely to internalize that behavior" (p. 102). Educational factors most associated with value by students are autonomy, personal choice, and accomplishment of mastery goals. Reverberating a similar definition, Vardi and Bunker (2001) related that, in order to create value for students, teachers should assess students on materials that the students feel are relevant and allow students to be involved in the

learning process while students become vested by active participation and feeling a sense of ownership.

Motivation, a key factor in student perceived value, is critical for success in the L2 setting. Cheng and Dörnyei (2007) strongly conveyed that

> it is fair to say that without sufficient motivation even the brightest learners are unlikely to persist long enough to attain any really useful language proficiency, whereas most learners with strong motivation can achieve a working knowledge of the L2, regardless of their language aptitude or any undesirable learning conditions. (p. 153)

Cheng and Dörnyei noted that much literature existed about motivational theories themselves; however, until recently not much research provided practical techniques for fostering motivation in the L2 classroom.

Rueda and Chen (2005) upheld the idea that student motivation was paramount for learning: "Motivational theories about foreign language learning have been valuable in explaining and enhancing learners' foreign language acquisition" (p. 14). Teachers need to discover their students' motivations, then, in order to be more effective in choosing assessments. Pintrich (2003) had already considered learning behavior a reflection of internal processes. He supposed that most students asked themselves several questions as they were learning. Students wanted to know why they were learning the material or the value of the lesson. They also questioned whether they would be successful or not. Finally, and essential to this research which asks for students' perceptions, Pintrich believed that students considered their own feelings about what they

were learning and how they were being assessed. Students stated that they felt bored, excited, disinterested, etc.

How then do teachers keep students motivated and add value to their methods and assessments? The value component of a task, or any assessment, is based on three things: importance, interest, and utility (Pintrich, 2003). Students' perceptions about the significance and practicality of something, as well as how attracted they are to the topic, affect classroom performance and assessment outcomes. The three alternative assessment types, Dynamic Assessment, specifically focused on conferencing and group strategy sessions; Task-based Assessment, specifically focused on real world communication; and Formative Assessment, specifically using peer- and self-evaluations were all chosen as viable alternative assessment options that would potentially be considered valuable by the students.

Conclusion

Dynamic Assessment, Task-based Assessment, and self- and peer-evaluation were chosen as potential alternatives to traditional assessments in the wake of new demands on foreign language teachers at the secondary high school level. Porter, et al. (2011) upheld the usage of the Common Core State Standards (CCSS) partly because it places "more emphasis on 'demonstrate understanding'…than do the state standards" (p. 107), as well as greater importance in the area of writing. Dynamic Assessment promotes greater understanding of the subject area as well as learner self-awareness and teacher mindfulness concerning student needs. Task-based Assessment encourages authentic communication and problem-solving in the target language. Self- and peer-evaluation inspire students to reflect and improve upon their writing in an authentic way. All three

assist in meeting the demands of the CCSS and standards for foreign language instruction:

> When reviewing the Mathematics and English Language Arts Common Core State Standards, it is clear the majority of the components mirror world languages. Perhaps the level and intensity of the knowledge of the content is not the same, but the intent of the standards certainly is… It is important that language educators move students from simple to complex language usage as they develop the ability to communicate in authentic contexts. (Heining-Boynton & Redmond, 2013, p. 52)

As today's secondary foreign language educators endeavor to teach and assess their students, Dynamic Assessment, Task-based Assessment, and self- and peer-evaluation should be considered as solid, data-driven choices. All three of these assessments have strong indications of learner value in the classroom and perceived value by the learners themselves.

Summary

With a comprehensive review of the related literature concerning Dynamic Assessment, Task-based Assessment, and Formative Assessment featuring self- and peer-evaluation, as well as student perceived value, this researcher was prepared to analyze data through the predetermined mixed method. Both quantitative and qualitative measures were incorporated.

CHAPTER 3

METHODOLOGY

Introduction

The literature review provided theoretical background as well as comprehensive information concerning the three alternative assessment types used in this study: Dynamic Assessment, Task-based Assessment, and Formative Assessment using Self- and Peer-Evaluations. Further, this review explored student motivation and perceived value as part of the education process. Information about the Common Core State Standards was also summarized, as it relates to foreign language educators. This chapter presents the data collection process involved in answering the five research questions for this study. A discussion of the statistical methods incorporated for data analysis will also be provided. The researcher will describe her chosen design, research procedures, population and sample demographics, and the statistical methods necessitated by the collected data. The chapter will conclude with some of this study's limitations.

Research Design

The purpose of this concurrent nested design study was to acquire the perceptions of first-year, high school foreign language students concerning the impact of alternative assessment types on their interest. Simultaneously, the researcher wanted to ascertain that the experimental group receiving alternative assessments learned equally or even exceeded the summative assessment scores of the control group who received more traditional-type assessment throughout the semester. To accomplish these goals of

discerning student perceptions with their test scores, the researcher primarily chose the following quantitative analyses: analysis of variance (ANOVA) with Bonferroni's post hoc to test all pairwise comparisons, independent-samples *t* test, and correlational statistics. The researcher asked School X teacher to keep a journal and surveyed the experimental group students post-study. The qualitative information gained from the teacher's journal notes and from the students' open-ended survey answers required coding answers, identifying themes, and discussing those themes.

The following research questions guided all procedural structure, data collection, and statistical analysis:

Research Question One: What perceptions do first-year foreign language students have of Dynamic Assessment, specifically focused on conferencing and group strategy sessions?

Research Question Two: What perceptions do first-year foreign language students have of Task-based Assessment, specifically focused on real world communication?

Research Question Three: What perceptions do first-year foreign language students have of Formative Assessment, specifically using peer- and self-evaluations?

Research Question Four: How do the summative assessment scores of the students who received interventions differ from those who did not?

Research Question Five: How do student perceptions of the alternative assessments correlate with their test scores?

So that this researcher could quantify the perceptions of the students, a simple survey was developed utilizing a four point, Likert-type scale (see Appendix A for the student perception surveys). Likert-type surveys provide a more objective manner in which to collect data so that it may be quantitatively analyzed (Gay, et al., 2009). By

quantifying the student perceptions, the researcher was better able to generalize the results. After completing each of the three alternative assessments, the students responded to two basic statements: *I was able to show the teacher how much I knew on Assessment __ (Blue, Red, Green)*, and *I would like my teacher to use Assessment ___ (Blue, Red, Green) again this year*. The choices on the four point Likert-type scale included Strongly Agree, Agree, Disagree, and Strongly Disagree. The researcher also developed an open-ended final survey that required the students to choose which assessment was the best for them and the worst, as well as the reasons for their choices. Before administering the surveys to the experimental group, the researcher gave the surveys to one of her own classes to look over and to make comments to ensure that the surveys were easily understood. Because of their simplicity, no piloting or changes needed to be made.

The researcher continued by developing, in cooperation with the classroom teacher, three unit assessments that met each unit's goals. Chapter two of the classroom teacher's textbook at School X, *Navegando 1* (Funston & Vargas Bonilla, 2005), was the starting point for the study. Chapter one is very introductory in nature and not suitable for this study; whereas, Chapter two employs the mastery of verb conjugation in connection with subject pronouns and is more appropriate for assessing competency. Dynamic Assessment was chosen for this unit for the following reasons: it could be combined with other forms of traditional assessment as a complement, could identify learning potential, and could classify student needs (Haywood & Lidz, 2007). Additionally, the classroom teacher incorporated Dynamic Assessment as a complement to a traditional summative exam. The Dynamic Assessment process was used to stimulate student growth and

discern student abilities through the teacher's hints, prompts, and leading questions in order to promote cognitive functioning (Haywood & Tzuriel, 2002; Lantolf & Poehner, 2004). In following a Vygotskian approach in using Dynamic Assessment, the researcher administered a pretest to the students, incorporated a structured intervention session with the students and allowed them to set goals, and then assessed the students using a posttest that aligned with the pretest in structure and content. The researcher provided the intervention after analyzing the pretest information and helped the students set goals.

Progressing to Chapter three of the textbook, *Navegando 1* (Funston & Vargas Bonilla, 2005), the classroom teacher and this researcher noted the learning objectives: to be able to discuss how to go somewhere on specific transportation, talk about places in a city and a restaurant, order from a menu, make introductions, and use the verb to go. Task-based Assessment seemed ideal for this unit. A unit assessment was designed to ascertain the student's ability to maneuver around a city and function in a restaurant setting by giving them real-world tasks. The tasks required of the students in the experimental group were communicative and met the unit learning objectives. Students were exposed to authentic language use in the classroom. They were assessed at the completion of Chapter three, primarily on their ability to produce dialog that was understandable and met the unit learning objectives. Grammar was a secondary consideration when grading since meaning takes precedence over grammar in a Task-based Assessment.

Finishing the study with Chapter four of the textbook, *Navegando 1* (Funston & Vargas Bonilla, 2005), the classroom teacher and this researcher directed themselves again to the learning outcomes: to talk about family and relationships, to describe

people's physical and personality traits, to state likes, to say what people do, and to show possession. Specifically, the students composed a pen pal letter using the unit objectives. As the Spanish 1 students progressed in knowledge and language ability, the need to work on writing skills in the target language became apparent. As the young learners began to write, it was essential that they received proper, positive feedback so that they improved and remained encouraged to continue writing. Formative Assessment using self- and peer-evaluation in a writing assignment worked well for this chapter. After successfully designing three unit assessments and the short student surveys, the researcher and classroom teacher proceeded to complete the classroom research.

Population

Considering all Spanish language learners at School X as potential participants in the study, the researcher had already narrowed the population to include only first year students. The sample comprised 48 students in three different sections, from an available 81 potential participants (n = 48). The researcher pre-determined that the participants in the first section of the day would be designated as the control group (n = 14), a sample of nine females and five males. The remaining two sections served as the experimental group (n = 34), a sample of 15 females and 19 males. School X's computer scheduling tool placed students into sections, with minor adjustments from counselors. The researcher did not have access to student placement or identification at any time to eliminate any potential bias on the part of the researcher.

Data Collection

Throughout the fall 2013 semester at School X, the researcher collected data from the participants. The students used a code for anonymity on all surveys and tests, instead

of their real names. The teacher and this researcher had the key which identified the code correlated to each student's name. This key was destroyed, post-study.

There was limited interaction between the researcher and the participants during the study. These times included the pre-study information session for all participants and the Dynamic Assessment intervention for the experimental group only. The classroom teacher administered the assessments and surveys. This process allowed the classroom teacher to conduct business as usual so that the control and experimental group classrooms were taught similarly, with exception of the unit assessments.

The participants completed each unit assessment and then immediately completed the survey for that assessment. The students did not know their score on the assessment at the time of the survey to avoid bias. The participants answered whether they felt that the assessment measured their knowledge and whether it should be administered again in the year, based on their own perceptions. The end-of-study survey was administered after all scores had been received. The classroom teacher reminded the students of each assessment type with a brief, unbiased description before they completed which assessment type they liked best, worst, along with their reasons why.

The researcher entered all data into an Excel spreadsheet as the year progressed. Responses to the four point Likert-type questions were coded as follows: strongly agree with 4, agree with 3, disagree with 2, and strongly disagree with 1. The researcher added the two questions' scores together for each student, for a possible agreeability score range of 2 to 8. The open-ended responses were coded as follows: Assessment Blue as best or worse with 1, Assessment Red with 2, and Assessment Green with 3. The students'

reasons for their rankings were categorized by the researcher and entered into a Word document. The teacher's brief journal notes were also similarly noted.

Concurrently, the researcher also entered individual test scores of the experimental group into Excel for all three assessments. Additionally, the final exam scores of every participant were compiled.

Analytical Methods

All data from the Excel spreadsheet was entered into Statistical Package for the Social Sciences Version 21.0 (SPSS) by the researcher. Research Questions One, Two, and Three were essentially the same question. All three were experimental and required the researcher to visually display the student perceptions of three different assessment types that were incorporated during the semester and compare the perceptions among the three assessments. An experimental design was used with an experimental group that experienced the intervention of alternative assessments and then were measured concerning the effects of said interventions (Leedy & Ormrod, 2013). This researcher introduced three different assessment types to the experimental group and measured their perceptions following each one and at the end of the study. The student perceptions were the dependent variables which could or could not have been influenced by the independent variables, the alternative assessments administered. Likert-scale survey questions were used as an ordinal level of measurement. The researcher manipulated between-subject data within the experimental group by comparing perceptions of assessment one to those of assessment two and those of assessment three. Since the control group did not experience the alternative assessments, no data existed for them in regard to Research Questions One, Two, and Three. To show the students' perceptions of

each assessment type, the researcher combined the two 4-point Likert-type question scores for a maximum of eight points and displayed the data using bar charts. The researcher also used the student data to rank the three alternative assessments in order of preference, based on information from the post-study survey, and then described these results. With three different assessments' perception scores (three variables), the researcher used a within subjects ANOVA with Bonferroni's post hoc to test all pairwise comparisons to determine the differences. This SPSS analysis showed the students' perceptions of each assessment type as well as a comparison among the three types.

The researcher also used an experimental design for Research Question Four as she needed to compare the final exam scores of the control group to those of the experimental group. This research design allowed the establishment of whether the alternative assessments administered to the experimental group impacted the participants' final exam scores in comparison with the control group members' scores. The researcher manipulated the independent variable (assessing the experimental group through alternative measures) to compare to the established dependent variable (using traditional means to assess the control). The researcher used an interval level of measurement by obtaining data from test scores. The independent variable manipulation was between-subjects, comparing the scores of the control to those of the experimental groups. Because two different groups' scores exist, the researcher examined the differences between the groups using an independent-samples *t* test to disclose *p* values, means, and standard deviations of the two groups. This information was visually displayed using box plots.

Research Question five was correlational in that it was a "statistical investigation of the relationship between two or more variables" (Leedy & Ormrod, 2013, p. 100) and was not concerned with revealing causes. The researcher was interested in showing a relationship, if any, between alternative assessment scores and students' perceptions of each assessment type. To that effect, the researcher used Spearman's Rho correlation coefficient to determine a positive, negative, or nonexistent correlation between the students' perceptions and their alternative test scores. For each assessment, the researcher correlated the student's alternative assessment score to his/her combined survey question scores. The relationships were displayed using scatter plots for the Spearman's Rho analysis.

Limitations

Some limitations existed for this study. One limitation was the sample size, as well as the location. With only 48 total participants, the study results may not be easy to generalize to the wider national audience of second language learners. Also, the student population was not very diverse, generally Caucasian and not low-income; a more diverse population may have garnered different results and might represent the general population more accurately. This study also took place in only one high school in a large city. Expanding the study to more schools, some urban and rural may create a more generalizable study. Another potential limitation was the answers on the student surveys. Being young, some students may have carelessly completed the surveys just to fulfill their obligations and not genuinely answered the questions. Students may have also been too thoughtful about their answers, in an attempt not to offend the researcher, after meeting her during the orientation session and during the Dynamic Assessment

intervention session. Due to restrictions placed on the researcher regarding her time a School X, the researcher was only allowed to create one alternative assessment per unit. With an expanded experimental group, the researcher could have used all three alternative assessments with each unit, potentially taking away any researcher bias that inadvertently may have resulted from creating only one alternative per unit. Additionally, if the students could have been exposed to the three alternative assessments for two units each, instead of one, their perceptions may have changed.

Summary

The previous information provided detailed explanations of the methodology used during the course of this study, completed during the fall of 2013. The researcher provided a documentary of the process and rationale involved in creating the alternative assessments and students' surveys. An explanation of the analysis required, based on each research question, was provided, including dependent and independent variables as well as the SPSS analyses used. The succeeding chapter will address the exact results produced from the collected data with its interpretations with recommendations.

CHAPTER 4

FINDINGS AND CONCLUSIONS

Introduction

After designing a research study on alternative assessments in the L2 classroom and reviewing the available and applicable literature, the researcher collected the data and completed the Statistical Package for the Social Sciences Version 21.0 (SPSS) analyses. Subsequently, the researcher was able to determine any statistically significant results and report meaningful interpretations of the findings. Tables and figures will be provided as visual representations of the data. This chapter will present the SPSS data, analyze the information, and explore possible meanings and draw conclusions from the findings for the following Research Questions:

1. What perceptions do first-year foreign language students have of Dynamic Assessment, specifically focused on conferencing and group strategy sessions?
2. What perceptions do first-year foreign language students have of Task-based Assessment, specifically focused on real world communication?
3. What perceptions do first-year foreign language students have of Formative Assessment, specifically using peer- and self-evaluations?
4. How do the summative assessment scores of the students who received interventions differ from those who did not?
5. How do student perceptions of the alternative assessments correlate with their test scores?

Findings

Research Findings: Questions One through Three

Research Questions One, Two, and Three required descriptive statistics that measured students' perceptions following the dynamic, task-based, and formative unit assessments. To show the students' perceptions of each assessment type, the researcher combined the two 4-point Likert-type question scores for a maximum of eight points and displayed the data using bar charts (Figures 1, 2, and 3). Figure 1 indicates that 25 of the 34 students or 73.5% in the experimental group agreed that the Dynamic Assessment showed the teacher how much they knew and favored using this type of assessment again in the future. None of the students strongly disagreed that this type of assessment should be continued nor that it showed how much they knew. Out of the 34 students, only 6 or 17.6% had disagreement to the Dynamic Assessment.

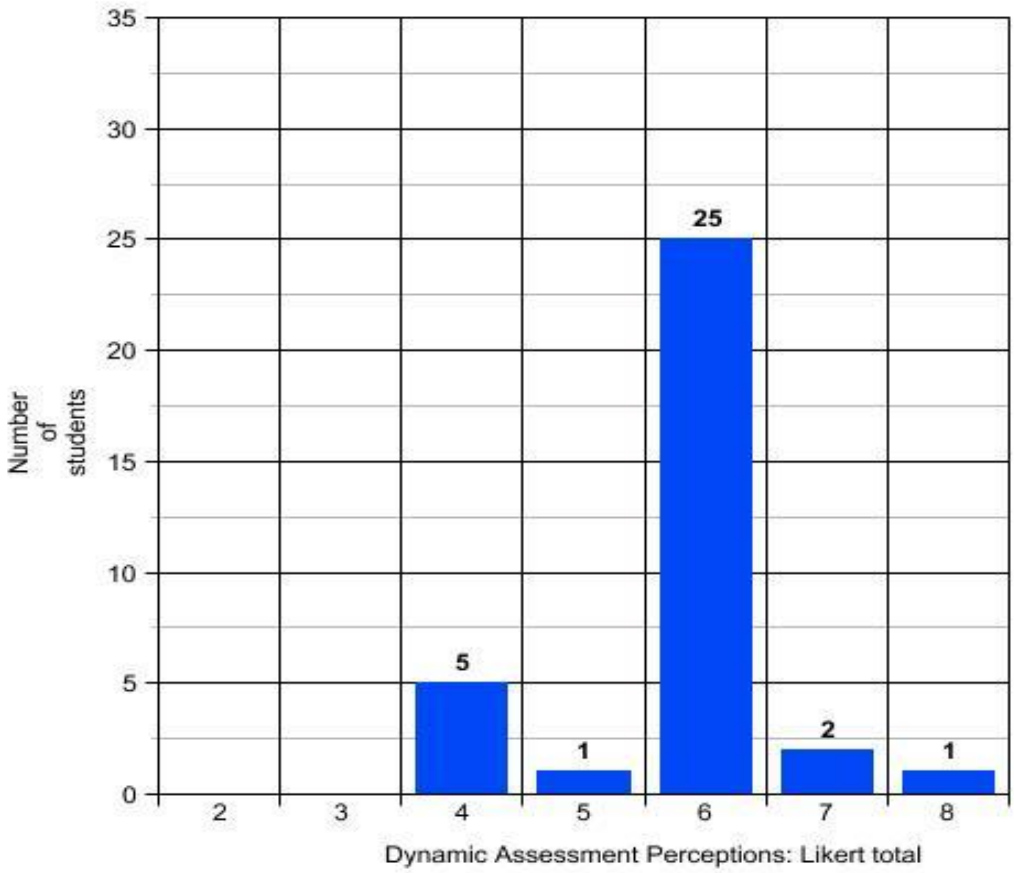

Figure 1. Dynamic Assessment perceptions: Blue.

Figure 2 shows that five of the 34 students or 14.7% strongly disagreed that the Task-based Assessment administered showed what they knew and that the teacher should use this assessment type later during the year. Twelve of the 34 or 35.3 had a mixed opinion of agreement with disagreement on the assessment. Overall, no students strongly agreed concerning Task-based Assessment.

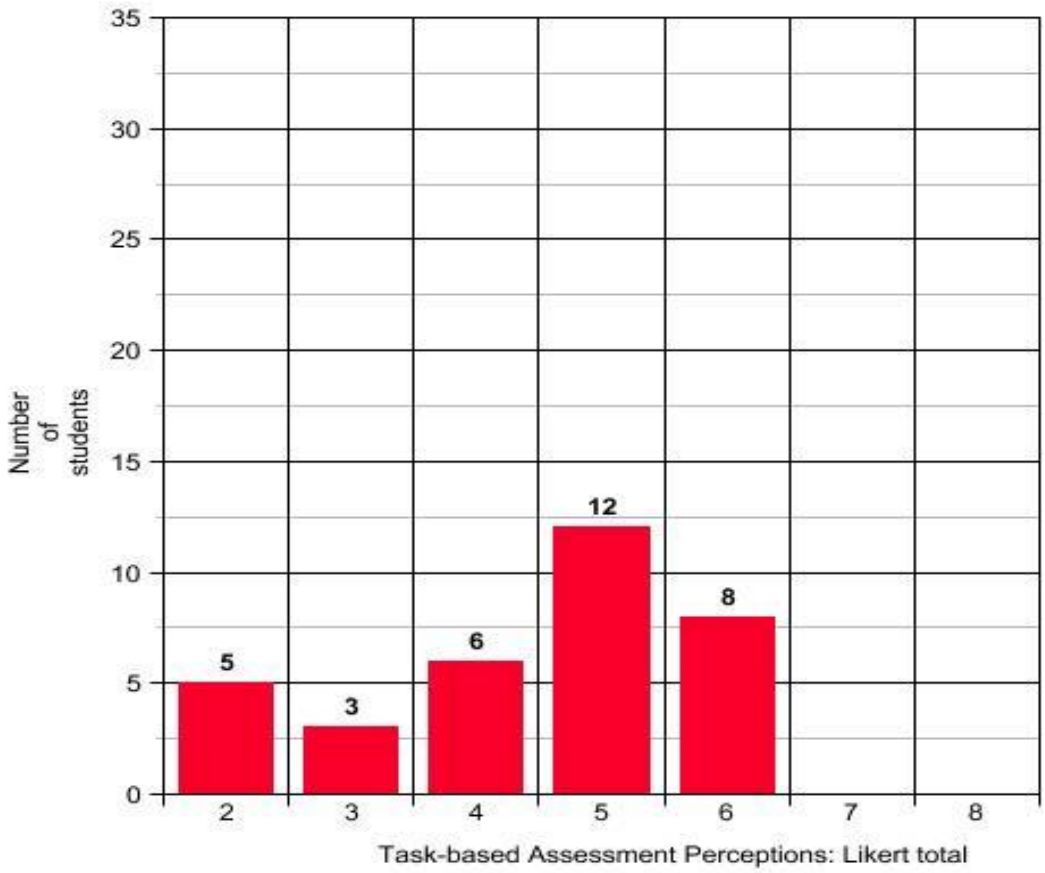

Figure 2. Task-based Assessment perceptions: Red.

Figure 3 illustrates the overall favorability of Formative Assessment. For agreeing or strongly agreeing that this assessment showed the teacher what they knew and should be used again, 28 of the 34 students or 82.4% fell into these categories. None of the students strongly disagreed on both questions, and only 2 or 5.9% strongly disagreed. One student did not believe that the assessment demonstrated knowledge and another that the assessment should be administered at another time. The Formative Assessment garnered 10 indications of strongly agreeing, more than either of the other two assessments.

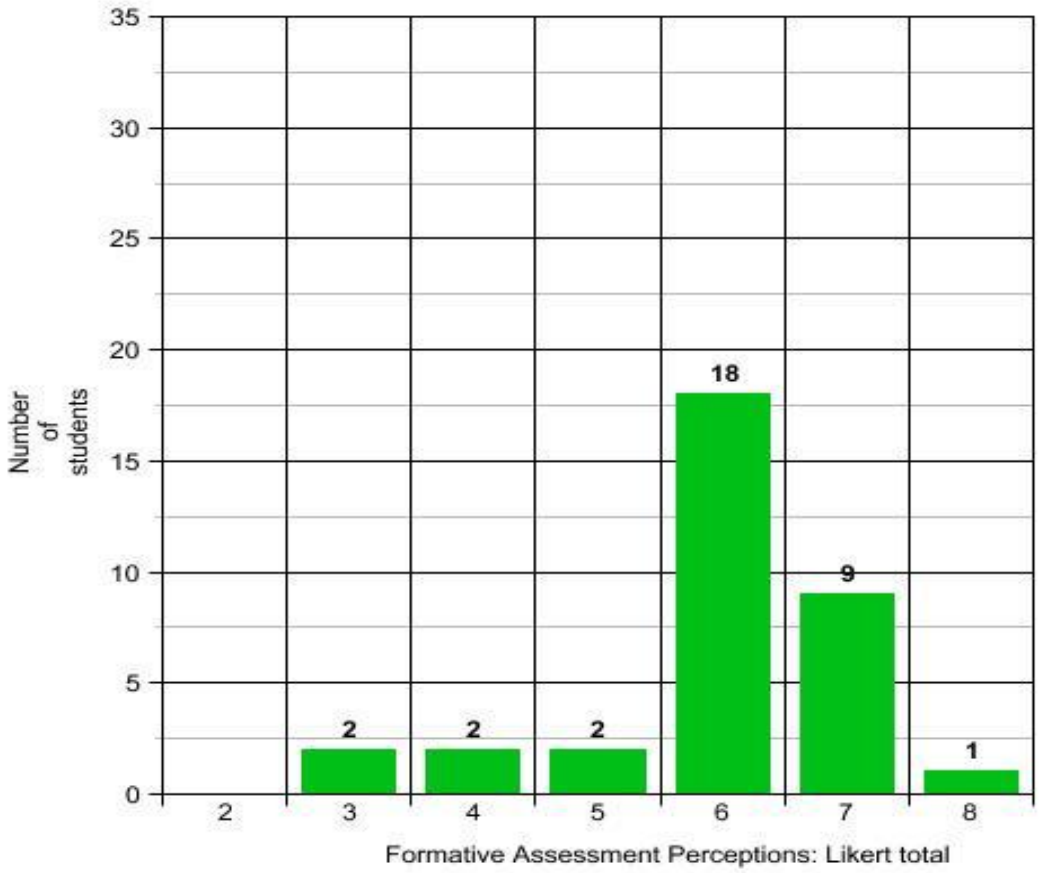

Figure 3. Formative Assessment perceptions: Green.

Overall, students were fairly agreeable to all three assessment methods, even though they preferred Dynamic and Formative Assessments over Task-based Assessments.

To finalize the quantitative analysis for Research Questions One, Two, and Three, the researcher computed a one-way, within subjects ANOVA to compare the effect various alternative assessments had on perceptions, specifically regarding Dynamic, Task-based, and Formative Assessment types. The means of student perceptions were associated with the student perceptions at the three different points after each assessment. The researcher found statistical significance, that at least one of the means for perceptions was different from the others: $F(2, 66.0)= 22.75$, $p< .001$, partial $\eta^2= .41$. Using

Bonferroni's correction for family-wise errors, the student perceptions of Dynamic Assessment (Blue) did not differ greatly from the perceptions of Formative Assessment (Green): p=1.0, which is not statistically significant. However, when comparing the student perceptions of Dynamic Assessment (Blue) to Task-based Assessment (Red), p<.001 was statistically significant, with Dynamic Assessment being better perceived than Task-based Assessment. Similarly, Task-based Assessment (Red) perceptions compared to those of Formative Assessment (Green), p<.001 which indicated statistical significance, as Formative Assessment was better perceived than Task-based Assessment. The students had a lower perception of Task-based Assessment than they held of either Dynamic or Formative Assessment; while their perceptions of both Dynamic and Formative Assessments were similar. An eta-square of .41 showed a large effect size.

From the post-study survey, the researcher tabulated the number of students citing each assessment type as the best and worst for reflecting knowledge of the material. Twenty-five of the 34 students or 73.5% recorded Formative Assessment, in which they wrote and used self- and peer-evaluation, as the type that best reflected their knowledge, as shown below in Figure 4.

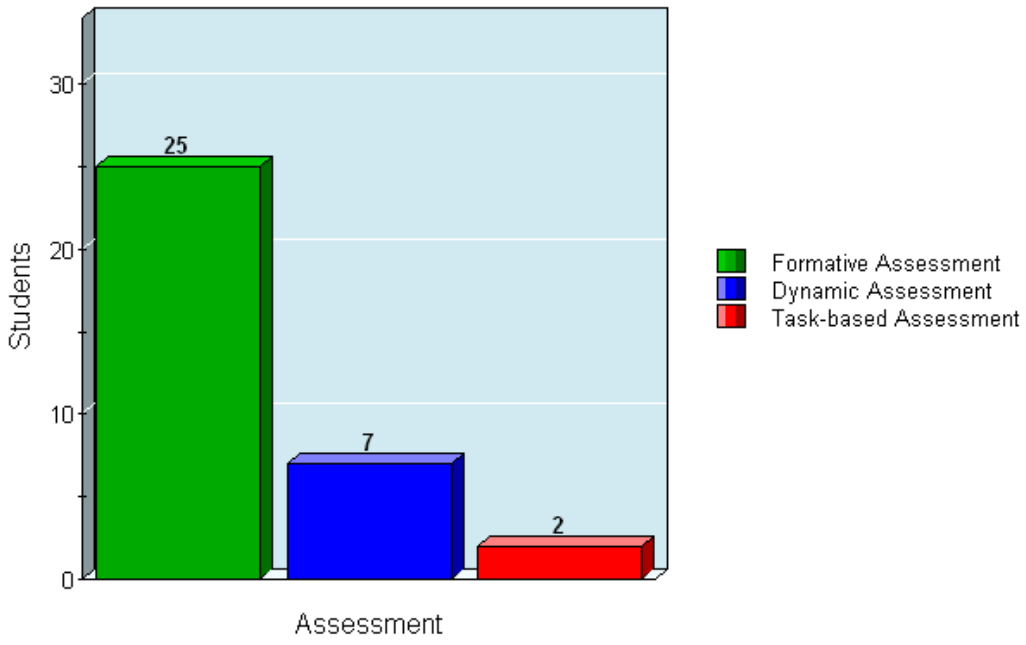

Figure 4. Best-Perceived Assessment type.

For the worst-perceived assessment, Task-based Assessment incurred 29 of the 34 students' choices, or 85.3%, as not reflecting their knowledge, as shown below in Figure 5. None of the students chose Formative Assessment as their least favored.

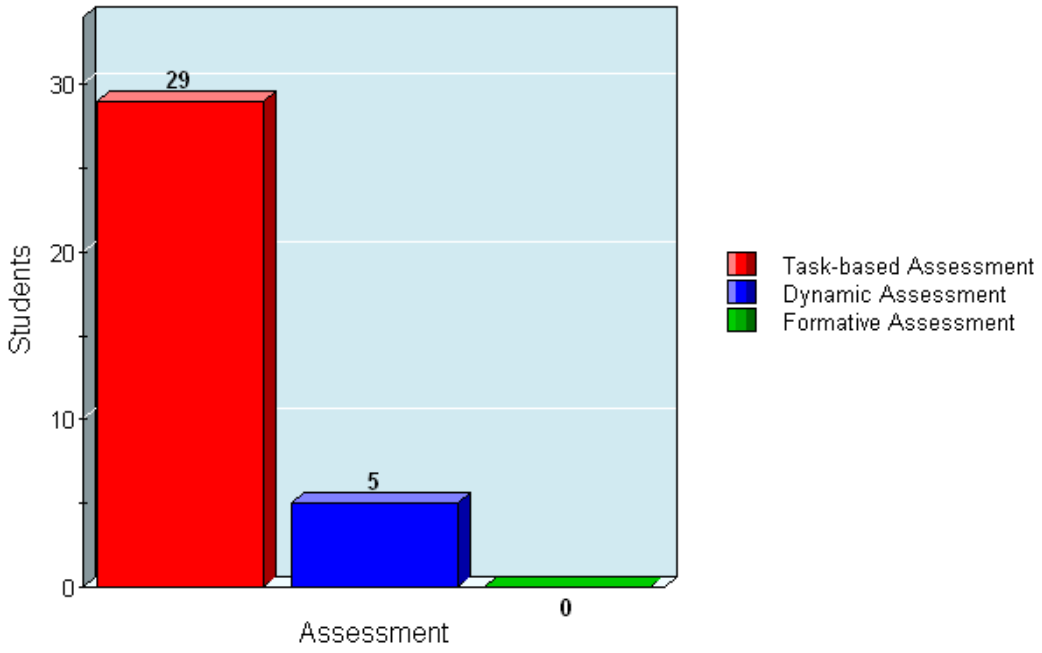

Figure 5. Worst-Perceived Assessment type.

Using qualitative procedures, the open-ended responses given by the students on the post-study survey were coded. Then they were assigned themes concerning the students' perceptions regarding why an assessment type was best or worst for knowledge reflection. Tables 1 and 2 depict those themes with the number of students claiming them. From Table 1, the eight students remarking that test format was crucial for the best perceived assessment also included specific details, such as, appreciating the revision process on the Formative Assessment, feeling advantaged from experiencing the intervention and taking the pretest on the Dynamic Assessment, and experiencing better success from the Task-based Assessment's separate task sections. Formative Assessment using self- and peer-assessment incorporated writing to demonstrate Spanish proficiency and knowledge. When referring to his preference for Formative Assessment, one student stated, "It let me show that I can communicate." Another student commented, "I could pick what to say based on the words I knew." An additional remark concerning

Formative Assessment indicated, “I could test my knowledge and have someone revise it.” In all, the students clearly favored the Formative Assessment as being most reflective of their knowledge, with 25 of the 34 students or 73.5% choosing it.

Table 1

Best-Perceived Assessment: Themes

	Themes				
Assessment Type	Seemed easy/ Knew material	Freedom	Test Format	Practical/ Communicative	Did Not Respond
Formative Assessment-Green	10	9	2	4	0
Dynamic Assessment-Blue	1	0	5	0	1
Task-based Assessment-Red	0	0	1	0	1

In reference to the worst perceived assessment, as indicated in Table 2, twenty-two students perceived that the format of the task-based assessment was not clear to help them understand exactly what the specifications of the test were. One student wrote, “It was too much…” and another related, “It was confusing and hard.” Two students reported that they did not favor a pre/posttest design, as seen on the Dynamic

Assessment. Overall, the students relayed that the Task-based Assessment did not reveal their understanding as fully as did the Formative or Dynamic Assessments. Of the 37 experimental group comments, 32 of them or 94.1% cited Task-based Assessment as the alternative they least favored.

Table 2

Worst-Perceived Assessment: Themes

	Themes			
Assessment Type	Confusing	Hard/Did not Know or Understand Material	Test Format	Did Not Respond
Task-based Assessment-Red	17	9	5	1
Dynamic Assessment-Blue	0	2	2	1
Formative Assessment-Green	0	0	0	0

Note. Some students provided a compound answer, so the number of students reflected in the table does not correspond with the number of students in the experimental group.

The classroom teacher provided limited journal notes concerning students' behaviors and comments at the time of each alternative assessment, but also delivered his

perceptions of how each assessment impacted the classroom. He noted that the Dynamic Assessment (Blue) format with a pre- and posttest, as well as an intervention and group strategy session, seemed beneficial to most students. The two-test method helped students better anticipate what to expect early in the year. The teacher also stated that the pretest and intervention guided his instruction to more effectively meet the needs of the students. However, many students complained that Task-based Assessment (Red) was difficult, and some stated it was confusing. The teacher felt that task-based activities may have been better suited as an in-class partner-practice activity or that some sort of task-based activity needed to be used during the unit and not just as the final assessment. Formative Assessment (Green) was well liked by the students. They commented to him that it was beneficial, practical, and allowed them to say whatever they wanted and avoid topics with which they had less comfort. Students appeared motivated to show off their knowledge for this assessment.

Research Findings: Question Four

To compare the mean of the final exam scores of the experimental group to that of the control group, an independent samples *t* test was calculated through SPSS. This information is visually displayed as a box plot in Figure 6, as shown below. There was not a statistically significant difference in the scores of the experimental group's final exam (M= 84.56, SD= 13.96) and from the final exam of the control group (M= 85.00, SD= 15.34); $t(46)= -.097$, $p = .923$, $d= .03$. These results suggest that the experimental group had comparable Spanish proficiency at the end of the semester to the control group. Both the control and experimental groups performed comparably on the final exam, even though they were assessed in different ways during the semester.

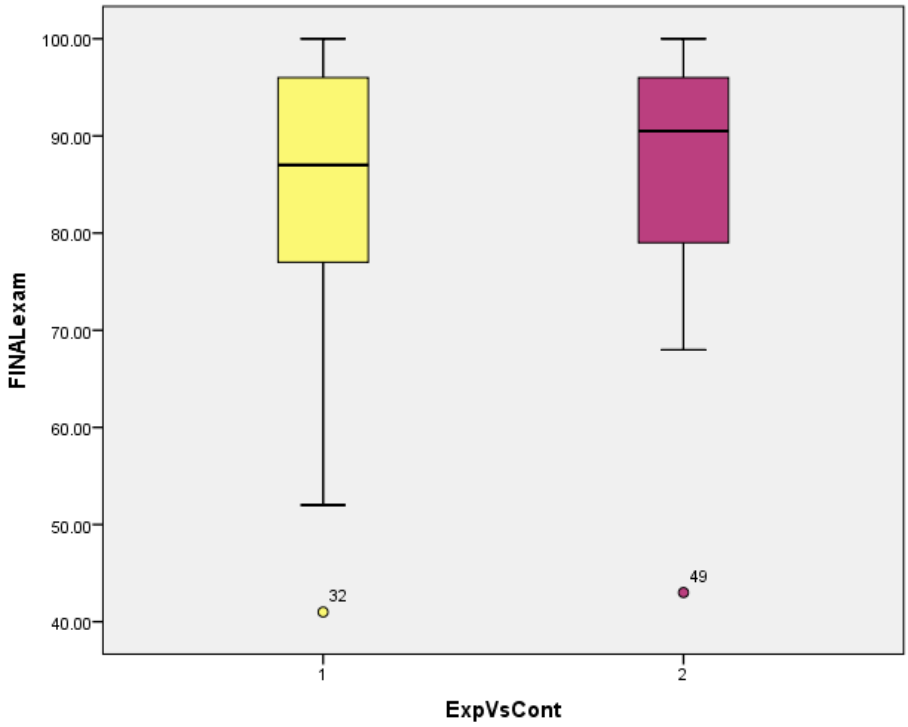

Figure 6. Final Exam scores. This figure illustrates the final exam scores of the experimental group in yellow in comparison to those of the control group in purple.

Research Findings: Question Five

Research Question Five was concerned with showing relationships between the students' perceptions of each alternative assessment and each student's score on that assessment. A Spearman's Rho correlation coefficient was computed to assess the possible relationships. Below, Figures 7, 8, and 9 illustrate the students' perceptions of an alternative assessment and their test scores on that particular assessment.

When analyzed, the Dynamic Assessment (Blue) showed no correlation between the students' perceptions and their scores, *r s*(32)= .29, *p*=.09, with a medium effect size of .293. Some students positively perceived the Dynamic Assessment and had high scores, while others viewed it similarly but had average or low scores. The following, Figure 7, summarizes the results with a scatterplot.

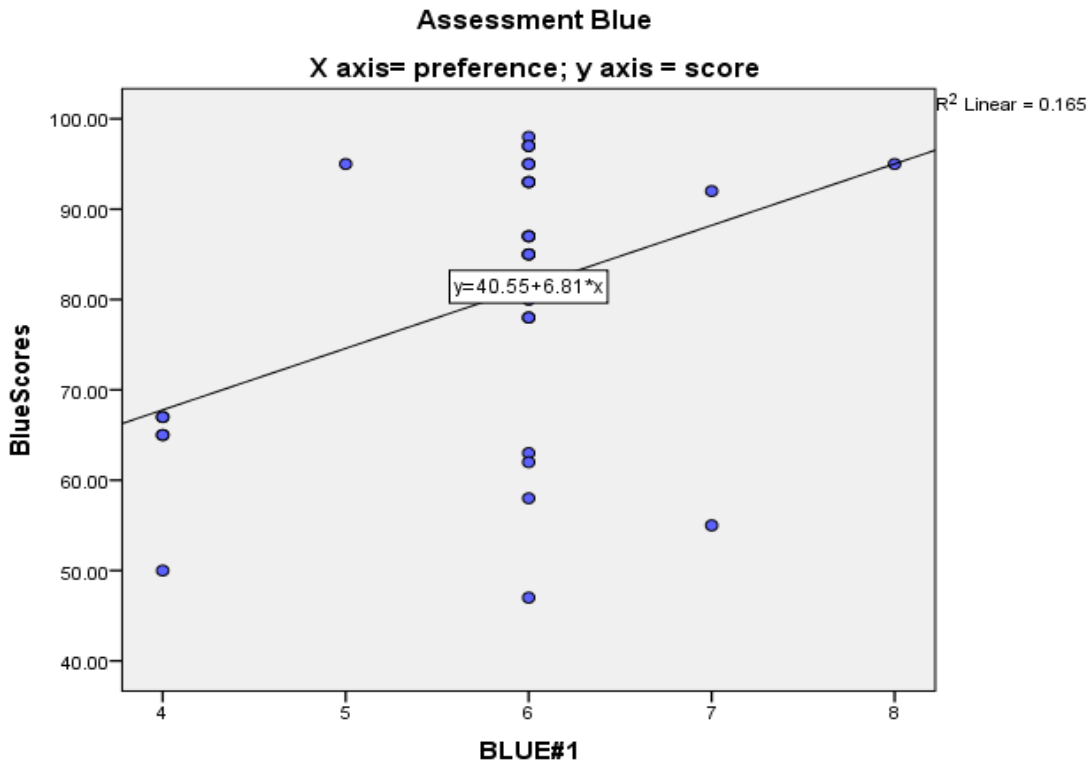

Figure 7. Dynamic Assessment (Blue) Correlation. This figure illustrates the correlation between the test scores (on a scale of 0 to 100) on Assessment Blue and each student's perception of the assessment (on a scale of 2 to 8).

In regard to the Task-based Assessment (Red), there was a significant positive correlation between the students' scores and their perceptions, *r s*(32)= .57, *p* < .001, with a large effect size of .574. As test scores rose, positive perceptions of the assessment also

rose; the converse was also true. A scatterplot below summarizes the results in Figure 8, as well as depicts the positive correlation with the ascending linear regression line.

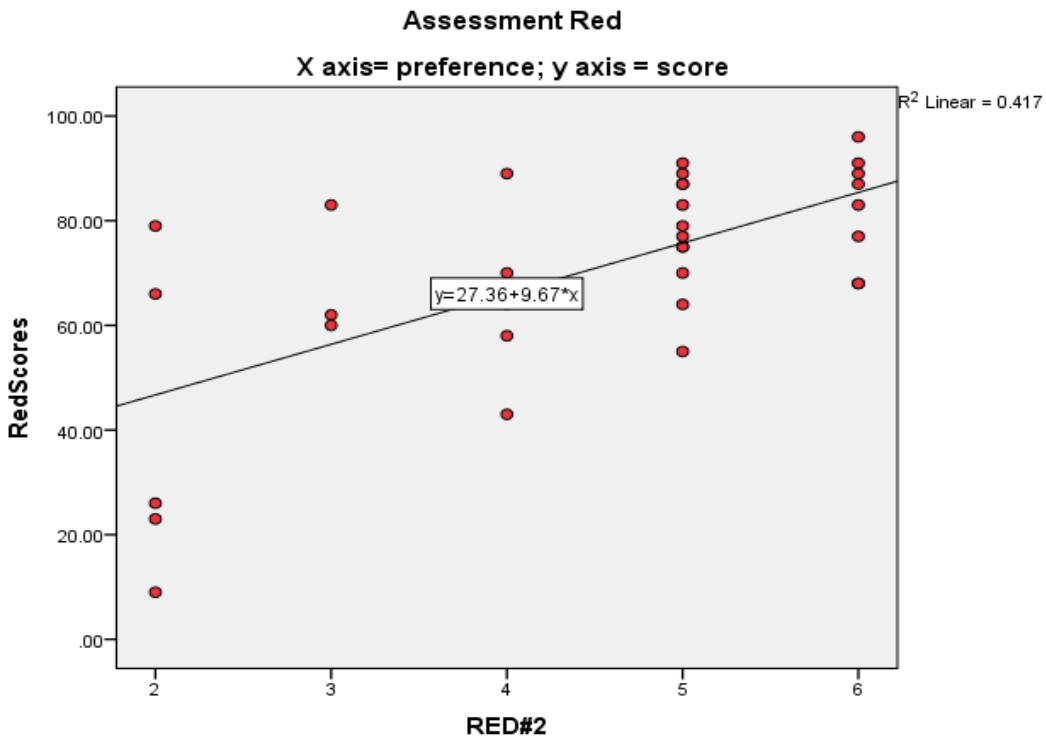

Figure 8. Task-based Assessment (Red) Correlation. This figure illustrates the correlation between the test scores (on a scale of 0 to 100) on Assessment Red and each student's perception of the assessment (on a scale of 2 to 8).

The Formative Assessment (Green) revealed no correlation between the assessment scores and the students' perceptions, $r\ s(32)= .10$, $p=.57$, with a small effect size of .101. The flat linear regression line on the scatterplot of Figure 9 indicates that no meaningful relationship existed between assessment scores and perceptions in relation to this assessment type.

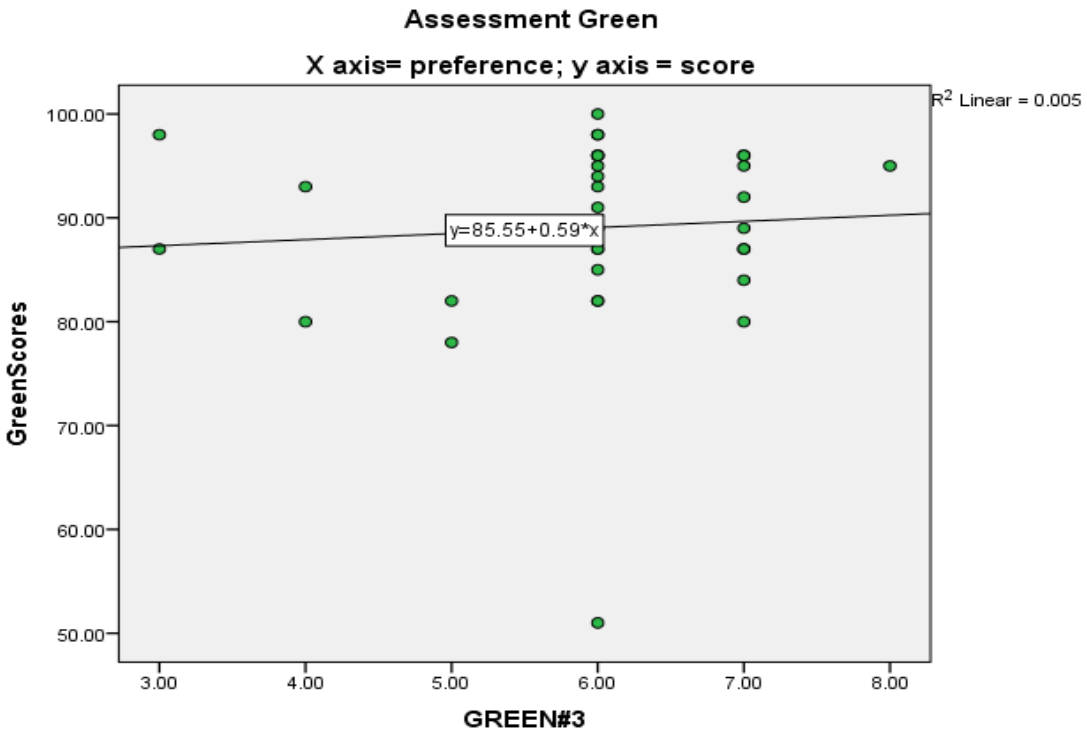

Figure 9. Formative Assessment (Green) Correlation. This figure illustrates the correlation between the test scores (on a scale of 0 to 100) on Assessment Green and each student's perception of the assessment (on a scale of 2 to 8).

Conclusions

It is noted here that many potential subjects for the experimental group were excluded from this study due to parental concerns about alternative assessment measures. These students' perceptions may have altered the data for all three assessment types.

Research Question One sought to determine the perceptions of first-year Spanish students about Dynamic Assessment. Twenty-eight of the 34 experimental group of students acknowledged Dynamic Assessment as agreeable or strongly agreeable, primarily in the agree category. Further, although Dynamic Assessment was not chosen as the best-perceived assessment type during the study, neither was it chosen as the worst.

The most noted positive features of the Dynamic Assessment were administration of the pre/posttest and the benefit of the intervention session. Based on student perceptions from the current study, it may be concluded that Dynamic Assessment was perceived as a favorable alternative assessment for foreign language classroom teachers. Dynamic Assessment should be part of the teaching repertoires in L2 classrooms, not simply due to its perceived favorability, but because Dynamic Assessment promotes self-awareness, teacher-learner needs awareness as a formative process, and is a proven method of increasing student achievement and motivation (Antón, 2009; Kozulin & Garb, 2004), as well as learner confidence (Mardani & Tavakoli, 2011).

In regard to Research Question Two, 20 of the 34 experimental group students noted that Task-based Assessment was agreeable to them, in that it demonstrated their knowledge and would be acceptable to use again. However, eight of the 34 students showed disagreeability or even strongly disagreed that Task-based Assessment was preferable. Task-based Assessment also ranked as the least preferred choice of the three alternatives. While Task-based Assessment may be an acceptable option, the results indicated that it was not highly preferable. While research supports the benefits of a task-based classroom, the data from this research may imply that using Task-based Assessments would not be highly recommended for the L2 teacher unless this method was used more consistently (Skehan & Foster, 1999) and within a communicative teaching atmosphere (Ellis, 2003). Without consistent practice of task-based pedagogy, students are too easily confused if L2 teachers only use this as an assessment, possibly negatively impacting scores and/or failing to demonstrate actual knowledge.

Formative Assessment perceptions' data needed to answer Research Question Three revealed that 28 of the 34 students agreed or strongly agreed with this alternative assessment's value. Ten of them showed levels of strong agreement. Additionally, Formative Assessment was established as the most favorable alternative assessment of the three, chosen by 25 of the 34 students; not one student ranked Formative Assessment as the worst choice. Given the student perceptions data, results established that Formative Assessment was well-perceived by the students in the experimental group and was the popular choice; thus, making Formative Assessment a valid alternative for L2 educators. Similar to the findings of Tamjid and Birjandi (2011), the data from this study may indicate that Formative Assessments featuring self- and peer-evaluation should be implemented by L2 teachers due to its highly agreeable perceptions by the students and previous research (Blue, 1994; Bryant & Carless, 2009; Pintrich, 2000) touting student self-awareness and improved personal responsibility. As L2 educators continue to find ways to create learner-centered environments, Formative Assessment using self- and peer-evaluation may provide one means to do just that.

Research Question Four, which compared the means of the control (M=85.0) and experimental (M=84.56) groups on the final exam, showed no significant difference. This data indicated that the end-of-semester knowledge for first-year Spanish students was equal. Students in both groups, as measured by a common cumulative assessment, had acquired the same knowledge and skills. It may be surmised that learners in neither group were adversely nor positively affected due to assessment types during the semester; therefore making alternative assessments viable choices for the foreign language teacher. While the future success of the Common Core State Standards remains to be seen,

researchers should study the effects of these new testing methods. In the meantime, L2 teachers will need to consider shifting from traditional assessments toward alternative assessments even though final knowledge and skills may not show any measured increase. One may ask why a teacher would make the effort to create alternative assessments if they would not yield a greater final knowledge. The reason for the shift is not based on end-game knowledge from a traditional exam but on the increased ability of students to incorporate higher-order thinking and to use critical thinking skills to solve real-world problems and use the language as an effective communication tool; these benefits far outweigh the time saved by teachers who continue to use traditional assessments. Additionally, in the future, final exams will most likely not be traditional in nature, as they currently stand, as teachers adapt their instruction and assessments to align with the Common Core State Standards. Final exams will change to reflect higher order thinking and be more writing intensive and real world oriented. According to the Marzano Center (2015):

> It is a shift in the philosophical thinking about the nature of teaching and learning. This shift basically says: We will no longer teach students to memorize by rote, to understand superficial facts and figures without more nuanced understanding, applicable to real-world problems. Rather, we will teach them to analyze...We will ask them to use complex cognitive skills to analyze the very complex problems they face as citizens in the 21st century. (para. 5)

Alternative assessments will lend themselves to preparing for these soon-to-come new formats for final exams.

When comparing the perceptions of the experimental group students to each one's alternative assessment scores for Research Question Five, a significant positive relationship existed between the students' perceptions of Task-based Assessment and their assessment scores ($p<.001$) indicating that Task-based Assessment did factor into performance. As positive perceptions increased, so did assessment scores. No correlation existed between the student perceptions of Dynamic Assessment and test scores ($p=.09$) or between Formative Assessment perceptions and test scores ($p=.57$). Since these analyses established that no statistical significance existed, student perceptions did not affect performance, positively or negatively, in relation to Dynamic and Formative Assessments.

While student perceptions are important (Birjandi & Tamjid, 2010; Cheng & Dörnyei, 2007), the results from this study indicate that, overall, students are agreeable to a variety of assessment types and perceive them as accurate reflections of their knowledge. It may be noted, though, that some assessment types need to be practiced and modeled throughout a unit to increase their effectiveness and perceived value. For example, Task-based Assessment was perceived as confusing to students when experienced for the first time on an end-of-unit assessment. Overall, data did not support that student perceptions affected the outcomes required on a Dynamic or Formative unit assessments. Data from the Task-based Assessment did indicate a positive correlation between student perceptions and the scores earned from this testing format. While student perceptions and choices are important, alone, they are not a reliable indicator on which to base the relationship between perceptions and scores. For instance, students do not necessarily have the maturity or knowledge to understand the value of certain activities or

assessments. Finding a middle ground in which both students and teachers perceive value from an assessment is ideal.

Implications and Recommendations

Upon completion of this study concerning perceptions of L2 learners about alternative assessments, this researcher has a desire to use data-driven pedagogy to propel her students. Considering the data from this study, students being assessed non-traditionally performed equally with those who were traditionally assessed, so L2 educators should strongly consider administering alternative assessments that also align with Common Core State Standards (CCSS). Poehner and van Compernolle (2011), Sidek (2012), and Tamjid and Birjandi (2011) all supported the need to move away from traditional assessments, as does this researcher. Administering assessments that promote critical reading, writing, and higher-order thinking skills is crucial as teachers' successes, evaluations, and retentions will be measured, in part, through their students' scores (Porter, et al., 2011). Clearly, while traditional assessments can measure student knowledge and skills, traditional assessments do not provide a foundation for students who will be required to critically operate on the new CCSS assessments (American Council on the Teaching of Foreign Languages [ACTFL], 1998; Illinois State Board of Education, 2012; National Governors Association Center for Best Practices, & Council of Chief State School Officers, 2010), so L2 educators should contemplate using Dynamic Assessments, Task-based Assessments, and Formative Assessments incorporating self- and peer-evaluation into their repertoire. Dynamic Assessment encourages learner and teacher awareness and would benefit L2 educators in formatively assessing students, given its pre/post-test format and intervention sessions. Task-based Assessment

incorporates authentic communication and problem-solving as its basis, components of the CCSS. Student reflection used for improvement in writing skills through Formative Assessment would further promote CCSS's resolve for writing proficiency.

In regard to the question of student perceptions of alternative assessments in this study, students provided a variety of ideas. While only Formative Assessment merited statistical significance as a highly favored assessment choice, all three types, Dynamic, Task-based, and Formative Assessments, received positive perception overall by the students. Given this information, foreign language educators must address the issue of effective assessment types while creating an atmosphere conducive to learning for the students. Seeking student perceptions while creating quality alternative assessments is a recommendation for L2 educators.

As the Alliance for Excellent Education (2011) reported the shift in expectations for both students and teacher alike, this researcher also believes that educators must demand more from students in the way of complex yet realistic thinking. This researcher recommends that L2 teachers break out of their traditional molds and create alternative assessments that require higher order thinking, communication, and accountability, such as those used in this study. Mandell (1999), Poehner & van Compernolle (2011), Sidek (2012), and Tamjid & Birjandi (2011) proposed alternative assessment methods which measured student outcomes but also served to increase critical thinking skills and/or real-world application. This researcher concentrated on Dynamic Assessment, Task-based Assessment, and Formative Assessment focused on self- and peer-evaluation so that students were required to explain, create, and write more. Generating a more communicative L2 classroom is this researcher's suggestion for teachers today as

communication promotes higher-order thinking, greater engagement, and personal responsibility, as well as provides realistic opportunities to learn and grow (American Council on the Teaching of Foreign Languages [ACTFL], 1998; Burkšaitienė & Teresevičienė, 2008; Riedinger, 2006). Creating alternative assessments may be a way to positively influence competencies for the L2 teacher and is highly advocated by this researcher.

The future is upon us as modern-day educators with the implementation of the CCSS. Regardless of opinions about this shift, students can benefit from their teachers' movement toward alternative assessments. Dynamic Assessments promote student growth through a process: pre/posttest method with interventions, remediation, and self-evaluation. Interaction facilitates learning and guides instruction (Feuerstein, et al., 1987). Task-based Assessment requires students, using an authentic approach, to perform tasks and communicate to solve a real-world issue (Byrnes, 2002; Ellis, 2006; Foster & Skehan, 1997). Self- and peer-evaluation as a Formative Assessment encourages student motivation and growth through personal responsibility while it builds communication through writing skills (Black & Wiliam, 1998a; Bryant & Carless, 2009; Cartney, 2010; Tamjid & Birjandi, 2011; Yang, 2011). Employing these assessments in the L2 classroom would provide a rich texture from which students can become confident and proficient learners.

Student perceptions and motivation will continue to be a common thread in the educational realm. As teachers move toward facilitation of learning and a learner-centered classroom, away from teacher-centered instruction, student perceived value will play an even larger role (Burkšaitienė & Teresevičienė, 2008). Student motivation in

correlation with alternative assessments is a topic that will warrant further investigation. Specifically, researchers should correlate their students' perceptions with a variety of alternative foreign language assessment scores. As a continuation of this study, researchers should continue to procure and evaluate student perceptions as a part of the assessment process.

Further investigations concerning alternative assessments are needed, especially after L2 educators are able to experience the outcomes of the CCSS and see, first-hand, the types of precise assessments being administered and their direct effect on the classroom teacher. As education continues to change through the years, more studies will be needed to validate assessment types and their effectiveness, as well as perceived value. Further study in the files of alternative assessments including a larger and more diverse population is recommended. Additionally, the creation of a Dynamic Assessment, Task-based Assessment, and Formative Assessment for all three units with the experimental group being subdivided into three sections could be advantageous. For each unit, the researcher would have perceptions and scores for each of the three alternative assessment types and scores for a traditional assessment from the control group. By doing so, the researcher would potentially acquire even more specific perceptions for the alternative assessments, without the possibility of researcher error in choosing the alternative assessment for each unit. Finally, since no surveys were administered to gauge the perceptions of the control group students taking traditional assessments for this study; no student buy-in can be presumed for this assessment type, suggesting that procuring the perceptions of the control group could be valuable as comparative data.

Teachers will continue to design assessments to meet their programs' outcomes but should always consider their students' perceptions, alternatives to traditional design, and the bonus effects of creating assessments that lend themselves to higher-order thinking and real-world application. Duncan (2014) purported the following:

> Assessment is the key to language learning. Only through the assessment lens can learners find out if they are meeting their goals, to what extent those goals are being met, and what they still need to do. Through the feedback received, our students' motivational fires are fed. (p. 19)

It is a primary responsibility as L2 educators to find and develop assessments for our students; Dynamic Assessment, Task-based Assessment, and Formative Assessment using self- and peer-evaluation should be strongly considered.

REFERENCES

Ableeva, R. (2008). The effects of dynamic assessment on L2 listening comprehension. In J. Lantolf & M. Poehner (Eds.), *Sociocultural Theory and the Teaching of second language*. London, England: Equinox.

Adair-Hauck, B., Glisan, E. W., Koda, K., Swender, E. B., & Sandrock, P. (2006). The integrated performance assessment (IPA): Connecting assessment to instruction and learning. *Foreign Language Annals, 39*(3), 359-382. doi:10.1111/j.1944-9720.2006.tb02894.x

Ahmadian, M. J. (2012). The relationship between working memory capacity and L2 oral performance under task-based careful online planning condition. *TESOL Quarterly, 46*(1), 165-175. doi:10.1002/tesq.8

Alliance for Excellent Education. (2011). A time for deeper learning: Preparing students for a changing world. *Policy Brief*. Retrieved from http://www.all4ed.org

Alonso Alonso, M. R., & Palacios Martínez, I. (2005). Evaluación del alumno, técnicas y elementos en el proceso de aprendizaje de segundas lenguas. *Encuentro, 15*, 5-13. Retrieved from http://www.encuentrojournal.org/textos/02_Alonso.pdf

American Council on the Teaching of Foreign Languages [ACTFL]. (1998). *Standards for foreign language learning: Preparing for the 21st century*. Retrieved from http://www.actfl.org/

Antón, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals, 43*(3), 576-599. doi:10.1111/j.1944-9720.2009.01030.x

Asghar, A. (2010). Reciprocal peer coaching and its use as a Formative Assessment strategy for first-year students. *Assessment & Evaluation in Higher Education, 35*(4), 403-417. doi:10.1080/02602930902862834

Bakar, K. A., Sulaiman, N. F., Akhtar, Z., & Rafaai, M. (2010). Self-determination theory and motivational orientations of Arabic learners: A principal component analysis. *GEMA Online Journal of Language Studies, 10*(1), 71-86.

Birjandi, P., & Tamjid, N. H. (2010). The role of self-assessment in promoting Iranian EFL learners' motivation. *English Language Teaching, 3*(3), 211-220. Retrieved from http://www.ccsenet.org/journal/index.php/elt/article/view/7234/5586

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom *Phi Delta Kappan, 86*(1), 8-21. Retrieved from http://web.uvic.ca/~gtreloar/Assessment

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-61. Retrieved from http://wiki.biologyscholars.org/@api/deki/files/1278/=black_p_etal_1998b%255b1%255d.pdf

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 139-148. Retrieved from http://blog.discoveryeducation.com/assessment/files/2009/02/blackbox_article.pdf

Blanche, P. (1988). Self-assessment of foreign language skills: Implications for teachers and researchers. *RELC Journal, 19*(1), 75-93. doi:10.1177/003368828801900105

Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. *Language Learning, 39*(3), 313-338. doi:10.1111/j.1467-1770.1989.tb00595.x

Block, D. (2004). Review of Rod Ellis' task-based language learning and teaching. *Language Learning Journal, 29*, 18-22.

Blue, G. (1994). Self-assessment of foreign language skills: Does it work? *CLE Working Papers, 3*, 18-35. Retrieved from http://files.eric.ed.gov/fulltext/ED396569.pdf

Boud, D. (2007). Reframing assessment as if learning were important. In D. Boud & N. Falchikov (Eds.), *Rethinking Assessment in Higher Education: Learning for the longer term*. New York, NY: Routledge.

Boud, D., & Lee, A. (2005). "Peer learning" as pedagogic discourse for research education. *Studies in Higher Education, 30*(5), 501-516. doi:10.1080/03075070500249138

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education.

Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32*(4), 653-675. doi:10.2307/3587999

Bryant, D., & Carless, D. (2009). Peer assessment in a test-dominated setting: Empowering, boring or facilitating examination preparation? *Educational Research for Policy and Practice, 9*(1), 3-15. doi:10.1007/s10671-009-9077-2

Budoff, M. (1968). Learning potential as a supplementary testing procedure. In J. Helmuth (Ed.). *Learning Disorders, 3* (pp. 295-343). Seattle, WA: Special Child.

Burkšaitienė, N., & Teresevičienė, M. (2008). Integrating alternative learning and assessment in a course of English for law students. *Assessment & Evaluation in Higher Education, 33*(2), 155-166. doi:10.1080/02602930601125699

Butler, Y. G. (2011) The implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annual Review of Applied Linguistics, 31*, 36-57. doi:10.1017/S0267190511000122

Büyükkarci, K. (2009). A critical analysis of task-based learning. *Kastamonu Education Journal, 17*(1), 313-320. Retrieved from http://www.kefdergi.com/pdf/cilt17_no1/313.pdf

Bygate, M., Skehan, P., & Swain, M. (Eds.) (2001). *Research pedagogic tasks, second language learning, teaching and testing*. Harlow, England: Longman.

Byrnes, H. (2002). The role of task and Task-based Assessment in a content-oriented collegiate foreign language curriculum. *Language Testing, 19*(4), 419-437. doi:10.1191/0265532202lt238oa

Campione, J. C., & Brown, A. L. (1985). Dynamic assessment: One approach and some initial data. *Center for the Study of Reading, 361*, 1-56. Retrieved from https://www.ideals.illinois.edu

Carless, D. (2007). The suitability of task-based approaches for secondary schools: Perspectives from Hong Kong. *System, 35*(4), 595-608. doi:10.1016/j.system.2007.09.003

Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment & Evaluation in Higher Education, 35*(5), 551-564. doi:10.1080/02602931003632381

Cheng, H.-F., & Dörnyei, Z. (2007). The use of motivational strategies in language instruction: The case of EFL teaching in Taiwan. *Innovation in Language Learning and Teaching, 1*(1), 153-174. doi: 10.2167/illt048.0

Chicago Tribune. (2012, October 30). Illinois school report cards. Retrieved from http://schools.chicagotribune.com/

Colby-Kelly, C., & Turner, C. (2007). AFL research in the L2 classroom and evidence of usefulness: Taking formative assessment to the next level. *The Canadian Modern Language Review, 64*(1), 9-38. doi:10.3138/cmlr.64.1.009

Coronado-Aliegro, J. (2006). *The effect of self-assessment on the self-efficacy of students studying Spanish as a foreign language*. (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh, PA. Retrieved from http://challenger.library.pitt.edu/ETD/available/etd-12012006-121145/unrestricted/Coronado-AliegroJ2_etd2006.pdf

Deakin-Crick, R., Lawson, H., Sebba, J., Harlen, W., & Yu, G. (2005) Systematic review of research evidence of the impact on students of self- and peer-assessment. Protocol. In *Research Evidence in Education Library*. London, England: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Retrieved from http://eppi.ioe.ac.uk/eppiwebcontent/reel/review_groups/assessment/assessment_protocol5.pdf

Downing, R. (2012). The development of plurilingual competence through authentic assessment and self-assessment: Case study. *VIGO International Journal of Applied Linguistics, 9*, 63-84. Retrieved from http://webs.uvigo.es/vialjournal/pdf/Vial-2012-Article3.pdf

Duncan, G. (2014). Embracing the paradigm shift in learning and assessment. *The Language Educator, 9*(2), 18-19.

East, M. (2012). *Task-based language teaching from the teachers' perspective*. Amsterdam, The Netherlands: John Benjamin Publishing.

Eberly Center for Teaching Excellence. (n. d.). *Whys and hows of assessment*. Retrieved from http://www.cmu.edu/teaching/assessment/basics/formative-summative.html

Education Northwest. (n.d.). *Common core state standards*. Retrieved from http://educationnorthwest.org/common-core

Ellis, R. (1997). *Second language acquisition.* H. G. Widdowson (Ed.). Oxford, UK: Oxford University Press.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research, 4*(3), 193-220.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University Press.

Ellis, R. (2006). The methodology of task-based teaching. *Asian EFL Journal, 8*(3). Retrieved from http://www.asian-efl-journal.com/Sept 6 re.php

Feuerstein, R., Feuerstein, S., Falik, L., & Rand, Y. (1979). *Dynamic assessments of cognitive modifiability*. Jerusalem, Israel: ICELP Press.

Feuerstein, R., Rand, Y., Jensen, M., Kaniel, S., & Tzuriel, D. (1987). Prerequisites for assessment of learning potential: The LPAD model. In C. Lidz (Ed.), *The Dynamic Assessment*. New York, NY: Guilford Press.

Foster, P. (1999). Task-based learning and pedagogy. *ELT Journal, 53*, 69-70.

Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition, 18*, 299-324.

Foster, P., & Skehan, P. (1997). Modifying the task: The effects of surprise, time and planning type on task based foreign language instruction. *Thames Valley University Working Papers in English Language Teaching, 4*, 86-109.

Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research, 3*(3), 215-247.

Funston, J. F., & Vargas Bonilla, A. (2005). *Navegando 1*. St. Paul, MN: EMC.

Garb, E. (1997). *Dynamic assessment as a teaching tool*. Retrieved from http:www.etni.org

Gardner, R. C. (1985). *School psychology and second language learning: The role of attitudes and motivation.* London, England: Edward Arnold.

Gay, L. R., Mills, G. E., & Airasian, P. (2012). *Educational research: Competencies for analysis and application* (10th ed.). Boston, MA: Pearson.

Geeslin, K. L. (2012). Student self-assessment in the foreign language classroom: The place of authentic assessment instruments in the Spanish language classroom. *Hispania, 86*(4), 857-868. Retrieved from http://www.jstor.org/stable/20062958

Hamidi, E. (2010). Fundamental issues in L2 classroom assessment practices. *Academic Leadership Journal*, *8*(2). Retrieved from http://www.sisd.net/cms/lib/TX

Hanrahan, S. J., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research & Development, 20*(1), 53-70. doi:10.1080/07294360123776

Hashemia, M., Azizinexhadb, M., & Darvishi, S. (2012). Using task-based language teaching, learning practically in English classes. *Procedia- Social and Behavioral Sciences, 31*, 526-529. doi:10.1016/j.sbspro.2011.12.098

Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in Practice: Clinical and Educational Application.* New York, NY: Cambridge University Press.

Haywood, H. C., & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education, 77*(2), 40-63. doi:10.1207/S15327930PJE7702_5

Hedgcock, J., & Lefkowitz, N. (1992). Collaborative oral/aural revision in foreign language writing instruction. *Journal of Second Language Writing, 1*, 255-276.

Heining-Boynton, A. L., & Redmond, M. L. (2013). The common core framework and world languages: A wake-up call for all. *The Language Educator, 8*(1), 52-56. Retrieved from http://www.actfl.org/sites/default/files/pdfs/TLE_pdf/TLE_Jan13_Article.pdf

Hill, K., & Sabet, M. (2009). Dynamic speaking assessments. *TESOL Quarterly, 43*, 537-545.

Hismanoglu, M., & Hismanoglu, S. (2011). Task-based language teaching: What every EFL teacher should do. *Procedia- Social and Behavioral Sciences, 15*, 46-52. doi:10.1016/j.sbspro.2011.03.049

Holec, H. (1979). *Autonomy in foreign language learning*. Oxford: Pergamon Press for the Council of Europe.

Illinois State Board of Education. (2012). Realizing Illinois. Retrieved from http://www.isbe.net/common_core/default.htm.

Illinois school report cards. (2012). *Chicago Tribune*. Retrieved from http://schools.chicagotribune.com/

Jabbarpoor, S. (2011).Task-based language assessment: Difficulty-based task items sequencing and task-based test performance. *IPEDR, 26*, 219-223. Retrieved from http://www.ipedr.com/vol26/45-ICLLL%202011-L10056.pdf

Jeon, I., & Hahn, J. (2006). Exploring EFL teachers' perceptions of task-based language teaching: A case study of Korean secondary school classroom practice. *Asian EFL Journal Quarterly, 8*(1), 123-143. Retrieved from http://asian-efl-journal.com/March_06.pdf#page=123

Johnson, K. (2000). What task designers do. *Language Teaching Research, 4*(3), 301-321. doi:10.1177/136216880000400306

Kaufman, A., & Dodge, T. (2009). Student perceptions and motivation in the classroom: Exploring relatedness and value. *Social Psychology of Education, 12*, 101-112. doi:10.1007/s11218-008-9070-2

Ke, C. (2006). A model of formative task-based language assessment for Chinese as a foreign language. *Language Assessment Quarterly*, *3*(2), 207-227. doi:10.1207/s15434311laq0302

Kim, H. (2002). Task-based performance assessment for teachers: Key issues to consider. *TESOL & Applied Linguistics*, *4*(2), 1-5. Retrieved from http://www.tc.columbia.edu/academic/tesol/WJFiles/pdf/HyunjooForum.pdf

Klapper, J. (2003). Taking communication to task? A critical review of recent trends in language teaching. *Language Learning Journal*, *27*, 33-42.

Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International*, *23*(1), 112-127. doi:10.1177/0143034302023001733

Kozulin, A., & Garb, E. (2004). Dynamic assessment of literacy: English as a third language. *European Journal of Psychology of Education*, *19*(1), 65–77. doi:10.1007/BF03173237

Krashen, S. D. (1981). *Second language acquisition and second language learning.* Oxford, UK: Pergamon Press.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford, UK: Pergamon Press.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications.* London, England: Longman.

Lantolf, J. P. (2009). Dynamic assessment: The dialectic integration of instruction and assessment. *Language Teaching*, *42*(3), 355-368. doi:10.1017/S0261444808005569

Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: bringing the past into the future. *Journal of Applied Linguistics, 1*(1), 49-72. doi:10.1558/japl.1.1.49.55872

Lantolf, J. P., & Poehner, M. E. (2006). *Dynamic assessment in the foreign language classroom: A teacher's guide*. University Park, PA: CALPER.

Lantolf, J. P., & Poehner, M. E. (2010). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research, 15*(1), 11-33. doi:10.1177/1362168810383328

Lantolf, J. P., & Poehner, M. E. (2011). *Dynamic assessment in the foreign language classroom. A teacher's guide*. (2nd ed.) University Park, Pennsylvania: The Pennsylvania State University, Center for Advanced Language Proficiency Education and Research.

Lee, I. (1998). Supporting greater autonomy in language learning. *ELT Journal, 52*(4), 282-290. doi:10.1093/elt/52.4.282

Lee, J. (2000). *Tasks and communicating in language classrooms*. Boston, MA: McGraw-Hill.

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research: Planning and design* (10th ed.). Boston, MA: Pearson.

Lindblom-Ylänne, S., Pihlajamäki, H., & Kotkas, T. (2006). Self-, peer-, and teacher-assessment of student essays. *Active Learning in Higher Education, 7*(1), 51-62. doi:10.1177/1469787406061148

Little, D. G. (1991). *Learner autonomy: Definitions, issues and problems* (Vol. 1). Dublin, Ireland: Authentik Language Learning Resources.

Little, D. (2005). The common European framework and the European language portfolio: Involving learners and their judgements in the assessment process. *Language Testing, 22*(3), 321–336. doi:10.1191/0265532205lt311oa

Littlewood, W. (2004). The task-based approach: Some questions and suggestions. *ELT Journal, 58*(4), 319-326.

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching, 40*(3), 241-249.

Liu, N., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education, 11*(3), 279-290. doi:10.1080/13562510600680582

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.). *Modeling and Assessing Second Language Acquisition*. Bristol, UK: Multilingual Matters.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.). *Handbook of Second Language Acquisition*. (pp. 413-468). San Diego: Academic Press.

Long, M. H. (1997). *Focus on form in task-based language teaching*. Paper presented at the Fourth Annual McGraw-Hill Satellite Teleconference. Abstract retrieved from http://www.mhhe.com/socscience/foreignlang/conf/option3.htm

Lunt, I. (1993). The practice of assessment. In H. Daniels (Ed.). *Charting the Agenda: Educational Activity after Vygotsky.* London, England: Routledge.

Machemer, P. L., & Crawford, P. (2007). Student perceptions of active learning in a large cross-disciplinary classroom. *Active Learning in Higher Education, 8*(1), 9-30. doi:10.1177/1469787407074008

MacIntyre, P., MacKinnon, S., & Clement, R. (2009). In D. Zoltan & E. Ushioda (Eds.), *Motivation, Language identity, and the L2 Self,* (pp. 193-214). Bristol, UK: Multilingual Matters.

Mandell, P. (1999). On the reliability of grammaticality judgement tests in second language acquisition research. *Second Language Research, 15*(1), 73-99. doi:10.1191/026765899667538959

Mardani, M., & Tavakoli, M. (2011). Beyond reading comprehension: The effect of adding a dynamic assessment component on EFL reading comprehension. *Journal of Language Teaching and Research, 2*(3), 688-696. doi:10.4304/jltr.2.3.688-696

Marzano Center, The. (2015). Common core and the art and science of teacher series. Retreived from http://www.marzanocommoncore.com/

Meurers, D., Ott, N., & Ziai, R. (2010). Compiling a task-based corpus for the analysis of learner language in context. *Proceedings of Linguistic Evidence, Tübingen*, 214-217. Retrieved from http://www.sfs.uni-tuebingen.de/~dm/papers/meurers-ott-ziai-10.pdf

Morrow, L. M., Shanahan, T., & Wixson, K. K. (Eds.). (2012). *Teaching with the common core standards for English language arts*. New York, NY: Guilford Press.

Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. New York, NY: Taylor & Francis.

National Capital Language Resource Center. (n.d.). The essentials of language teaching. Retrieved from http://nclrc.org/essentials.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common core state standards initiative.* Retrieved from http://www.corestandards.org/

Norris, J. M. (2009). Task-based teaching and testing. In M. H. Long & C. J. Doughty (Eds.). *The Handbook of Language Teaching* (pp. 578-594). Singapore, Malaysia: Blackwell.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge, UK: Cambridge University Press.

Oscarson, M. (1978). *Approaches to self-assessment in foreign language learning.* Strasbourg, France: Council of Europe.

Oscarson, M. (1984). *Self-Assessment of Foreign Language Skills: A Survey of Research and Development Work*. Strasbourg, France: Council of Europe.

Oscarson, M. (1998, Nov.). Learner Self-Assessment of Language Skills. *IATEFL TEA SIG Newsletter*. Retrieved from http://rcel.enl.uoa.gr/kpg/texts/IATEFL%20TEA%20SIG-EALTA%20Barcelona%20Conference%20Proceedings.pdf

Palinscar, A., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*, 117-175. Retrieved from http://people.ucsc.edu/~gwells/Files/Courses_Folder/ED%20261%20Papers/Palincsar%20Reciprocal%20Teaching.pdf

Palloff, R. M., & Pratt, K. (2007). *Building online learning communities: Effective strategies for the virtual classroom*. San Francisco, CA: Jossey-Bass.

Peña, E., Quinn, R., & Iglesias, A. I. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. *Journal of Special Education*, *26*(3), 269-280. doi:10.1177/002246699202600304

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 451-502). San Diego, CA: Academic.

Pintrich, R. R. (2003). A motivational science perspective on issues in motivation terminology, theory, and research. *Journal of Educational Psychology, 95*, 667-686. doi:10.1037/0022-0663.95.4.667

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, *9*(3), 233–265. doi:10.1191/1362168805lr166oa

Poehner, M. E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *Modern Language Journal*, *91*(3), 323-340. doi:10.1111/j.1540-4781.2007.00583.x

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. New York, NY: Springer.

Poehner, M. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly*, *43*(3), 471-491. doi:10.1002/j.1545-7249.2009.tb00245.x

Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's teaching-assessment dialectic and L2 education: The case for dynamic assessment. *Mind, Culture, and Activity, 17*, 312-330.

Poehner, M. E., & van Compernolle, R. A. (2011). Frames of interaction in dynamic assessment: Developmental diagnoses of second language learning. *Assessment in Education: Principles, Policy & Practice*, *18*(2), 183-198. doi:10.1080/0969594X.2011.567116

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011). Common core standards: The new U. S. intended curriculum. *Educational Researcher, 40*(3), 103-116. doi:10.3102/0013189X11405038

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford, England: Oxford University Press.

Richards, J., Platt, J., & Weber, H. (1985). *Longman dictionary of applied linguistics*. London, England: Longman.

Richards, J., & Rodgers, T. (2001). *Approaches and methods in language teaching*. Cambridge, England: Cambridge University Press.

Riedinger, B. (2006). *Mining for meaning: Teaching students to reflect*. In A. Jafari, & C. Kaufman (Eds.). *Handbook of research on eportfolios* (pp. 90-101). Hershey, PA: Idea Group.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring 76. interactions in a componential framework. *Applied Linguistics, 22*(1), 27-57.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45-105. Retrieved from http://www.hawaii.edu/sls/sls/wp-content/uploads/2011/06/Robinson.pdf

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching, 45*(3), 193-213.

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching, 45*(3), 161-176.

Robson, C. (2011). *Real world research* (3rd ed.). Chichester, UK: Wiley.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing Journal, 15*(1), 1-20. doi:10.1177/026553229801500101

Ross, S. J. (2005). The impact of assessment method on foreign language proficiency growth. *Applied Linguistics, 26*(3), 317-342. doi:10.1093/applin/ami011

Rueda, R., & Chen, C. B. (2005). Assessing motivational factors in foreign language learning: Cultural variation in key constructs. *Educational Assessment, 10*(3), 209-229. doi:10.1207/s15326977ea1003_4

Sadeghi, K., & Khanahmadi, F. (2011). Dynamic assessment of L2 grammar of Iranian EFL learners: The role of mediated learning experience. *International Journal of Academic Research, 3*(2), 931-937. Retrieved from http://scholar.googleusercontent.com/scholar?q=cache:sHQXjdWcI_wJ:scholar.google.com/+iranian+efl+learners+sadeghi&hl=es&as_sdt=0&as_ylo=2011&as_yhi=2012

Schunk, D. H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist, 40*(2), 85-94. Retrieved from http://libres.uncg.edu/ir/uncg/f/D_Schunk_Self_2005.pdf

Schütz, R. (2007). Stephen Krashen's theory of second language acquisition. *English Made in Brazil*. Retrieved from http://www.sk.com.br/sk-krash.html

Sebba, J., Crick, R. D., Yu, G., Lawson, H., Harlen, W., & Durant, K. (2008). Systematic review of research evidence of the impact on students in secondary schools of self and peer assessment. Technical report. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Retrieved from http://www.educationscotland.gov.uk/Images/Systematic%20review%20of%20research%20evidence_tcm4-529527.pdf

Shrum, J. L., & Glisan, E. W. (2010). *Teacher's handbook: Contextualized language instruction*. Boston, MA: Heinle, Cengage Learning.

Sidek, H. M. (2012). EFL reading instruction: Communicative task-based approach. *International Journal of Instruction, 5*(2), 109-128. Retrieved from http://www.e-iji.net/dosyalar/iji_2012_2_7.pdf

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38-62. doi:10.1093/applin/17.1.38

Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics, 18*, 268-286.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*(3), 185-211. doi:10.1177/136216889700100302

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93-120.

Speece, D. L., Cooper, D. H., & Kibler, J. M. (1990). Dynamic assessment, individual differences, and academic achievement. *Learning and Individual Differences, 2*(1), 113-127. doi:10.1007/978-0-387-30424-3_181

Sugita, Y. (2009). The development and implementation of task-based writing performance assessment for Japanese learners of English. *Pan-Pacific Association of Applied Linguistics, 13*(2), 77–103. Retrieved from http://www.eric.ed.gov/PDFS/EJ868853.pdf

Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics, 26*(3), 376-401.

Tamjid, N. H., & Birjandi, P. (2011). Fostering learner autonomy through self- and peer-assessment. *International Journal of Academic Research, 3*(5), 245-252.

Tavakoli, P., & Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 61*(s1), 37-72. doi:10.1111/j.1467-9922.2011.00642.x

Tulung, G. (2008). On examining communicative tasks in second language acquisition. *TEFLIN Journal, 19*(2), 105-118. Retrieved from http://pdfsdb.com/pdf/on-examining-communicative-tasks-in-second-language-acquisition-41138220.html

VanPatten, B., & Williams, J. (2007). *Theories in second language acquisition: An introduction*. Mahwah, NJ: Lawrence Erlbaum.

Vardi, I., & Bunker, A. (2001). The relationship between the perceived value of supplementary online components, and student roles and responsibilities. Paper presented at the Meeting at the Crossroads. Proceedings of the 18th Annual Conference of the Australasian society for Computers in Learning in Tertiary Education (ascilite), Melbourne, Australia.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wei, L. (2011). Formative Assessment in classrooms: Operational procedures. *Journal of Language Teaching and Research*, *2*(1), 99-103. doi:10.4304/jltr.2.1.99-103

Willis, D., & Willis, J. (2007). *Doing task-based teaching.* Oxford, UK: Oxford University Press.

Yang, Y. F. (2011). A reciprocal peer review system to support college students' writing. *British Journal of Educational Technology*, *42*(4), 687-700. doi:10.1111/j.1467-8535.2010.01059.x

Appendix

Student Perceptions Surveys

Student Perceptions of Assessment Blue

My Username_____

Grade on Assessment Blue_____ (The teacher will complete this line.)

Please circle one answer for each question.

Circle the answer that best fits how you feel.

1. I was able to show the teacher how much I knew on Assessment Blue.

Strongly Agree Agree Disagree Strongly Disagree

2. I would like my teacher to use Assessment Blue again this year.

Strongly Agree Agree Disagree Strongly Disagree

Student Perceptions of Assessment Red

My Username_____

Grade on Assessment Red_____ (The teacher will complete this line.)

Please circle one answer for each question.

Circle the answer that best fits how you feel.

1. I was able to show the teacher how much I knew on Assessment Red.

Strongly Agree Agree Disagree Strongly Disagree

2. I would like my teacher to use Assessment Red again this year.

Strongly Agree Agree Disagree Strongly Disagree

Student Perceptions of Assessment Green

My Username_____

Grade on Assessment Green_____ (The teacher will complete this line.)

Please circle one answer for each question.

Circle the answer that best fits how you feel.

1. I was able to show the teacher how much I knew on Assessment Green.

Strongly Agree Agree Disagree Strongly Disagree

2. I would like my teacher to use Assessment Green again this year.

Strongly Agree Agree Disagree Strongly Disagree

Student Perceptions of Assessments Blue, Red, and Green

My Username_____

Please answer each question as completely as possible.

1. Which Assessment was the best reflection of my knowledge,

   Blue, Red, or Green?

Why? Try to give specific details.

2. Which Assessment was the weakest reflection of my knowledge,

   Blue, Red, or Green?

Why? Try to give specific details.