

ABSTRACT

LATHERS, JOHN DANIEL. Designing an Instrument to Identify Instructor Characteristics and Student Reactions in U.S. Marine Corps Vocational Training. (Under the direction of James Bartlett, II.)

The purpose of the study was to examine generic instructor characteristics capable of predicting learning outcomes for application to the evaluation process of formalized trainers within U.S. Marine Corps vocational specialty courses. A survey instrument was created based on research findings from multiple training disciplines and educational settings to assess student's reaction to perceived instructor quality and associated performance. The study's methodology implemented a non-experimental research design. A nonrandom purposive sample of participants selected from entry-level U.S. Marines currently undergoing vocational training was used for the study. Data analysis techniques included exploratory factor analysis. Means, standard deviation, frequencies, and percents were computed for descriptive data. Step-wise multiple regression analysis was used to examine the relationships between instructor characteristics and student's perception of instructor quality and student learning, as evidenced by post-lesson test scores.

By examining participant ratings of perceived instructor characteristics, the initial findings allowed for the creation of an instrument. Subsequent findings indicate that participants

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE

2010

2. REPORT TYPE

3. DATES COVERED

00-00-2010 to 00-00-2010

4. TITLE AND SUBTITLE

Designing an Instrument to Identify Instructor Characteristics and Student Reactions in U.S. Marine Corps Vocational Training.

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

North Carolina State University, Raleigh, NC, 27695

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The purpose of the study was to examine generic instructor characteristics capable of predicting learning outcomes for application to the evaluation process of formalized trainers within U.S. Marine Corps vocational specialty courses. A survey instrument was created based on research findings from multiple training disciplines and educational settings to assess student's reaction to perceived instructor quality and associated performance. The study's methodology implemented a non-experimental research design. A nonrandom purposive sample of participants selected from entry-level U.S. Marines currently undergoing vocational training was used for the study. Data analysis techniques included exploratory factor analysis. Means, standard deviation, frequencies, and percents were computed for descriptive data. Step-wise multiple regression analysis was used to examine the relationships between instructor characteristics and student's perception of instructor quality and student learning, as evidenced by post-lesson test scores. By examining participant ratings of perceived instructor characteristics, the initial findings allowed for the creation of an instrument. Subsequent findings indicate that participants positively rated their instructors on effective delivery and concern for learning while concurrently rating instructors less high on negativity items. Participants also rated their instructor positively in overall instructor quality. Post-lesson test scores indicate that nearly 90% of participants mastered the lesson objectives, as taught by the rated instructor, at the outset without remediation. Data analysis revealed correlations that indicate that none of the instructor characteristics had significant relationships with, or could explain a significant amount of variance of, student's perception of instructor quality or post-lesson test scores. The study's findings question the validity of employing Donald Kirkpatrick's evaluation model, and more specifically, the student reaction phase, to measure vocational training effectiveness when perceived instructor characteristics are unrelated to instructor quality, and more importantly, student performance.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 295	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Standard Form 298 (Rev. 8-98)
 Prescribed by ANSI Std Z39-18

positively rated their instructors on effective delivery and concern for learning while concurrently rating instructors less high on negativity items. Participants also rated their instructor positively in overall instructor quality. Post-lesson test scores indicate that nearly 90% of participants mastered the lesson objectives, as taught by the rated instructor, at the outset without remediation. Data analysis revealed correlations that indicate that none of the instructor characteristics had significant relationships with, or could explain a significant amount of variance of, student's perception of instructor quality or post-lesson test scores. The study's findings question the validity of employing Donald Kirkpatrick's evaluation model, and more specifically, the student reaction phase, to measure vocational training effectiveness when perceived instructor characteristics are unrelated to instructor quality, and more importantly, student performance.

© Copyright 2010 by John Daniel Lathers

All Rights Reserved

Designing an Instrument to Identify Instructor Characteristics
and Student Reactions in U.S. Marine Corps Vocational Training

by
John Daniel Lathers

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Education

Adult and Higher Education
Raleigh, North Carolina

2010

APPROVED BY:

Dr. Diane D. Chapman

Dr. D. Jan W. Lucas

Dr. Julia Storberg-Walker

Dr. James E. Bartlett, II
Chair of Advisory Committee

DEDICATION

I humbly dedicate this dissertation to my children for their unceasing inspiration and encouragement to first begin and then follow through to achieve my goals - all the while suffering the brunt of a distracted father. My oldest John, who has already far surpassed his father in capabilities; Alexendar who never ceases to make me proud by his sincerity and well-defined sense of humor; my daughter, Gabrielle, my brilliant sunshine, whose intellect and dedication often leaves me speechless; Nicholas with his intense drive and complex wit; and Nathan who interests, skills and abilities frequently out shadow his peers.

BIOGRAPHY

John Daniel Lathers, born August 25, 1961, is the son of Lawrence Lathers and the former Evelyn McCune of Ypsilanti, Michigan. John attended public schools in Washtenaw County, Michigan and graduated from Ypsilanti High School in 1979.

Upon graduation, John enlisted in the U.S. Marine Corps to serve as a Marine Air Traffic Control Navigational Systems Repair Technician, but also performed a variety of other vocations including: micro-miniature electronics repair, airfield operations, family resource center office manager, U.S. Marine Corps recruiter, CT-39 Saberliner jet-aircraft plane captain, technical formal school instructor, and curriculum course manager. He retired as a Master Sergeant of Marines in 2000, after 21 years of honorable service.

Concurrent with military deployments and frequent relocations, John also worked after hours as an adjunct professor, septic tank builder and drain field installer, short order cook and waiter, graphic artist, yacht engine repair and dock hand, painter, roofer, house builder and ceramic tile installer, salvage operator, and fishing-pier pylon installer.

While serving on active-duty, John graduated from the University of North Carolina at Wilmington, in 1991, with a Bachelor of Arts with honors in Psychology. His undergraduate dissertation was titled: *Implicit and Explicit Memory Retention in the Presence of an Olfactory Cue*. In 1993, he graduated from Webster University with a Master of Arts in Counseling. After retiring from active-duty service, John graduated from North Carolina State University in 2005, having earned a Master of Education degree.

In 2000, John became an instructional systems supervisor and training specialist with the U.S. Marine Corps Combat Service Support Schools as a civilian federal employee. Since then, he has assisted the formal schools with the creation of academic policy and the functional design of training facilities, staff development and training, as well as the establishment of academic training standards used in the analysis, development, implementation and evaluation of the vocational training of entry-, career- and advanced-level U.S. Marines, both active and reserve, as well as service members from other U.S. Armed Forces and international military officers.

John Daniel Lathers married the former Belinda Sue Kellum of Jacksonville, North Carolina on 25 November, 1989. They have four children, Alexandar Michael, Gabrielle Mariah, Nicholas Scott and Nathan Stuart, ages 17, 15, 14 and 14, respectfully. They currently live in Jacksonville, North Carolina with their two golden retrievers and rescued Persian cat.

ACKNOWLEDGEMENTS

I am indebted to a great many people for their contributions to this study. First and foremost, my considerable appreciation and heartfelt thanks is given to my dissertation chair and mentor, Dr. James Bartlett II, whose knowledge, practicality, and genuine and sincere encouragement made this study possible. Special thanks are also due to Dr. Michelle Bartlett for her considerable expertise, critical eye, and thoughtfulness for allowing me to often invade the peaceful sanctity of her (and James') home.

I would also like to thank my other committee members for their time and support during this lengthy process; Dr. Diane Chapman who has suffered my questions and puzzlement the longest, Dr. Jan Lucas who provided invaluable CTE insight that has benefited me greatly in the military training arena, and Dr. Julie Storberg-Walker who was surprised that I was funnier than I appeared.

The encouragement and support of the professionals and students at Marine Corps Combat Service Support Schools deserve generous appreciation for their support of this study. My immediate supervisor, Lieutenant Colonel Kevin M. Barth, USMC (ret.), was especially supportive and accommodating in my academic and professional pursuits. The Marine Corps base

Commanding Officer, Colonel Richard Fullerton, USMC, and many of his subordinate officers, were also particularly supportive and cooperative with the data collection process.

Also considerable thanks goes out to my co-workers and friends - Mr. Robert Fahy, Mrs. Tiffany Green, Mr. John Jones, Mrs. Gail Ward-Warner, Mr. Dan Towler, Mrs. Heidi Rawson, Mr. Joe Lomangino, Mr. Shane Culbreath and Mr. Gary Lambertsen - for suffering my rough theoretical constructs, formative conjectures, and unrelenting personal anecdotes and stories. I couldn't have completed this process without their straightforward and impermeable honesty.

The overwhelming majority of my thanks goes out to my wonderful family; my wife, Belinda, for her commitment during the many long commutes and the many hours locked away researching and typing, and my unique children, John, Alex, Gabby, Nick and Nate, who never cease to amaze me - as children often do.

TABLE OF CONTENTS

LIST OF TABLES.....	xii
LIST OF FIGURES.....	xiv
Chapter 1	NATURE OF THE PROBLEM..... 1
	Statement of the Problem..... 4
	The Purpose of the Study..... 6
	Research Objectives..... 7
	Research Questions..... 8
	Research Methodology..... 9
	Theoretical Framework..... 10
	Conceptual Framework..... 12
	Significance of the Study..... 16
	Delimitations..... 17
	Organization of the Study..... 18
Chapter 2	REVIEW OF LITERATURE..... 21
	Literature Review of Student Evaluation..... 22
	Student Satisfaction..... 22
	Consumer Model..... 23
	Criticism of the Consumer Model..... 27
	Correlation of Student Ratings..... 31
	Perception of Effective Teaching..... 33
	Effect on Instructor Performance..... 38
	Student Ratings on Improving Teaching..... 43
	Theoretical Framework..... 49
	Humanism..... 51
	Adult Education..... 53
	Andragogy..... 54
	Social Learning..... 60

Table of Contents (Continued)

	Behaviorism.....	62
	Cognitivism.....	64
	Systematic Instructional Design.....	66
	Summary of Theoretical Framework.....	74
	Background Information - The U.S. Armed Forces.....	74
	Secondary Societal Purpose.....	78
	Myth of the Representative Population.....	79
	Enlistment Process.....	83
	Qualifications.....	85
	Attrition.....	86
	ASVAB and AFQT.....	88
	Occupational Specialty.....	92
	Training - Initial.....	93
	Training - Vocational.....	94
	Formal School Vocational Training.....	95
	Military Evaluation of Training.....	96
	The Role of the Military Vocational Instructor.....	98
	Summary of Literature Review.....	100
Chapter 3	METHODOLOGY.....	102
	Research Design.....	102
	Research Methods.....	105
	Research Questions.....	106
	Research Variables Used in the Study.....	107
	Population.....	109
	Sample and Sample Size.....	110

Table of Contents (Continued)

	Instrument Development - Item Generation.....	111
	Factor Naming.....	118
	Data Collection.....	125
	Survey Administration.....	126
	Statistical Data Analysis.....	128
	First Round Field Test.....	132
	Demographics.....	132
	Addressing Missing Responses.....	134
	Identification of Case Outliers.....	134
	Exploratory Factor Analysis.....	137
	Revised Factor Naming.....	153
	Reliability of Instrument.....	155
	Summary of Methodology.....	159
Chapter 4	ANALYSIS OF DATA.....	162
	Demographics.....	162
	Summary of Participant Demographics.....	165
	Research Questions.....	165
	Revised Instrument Analysis.....	168
	Revised Instrument Item Outliers.....	169
	Exploratory Factor Analysis.....	170
	Final Instrument Factor Naming.....	175
	Post-Lesson Test Scores.....	180
	Summary of Analysis.....	186
Chapter 5	SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS....	189
	Purpose of the Study.....	190
	The Goal of the Study.....	190

Table of Contents (Continued)

Data Collection.....	191
Summary of Participant Demographics.....	191
Summary of U.S. Marine Corps Vocational Instructors.....	192
Research Objectives.....	193
Summary of Chapters.....	195
Findings and Conclusions by Research Question.....	199
Overall Conclusions and Recommendations.....	208
Limitations.....	214
Recommendations for Future Research.....	215
REFERENCES.....	218
APPENDICES.....	244
Appendix A. Item Deletion List.....	245
Appendix B. Initial Factors, Sub-factors, and Items..	247
Appendix C. Script for Data Collection.....	255
Appendix D. Institutional Review Board Approval.....	266
Appendix E. Informed Consent to Release Form for Research.....	267
Appendix F. Initial Instrument.....	268
Appendix G. Frequencies and Percents of Round One Participants' Age, Gender, Education Level, and Ethnicity.....	270
Appendix H. Final Instrument.....	271
Appendix I. Revised Factors, Sub-factors, and Items..	272
Appendix J. Final Factors, Sub-factors, and Items....	276

LIST OF TABLES

Table 2.1	Educational Tiers of Active Duty Enlistees to Civilian Comparison Group.....	82
Table 2.2	Armed Forces Qualification Test (AFQT) Percentile Scores for Active Duty Enlistees and a Civilian Comparison Group.....	91
Table 3.1	Variables Used to Describe Instructors and Explain Student Reactions of Instructor Quality and Learning.....	109
Table 3.2	Statistical Analysis.....	129
Table 3.3	Cases Discarded as Outliers Due to Non-Response..	135
Table 3.4	Items Discarded as Outliers Due to Non-Response..	136
Table 3.5	Exploratory Factor Analysis of Initial Instrument.....	139
Table 3.6	Exploratory Factor Analysis Oblimin Rotation.....	141
Table 3.7	Factor Loading Coefficients within Eight Fixed Factors.....	145
Table 3.8	Exploratory Factor Analysis with Eight Fixed Factors.....	146
Table 3.9	Factor Loading Coefficients within Five Fixed Factors.....	148
Table 3.10	Exploratory Factor Analysis with Five Fixed Factors.....	149
Table 3.11	Factor Loading Coefficients within Three Fixed Factors.....	151
Table 3.12	Extracted Factors Using Oblimin Rotation and Three Fixed Factors.....	152
Table 3.13	Reliability of Three Factor Instrument.....	158
Table 4.1	Frequencies and Percents of Age, Gender, Education Level, and Ethnicity of Participants...	163
Table 4.2	Revised Instrument Cases Discarded Due to Non-Response.....	169
Table 4.3	Varimax Rotated Factor Matrix for Instructor Characteristics.....	172

List of Tables (Continued)

Table 4.4	Varimax Rotated with Three Fixed Factors for Instructor Characteristics.....	174
Table 4.5	Factor Loading Coefficients within Three Fixed Factors.....	175
Table 4.6	Number of Factor Items, Mean, Standard Deviation, Minimum and Maximum Response Ratings for Instructor Characteristics.....	176
Table 4.7	Mean, Standard Deviation, Minimum and Maximum Response Ratings of Averaged Scales for Perceived Instructor Characteristics.....	178
Table 4.8	Instructor Quality Mean, Standard Deviation and Response by Rating Scale.....	179
Table 4.9	Frequency and Percent of Post-Lesson Test Scores by Percentile.....	180
Table 4.10	Relationship between Perceived Instructor Quality, Post-Lesson Test Scores and Instructor Characteristics as Measured by Pearson's Product Moment Correlation.....	182
Table 4.11	Step-wise Multiple Regression Analysis Using Forced Entry to Explore if Instructor Characteristics Explain a Significant Amount of Variance in Perceived Instructor Quality.....	184
Table 4.12	Step-wise Multiple Regression Analysis to Explore if Perceived Instructor Characteristics Explain a Significant Amount of Variance in Post-Lesson Test Scores.....	186

LIST OF FIGURES

Figure 1.1	Theoretical Framework of Instructor Characteristics Related to Student's Reaction to Perceived Instructor Quality and Student's Learning.....	11
Figure 1.2	Conceptual Framework of Instructor Characteristics Related to Educational Outcomes of Student's Reaction to Perceived Instructor Quality and Student's Learning.....	13
Figure 3.1	Un-Rotated Scree Plot with Eigenvalue of One..	144
Figure 3.2	Oblimin Rotation Eigenvalue Scree Plot with Five Fixed Factors.....	147
Figure 3.3	Oblimin Rotation Eigenvalue Scree Plot with Three Fixed Factors.....	150
Figure 4.1	Varimax Rotation Scree Plot for Instructor Characteristics.....	173

CHAPTER 1

Nature of the Problem

Between January 2009 and October 2009, 2.9 million students graduated from high school. Of those graduates, 2.1 million (70.1%) continued their post-secondary studies by enrolling in 2- or 4-year colleges (Bureau of Labor Statistics, 2010a). Of the 879 thousand not enrolled in college, the majority will enter the civilian workforce and roughly 184 thousand will enlist in the U.S. Armed Forces to replace those who retire or leave the service at the end of their commitment (Bureau of Labor Statistics, 2010b).

Normally when a new employees enters the workforce they require initial training to gain organizational familiarity and proficiency in the skills they bring to their new vocation. With new systems innovations or modernized equipment, vested employees also frequently require and benefit from sustainment and refresher training. To address these reoccurring needs, U.S. organizations spent over \$134.07 billion dollars in 2008 on direct employee learning and development (Paradise & Patel, 2009). Therefore, it is not surprising that 28% of training organizations considered increasing effectiveness, concurrent with measuring the impact of the training programs, their top

priorities (Industry Report, 2007). This is no different for many non-profit companies or governmental agencies that must still obligate and justify resource allocations and cost expenditures.

In 2006, the U.S. Armed Forces provided training and work experience to more than 2.6 million people making it the largest single provider of on-the-job training and experience in the U.S. economy and "the largest employer of youth" (Asch, 1994, p. 1; Bureau of Labor Statistics, 2008; Goldberg & Warner, 1987). The U.S. Armed Forces is comprised of "dedicated, highly educated, and well-trained men and women" (U.S. Office of Management and Budget [OMB], 2005b, section 1.1) who possess diverse skills and abilities needed to "run a hospital, command a tank, program computer system, operate a nuclear reactor, or repair and maintain a helicopter" (Bureau of Labor Statistics, 2008, p. 1).

To meet the personal training needs, the Department of Defense obligates over \$12.8 billion dollars annually for basic skills and advanced training (Department of Defense Fiscal Year Budget, 2008). Another \$35 billion is allocated annually for training scenarios and training missions (National Training Systems Association, 2008). But, unlike many civilian companies, the U.S. Armed Forces develops their trainers from within as the

"crucial task of training can be accomplished only through the assignment of the most highly skilled and motivated non-commissioned officers who are led by experienced quality officers" (Cohen & Shelton, 1998, March 16). These training instructors are selected from the individual's vocational specialty based on "current qualifications and ability to fill a valid requirement" (Wolfowitz, 2007, p. 2).

Specific to the U.S. Marine Corps, instructors are required to attend a one-week formal course to gain the requisite skills necessary to train personnel (Formal School Instructor Course, 2008, p. 2). Evaluation within the formal vocational schools is mandatory and directs managers to evaluate instructor's performance and gather and analyze data on "student reaction to the instruction, student performance in the course, and graduate performance on the job" (Hanlon, 2003, p. 9).

With the considerable focus on measuring the impact of training in the workplace, most organizations - including the U.S. Armed Forces and the Department of Defense - generally rely on Donald Kirkpatrick's four-level model of evaluation techniques: reaction, learning, behavior, and results (Alliger & Janak, 1989; Faerman & Ban, 1993; Lee & Pershing, 1999; Plant & Ryan, 1994). Unfortunately, due to the complexity and cost "of

measuring behavioral change, most organizations have simply relied on reaction measures, generally using end-of-course evaluations that ask the participants if they liked the course and if they thought the material covered was relevant or useful" (Faerman & Ban, 1993, p. 300).

Even though "relatively little is known about characteristics of instructors that contribute to positive or negative evaluations from students" (Murray, Rushton & Paunonen, 1990, p. 250) most organizations only measure "tactical aspects such as completions and volumes—things that are relatively easy to measure, but aren't very valuable" (Industry Report, 2007, p. 17). This has led to "defining good teaching in terms of good scores on the student evaluation forms" which is "based on an analogy between the student and the consumer - the student, as the primary consumer of the teaching product, is in the best position to evaluate its worth" (Rodin & Rodin, 1972, p. 1166).

Statement of the Problem

Current vocational-training within the U.S. Armed Forces is clearly well-designed enough to produce technically proficient graduates, as evident by the operational success demonstrated in Iraq, Afghanistan, and multiple humanitarian aid missions throughout the world. Although their structure, roles, and

responsibilities may differ, the U.S. Armed Forces is similar to many large civilian organizations in the way they evaluate training effectiveness; both primarily rely on the student's reaction to the training and their satisfaction rating to measure training effectiveness. Like civilian organizations, this evaluation shortfall is likely due to the complexity and cost, when "measuring behavioral change, most organizations have simply relied on reaction measures" (Faerman & Ban, 1993, p. 300). This serves as an evaluation 'check-in the box' by recording trainee attendance and completion but does little to indicate if the training has successfully met the trainee's, and by extension, the organization's goals.

Additionally, student reactions to the training program, and more specifically an instructor's effectiveness, can impact the career potential of those same instructors (Schmoker, 2006; Simpson, 2008). Supervisors use the tallied student reactions as quantifiable surrogate appraisals for an instructor's performance (Fitzpatrick, 2004). This leads to negative consequences for the organization, the instructor, and the student. The organization can be negatively impacted due to potentially lowered staff morale, increased personnel turnover and the associated incidental costs to retrain new personnel. To the instructor

assigned to a training billet as a collateral duty, and who is usually only a subject matter expert in a specific vocational specialty and not a credentialed teacher, negative student ratings can be detrimental to promotion, longevity and morale. Lastly, the student may not achieve the highest possible learning outcomes which may negatively affect enthusiasm to learn and overall mastery of learning objectives.

The Purpose of the Study

The purpose of the study was to examine instructional techniques and attributes (henceforth called: instructor characteristics) for application in the evaluation process of formal instructors within U.S. Marine Corps vocational specialty courses. Once identified, these revealed instructor characteristics could be used to predict educational outcomes.

The study's purpose was to create an instrument capable of assessing the student's perception of their instructor using the instrument's instructor characteristics rating scale. Secondly, the study's purpose was to identify how students perceive their instructor using the revealed and rated instructor characteristics and post-lesson test scores. The study also explored relationships between revealed instructor characteristics and student's reaction to perceived instructor

quality and learning. Lastly, the study explored if certain factors explain student's perception of instructor quality and associated learning as evidenced by post-lesson test scores.

Research Objectives

There were three research objectives central in this study. They were:

Research Objective 1 - The first research objective was to determine consistent instructor characteristics, aligned into distinct factors, which could be used to predict student's reaction to instructor quality and student's performance on post-lesson test scores. The study will create an instrument to assess instructor quality based on student's reaction to perceived instructor characteristics and post-lesson test scores.

Research Objective 2 - Research objective two explored the instructor characteristics that explain student's scaled rating responses of satisfaction in regards to perceived instructor quality.

Research Objective 3 - Research objective three explored the instructor characteristics that explain student's scaled rating of satisfaction in regards to student learning as reflected by post-lesson test scores.

Research Questions

There were six research questions considered in the study:

Research Question 1 - *What are the instructor characteristics that can be used to describe instructors and explain student's reaction toward instructor quality and student learning of lesson material?*

Research Question 2 - *How do students perceive instructor characteristics as measured by the developed instrument?*

Research Question 3 - *What are the student reactions to perceived instructor quality as measured by the instrument and student learning as reported by post-lesson test scores?*

Research Question 4 - *Are there relationships between the instructor characteristics and student's reaction to perceived instructor quality as measured by the instrument and student learning as reported in post-lesson test scores?*

Research Question 5 - *Do instructor characteristics explain a significant amount of variance in perceived instructor quality?*

Research Question 6 - *Do instructor characteristics explain a significant amount of variance in post-lesson test scores?*

Research Methodology

The research methodology for the study employed a quantitative survey research design. The research methodology supports the collection of sampling reactions with validated collection instruments. A comprehensive review of literature was used to identify constructs and gather items to create the research instrument. All collected items were placed within themes and within themes the sub-functions and associated items were categorized. Participants were surveyed with a paper-based instrument. Two versions of the instrument were developed; an initial version used in the field test first round, and a revised instrument used in the second and final round. Preliminary data analysis was conducted to refine the instrument for appropriateness as a result of participant responses obtained during the first round.

Post-lesson test scores (as an indication of learning) and demographics were also collected from all participants. Data were collected from U.S. Marines currently enrolled in entry-level formalized military occupational (vocational) specialty training during the months of May 2010 through August 2010. Data were objectively analyzed using statistical procedures provided

by *Statistical Package for the Social Sciences (SPSS)*, version 18.0.

The study collected participant data in two rounds to develop the instrument for research objective one. Descriptive statistics (means and standard deviations) were used to meet research objective two and also to describe the participants. Exploratory factor analysis, correlations and multiple regressions were used to meet research objective three.

Theoretical Framework

The theoretical framework for the study was the product of a comprehensive review of social learning, humanistic, cognitivist and behaviorist theories. Instruction Systems Design (ISD) theory, as a product of behaviorism and cognitivism, was considered most appropriate for this study and is shown in Figure 1.1. A detailed review of potential learning theories considered for inclusion in this study is included in chapter two.

Today, the U.S. Armed Forces create learning situations based on a systems approach to training which employ much of the research gained from previous generations (Hanlon, 2003). Therefore, for the purpose of this study and its unique population group, we will focus on components of the SAT used in

U.S. Marine Corps vocational training today. Specifically the study will concentrate on the evaluation component used to measure training effectiveness from an instructional systems design (behaviorist + cognitivist) theoretical framework.

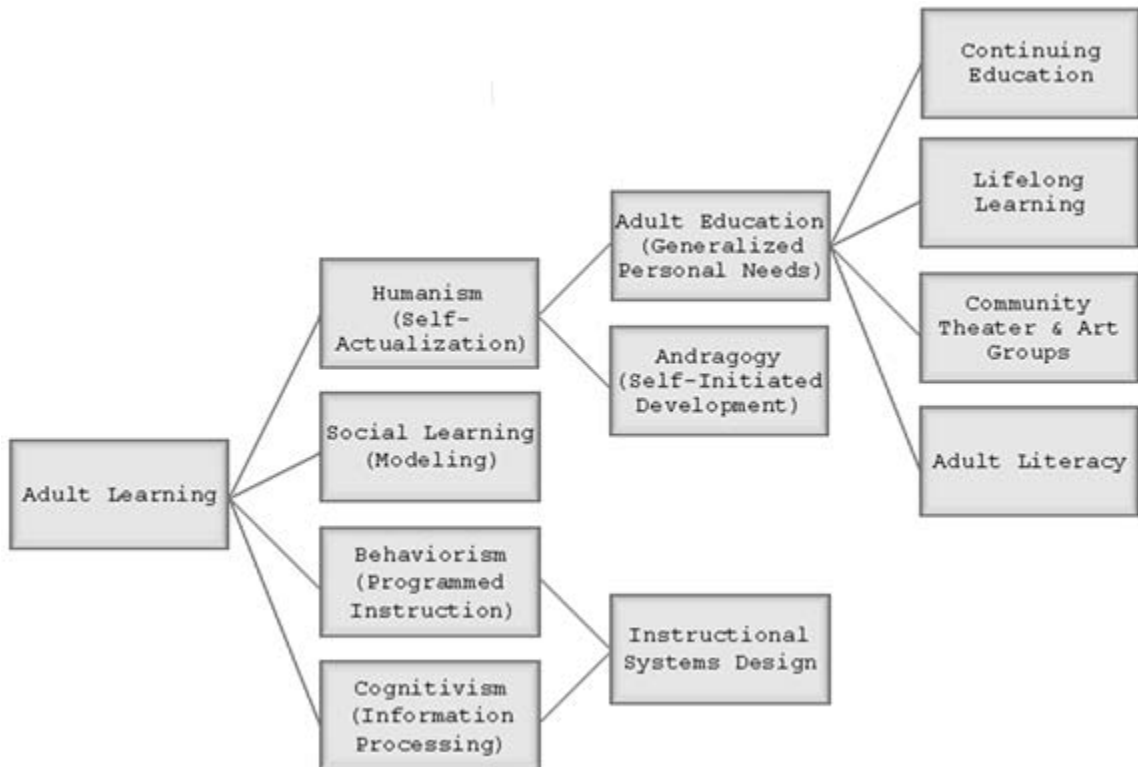


Figure 1.1.

Theoretical Framework of Instructor Characteristics Related to Student's Reaction to Perceived Instructor Quality and Student's Learning

Conceptual Framework

The conceptual framework (Figure 1.2) of the study indicates instructor characteristics as factors and student's reaction and learning as educational outcomes. Collected data included responses to items from the instrument used to assess participant reaction to perceived instructor characteristics, individual participant demographics, and participant post-lesson test scores.

The conceptual framework includes instructional and educational-related themes comprised of individual items (independent variables) that could be first identified and then clustered into related item factors based on participant's scaled rating. The dependent variables were the student's reaction to perceived instructor quality as measured by the study's instrument and student's evidence of learning reflected on post-lesson test score. To better conceptualize the framework of the study each research question is discussed in the following paragraphs.

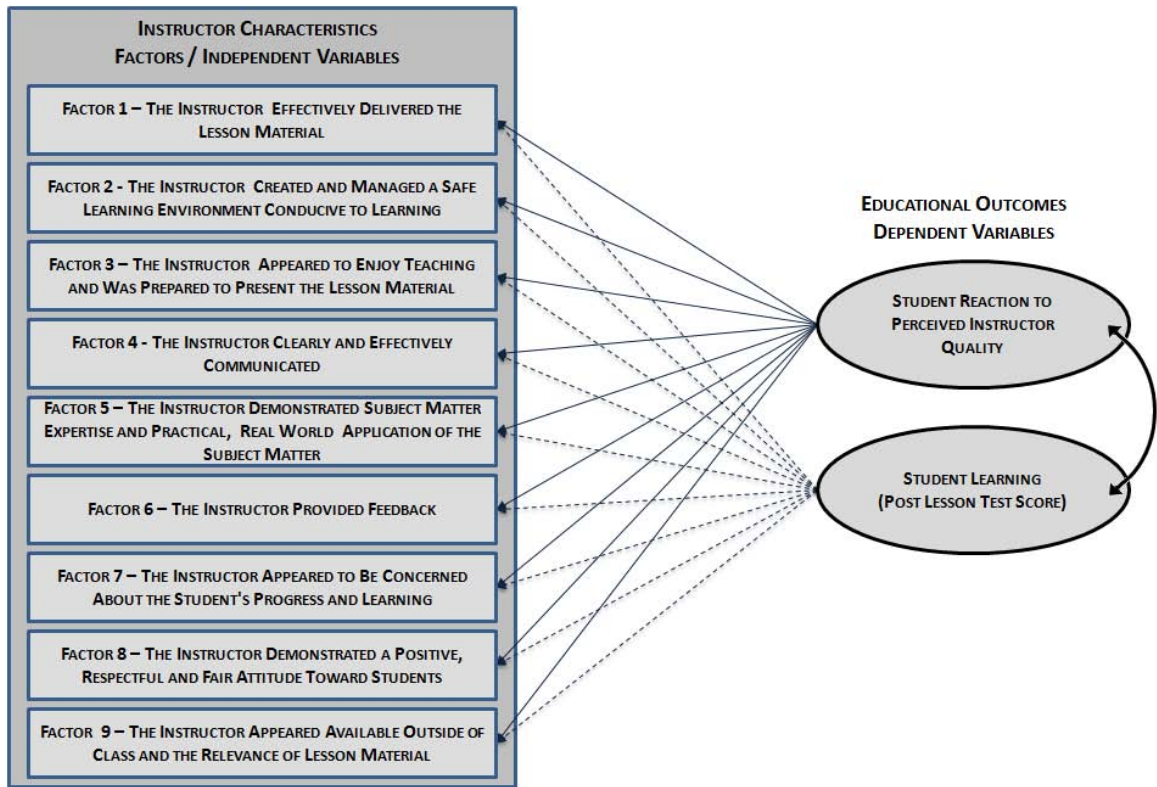


Figure 1.2.

Conceptual Framework of Instructor Characteristics Related to Educational Outcomes of Student's Reaction to Perceived Instructor Quality and Student's Learning

Research Question 1 - *What are the instructor characteristics that can be used to describe instructors and explain student's reaction toward instructor quality and student learning of lesson material?*

Research question one sought to identify instructor characteristics as individual items found during the literature

review of previously published and peer-reviewed instruments. These items were revised into an assessment instrument to collect student's perceived ratings of instructor characteristics as independent variables (interval). Instrument items that rated the student's perception of instructor quality were the dependent variables (interval). The student's post-lesson test scores were collected subsequently and also became dependent variables (ratio).

Research Question 2 - *How do students perceive instructor characteristics as measured by the developed instrument?*

Research question two examines how students perceive instructor characteristics as rated by the developed instrument. The independent variables used in this research question are the same variables summated into factors explored in question one. The dependent variable is the student's rating of perceived instructor quality (interval).

Research Question 3 - *What are the student reactions to perceived instructor quality as measured by the instrument and student learning as reported by post-lesson test scores?*

Research question three sought to determine the student's reaction to the instructor's quality as measured by the

instrument and the student learning as evidenced by the post-lesson test scores. The independent variables are the same as the two previous research questions. The dependent variables include the student's rating of satisfaction of the instructor's quality and the student's post-lesson test score as evidence of learning.

Research Question 4 - Are there relationships between the instructor characteristics and student's reaction to perceived instructor quality as measured by the instrument and student learning as reported in post-lesson test scores?

Research question four sought to determine if relationships existed between student's reaction to perceived instructor quality and student learning as evidenced by individual post-lesson test scores. The independent variables remain the same variables as the previous research questions; whereas, the dependent variables include the student's rating of instructor quality and the student's post-lesson test scores.

Research Question 5 - Do instructor characteristics explain a significant amount of variance in perceived instructor quality?

Research question five sought to identify the amount of variance found in the instructor's characteristics that explain

the student's reaction to perceived instructor quality. The independent variables remain the same variables as previous research questions while the dependent variable is the student's rating of the instructor's quality.

Research Question 6 - *Do instructor characteristics explain a significant amount of variance in post-lesson test scores?*

Research question six concludes this section and sought to identify the amount of variance found in the instructor's characteristics that explain student learning as evidenced by post-lesson test scores. The independent variables remain the same variables as previous research questions while the dependent variable is the student's post-lesson test scores.

Significance of the Study

Training and subsequent evaluation of intended outcomes is a resource intensive yet necessary endeavor regardless of the organization. Reviewing and compiling learner reactions to specific instructor characteristics as items within peer-reviewed instruments could lead to a more streamlined and cost effective tool. Once identified, this tool could be used to focus and enhance trainer developmental programs and by extension increase student satisfaction and performance.

The final instrument will benefit multiple stakeholders including the administration who will gain legitimate empirical assessments of the instructor providing training; the instructor who will be more apt to improve their training methods as a result of accurate assessments; and the student who will be able to provide relevant instructor assessments with the implication of their recommendations being employed in future training sessions.

The overarching goal of the study was to develop an instrument to assess instructor characteristics based on research findings from multiple training disciplines and educational settings for application to U.S. Marine Corps vocational specialty instructors. Additionally, a greater understanding of the relationship between student reactions to perceived instructor quality and student learning could provide empirical evidence for continued or reduced use. This evidence could be used to justify non-training time and validate administrative costs.

Delimitations

The study was bound by the specific population of U.S. Marine Corps vocational trainees at Marine Corps Base, Camp Johnson, North Carolina. Additionally, the study was delimited to the

specific population group being trained to support current U.S. national defense operational missions in Southwest Asia (Afghanistan, Iraq, etc.) as well as other global military obligations. Furthermore, the study was delimited by the use of literature to create constructs and operational statements designed to assess U.S. Marine Corps' vocational training effectiveness.

Organization of the Study

The study is structured around five chapters. Chapter one began by identifying the nature of the problem, the statement of the problem, the specified research objectives and associated research questions, and the theoretical and conceptual frameworks used to guide the study. The research design and methodology were introduced with the study's variables. Chapter one concludes by identifying the study's significance, delimitations and overall organization of the study.

Chapter two provides a thorough examination of literature related to student evaluation and satisfaction, the consumer model, effective teaching, and instructor performance. The chapter describes the adult learning theories (humanism, adult education, andragogy, social learning, behaviorism, cognitivism and systematic instructional design) examined and explains the

justification of the study's theoretical framework as it applies to U.S. Marine Corps vocational training. The chapter concludes with an extension review of the U.S. Armed Forces for the benefit of those non-military orientated readers. Specific areas discussed include the enlistment process, qualifications, assignment to occupational specialties, training (both initial and vocational), military evaluation of training, and the role of the military instructor specific to the U.S. Marine Corps.

The third chapter describes the research design, research objectives and associated research questions, the variables used in the study, the study's methodology, as well as the study's sample and participants. A detailed narrative of the instruments' creation, item selection, and critical review is included. Missing data and outlier responses were analyzed and discussed relative to the study. Exploratory factor analysis was conducted for item refinement during instrument creation and is comprehensively discussed. Finally, an overview of the data collection procedures, subsequent analysis, and assumptions of the findings are discussed.

Chapter four presents the analyzed findings from the collected data. Findings are presented that address the research objectives and research questions. A description of the

participant's demographics, correlations between the study's variables, and regression analysis conducted to explain relationships and variance between demographic variables and participant responses is provided.

The fifth and final chapter presents a detailed summary of the study's findings and conclusions. Lastly, recommendations for future research and limitations of the study are provided.

CHAPTER 2

Review of Literature

Chapter two is broken down into three primary sections; a literature review of student reactions and evaluation of effective teaching, a review of multiple learning theories, and a comprehensive background of the U.S. Armed Forces.

The chapter initially provides a review of literature explicit to this study focused on student evaluation, student satisfaction, perception of teaching effectiveness, and the effect of instructor performance for application within a military performance-based training setting. Chapter two also provides a detailed analysis of multiple learning theories in order to better identify a practical theoretical framework in support of this study and military vocational training. Specifically, humanism, adult education, andragogy, social learning, behaviorism, cognitivism and systems instruction design were considered for inclusion. Chapter two concludes with a detailed narrative of the U.S. Armed Forces with a specific concentration on vocational training within the U.S. Marine Corps.

Literature Review of Student Evaluation

A simple internet search in the ERIC database system identifies over 14,399 research articles related to student evaluations making an overview of teaching effectiveness difficult especially with the amount of conflicting research evident. Therefore, a review of student evaluation is warranted.

Student Satisfaction

Early in the 20th century, Edward Thorndike, while studying the human learning process, concluded that the connection between "situation and response by resulting satisfaction is better than the inhibition of alternative connections by discomfort" (Thorndike, 1910, p. 11); whereas, "a satisfying aftereffect strengthens somewhat the connection to which it is attached" (Thorndike, 1933, p. 15). Later in 1940, Lily Detchen explained that Thorndike had "formulated a psychological assumption, to some extent now substantiated by experimental evidence, that an attitude of satisfaction or annoyance is conducive to learning" (Detchen, 1940, p. 146).

A few years later, Ralph Tyler wrote *Basic Principles of Curriculum and Instruction* where he cautioned that a student's interests should be considered in the design of educational

objectives because "what one is interested in largely determines what he attends to and frequently what he does" (Tyler, 1949, p. 78). Therefore it is not surprising, with such distinguished educational leaders stressing the need for student satisfaction and focused interest, that Donald Kirkpatrick considered student reaction to training the first stage in his 1959 four-stage model of evaluation (Kirkpatrick, 1979). Thorndike, Tyler and Kirkpatrick focused on the role of student's satisfaction and that combined focused interest may have contributed to the consumer model of student evaluations where "the student, as the primary consumer of the teaching product, is in the best position to evaluate its worth" (Rodin & Rodin, 1972, p. 1166).

Consumer Model

The consumer model assumes that education and learning are similar to other fields of consumption and as such "the consumer is given an opportunity to rate the product" (Detchen, 1940, p. 147) and "as consumers of instruction, are best qualified to evaluate the product being offered" (Marsh, Fleiner & Thomas, 1975, p. 833). Proponents of the consumer model suggest "the student does have an opinion which he will record sincerely if given the opportunity and which can make a helpful contribution toward a better planning of his welfare" (Detchen, 1940, p. 146).

Kirkpatrick (1994) suggested that "evaluating reaction is the same thing as measuring customer satisfaction" (p. 27); therefore, effective training requires favorable reactions or trainees "will not be motivated to learn" (p. 27). Newstrom (1978) succinctly outlines Kirkpatrick's model by describing the process as "if trainees react favorably, they'll probably learn more; if they learn more, they'll probably change their behavior; and if they change their behavior, the usual indices of performance will improve" (p. 22). Kirkpatrick (1994) even referred to these reaction surveys as "happiness sheets" (p.27) but cautioned that they only rate what learners like about the training program - specifically "reactions both to the subject and to the leader" (p. 28) and that evaluators should sequentially implement all four levels of the evaluation model because each stage in the process "provides more valuable information" (p. 21).

Although, Kirkpatrick may never have meant for his model "to be more than a first, global heuristic for training evaluation" (Alliger & Janak, 1989, p. 339) countless future evaluators have continued to focus solely on Kirkpatrick's first level of evaluation - student reaction to training (Lee & Pershing, 1999). Thereafter, the consumer model afforded

students the opportunity to evaluate their teachers in the same manner that shoppers could evaluate a shopping experience. This created an atmosphere in academia where for decades the major argument over "college and university students' evaluation of their professors has been raging" (Raskin & Plante, 1979, p. 381). Theall and Franklin (2001) suggest that "few issues in higher education are as sensitive, divisive, and political as faculty evaluation and in particular the quality and value of the information provided by students in their evaluations of teachers and courses" (p. 45).

Even though Kirkpatrick (1979) was quick to warn that even if student reactions are superbly measured "there is no assurance that any learning has taken place" (p. 126) many educators were critical of his model emphatically stating that "liking training has little relationship to learning, learning has little relationship to using training" (Brinkerhoff, 1995, p. 387). This may have been because Kirkpatrick's model made "a lot of intuitive sense" (Brinkerhoff, 1995, p. 387) as its power lies in its "simplicity and its ability to help people think about training evaluation criteria" (Alliger & Janak, 1989, p. 331). But the simplicity of the model led to other criticisms that suggested by concentrating so heavily on evaluating student

reaction as a primary product - defined only by student happiness - the result has diverted attention away from what is truly important - evaluating student performance - in lieu of "activities that generate high ratings" (Holton, 1996, p. 11).

Researchers like Nancy Dixon (1990) dissected Kirkpatrick's model and concluded that there was no significant relationship between "trainee perceptions of enjoyment and their posttest scores", no significant relationship between "trainee perceptions of job relevance and their posttest scores" and no significant relationship between "perceptions of instruction skillfulness and their posttest scores" (p. 137).

Mathieu, Tannenbaum and Salas (1992) disagreed and conducted a similar analysis finding that only by examining "linear relationships" could researchers conclude that "reactions have no significant effects on learning" but that "participants' reactions to the program played a multifaceted role in linking individual and situational characteristics to other training effectiveness measures" and that "reactions are important for training effectiveness, but not in and of themselves" (p. 843).

Criticism of the Consumer Model

Critics of the consumer model argue that "students are incapable of evaluating in a valid manner the quality of instruction they receive and should not be given the opportunity to do so" (Simpson, 1995, p. 3) and that "students cannot judge all aspects of teaching effectiveness equally well" (McKeachie, 1979, 390) and "that by asking them to do so undermines faculty confidence in SETs [Student Evaluation of Teaching]" (Simpson & Signaw, 2000, p. 205).

As far back as 1940, Lily Detchen considered qualifying student opinions "to the degree to which the opinions of the number among the student body who are incompetent to judge can be eliminated" because the "validity of the student-rating device cannot be well established with the opinion of experts" (Detchen, 1940, p. 148-149).

There are a number of challenges to the consumer model and use of student evaluations which include evaluations that only serve as popularity contests, student's ability to accurately evaluate, the student's reason for attending training and perceived outcomes, student's dissimilar preferences of teaching style, the student's perception of the instructor's level of

caring (halo effect), and the emotional and psychological factors that may affect the student.

By tracing the history of student evaluations since the 1960s, Knapper (2001) determined that "questionnaire-based teacher ratings" were initiated by students as popularity contests "primarily as an aid to selecting courses and teachers" (p. 7). McDaniel (2006) echoes this caution suggesting that this type of consumer mentality where the student provides a usually anonymous customer satisfaction rating for services, in this case a lesson or course, they have bought and paid for "may not be in the best interest of faculty or students to assume this right and such a level of competence" (p. 8) because "student evaluations reflect popularity and other factors unrelated to teaching excellence" (Marsh, Fleiner & Thomas, 1975, p. 833). Furthermore, popularity is seen "as most related to evaluations, followed by the effectiveness of teaching" (Gross & Small, 1979, p. 218). Abrami and d'Apollania (1999) warn that "not every student registers for a course with the loftiest ambitions and purest intentions" and that others "may lack the necessary prerequisite skills and abilities to excel" (p. 520); therefore, their ability to accurately evaluate an instructor's performance may be limited.

A recent Nebraska court case amplifies the role of students rating teachers when a fifteen-year veteran teacher was fired as a result of two unsatisfactory performance ratings which included scores derived directly from student satisfaction surveys. During one year, the teacher received almost perfect ratings from peers and the principle and mostly superior marks from students; although, the overall student score was skewed when other students rated him poorly with complaints like "He gives us too much homework; he's "too pleasant and positive"; he should show "more movies" in class; and "he talks funny and doesn't have no humor" (Simpson, 2008, p. 20). The NEA union argued that the firing was "arbitrary and capricious" because it was based solely on student evaluations" and that it was unreasonable for the school to rely on the opinions of "emotionally troubled, educational disadvantaged" students who were "inexperienced in responsible decision making" (Simpson, 2008, p. 20). The U.S. District Court agreed with the NEA union and ordered the teacher reinstated with back pay and benefits.

McKeachie (1997) presents two challenges to the consumer model and the use of student evaluation. The first is that students may prefer a certain teacher style as "many students prefer teaching that enables them to listen passively - teaching

that organizes the subject matter for them and that prepares them well for tests" (McKeachie, 1997, p. 1219). The second is the student's perception of teacher's level of caring where "those students (frequently the less able) who feel that the teacher does not care about their learning develop a negative halo, whereas those who feel that the teacher cares about them develop a positive halo" (McKeachie, 1997, p. 1221).

Small, Hollenbeck and Haley (1982) caution that a student's "emotional state at the end of the semester, which is the time of instructor and course ratings, was related to instructor ratings" and that the more "hostile, anxious, and depressed" the student, the "lower the evaluation" (p. 207). Small, Hollenbeck and Haley (1982) suggest that a greater threat to the validity of student evaluations lies in the "potential confounds arising from psychological factors" particularly when a "close examination of this available research does not reveal a consistent pattern" (Small, Hollenbeck & Haley, 1982, p. 205). Finally, McDaniel (2006) questions the benefit of students evaluating teaching effectiveness when "the evaluation instrument, the insight of the student, the personality of the faculty member, the motivation and fairness of the young evaluator, and myriad other variables that make the fundamental validity of the process doubtful" (p.

8). In light of the conflicting research, the researcher or practitioner must ask: *What does correlate with student ratings?*

Correlation of Student Ratings

In 1972, Rodin and Rodin assessed the validity of two measures used to identify effective teaching; the objective criterion - based on what a student has learned; and the subjective criterion - based on student evaluations of teacher effectiveness. Rodin and Rodin (1972) professors in psychology and mathematics, respectively, were concerned over the easy implementation and widespread use of student evaluation forms where "good teaching is then defined as good scores on the student evaluation form" (p. 1164). They concluded, based on a 293-student undergraduate calculus class correlational study between the objective (student grade) and subjective (student evaluation) measures of teaching ability, that there was a negative relationship between student grades and student evaluation where the "instructor with the three lowest subjective scores received the highest objective scores" and the "instructor with the highest subjective rating was lowest on the objective measure" (p. 1165). In 1974, Rodin and Rodin reemphasized that "the most striking thing about studies relating student achievement to student ratings is the inconsistency of the results. Very high positive, very high

negative, moderate positive, and inconsistent correlations have been reported" (p. 56).

Other researchers, like Centra in 1977, came to different conclusions and reported the "highest correlation with achievement were for ratings of the value of the course to students, followed by the ratings of teacher effectiveness and ratings of lecture quality" (p. 20) as well as "ratings of course objectives and organization" (Centra, 1977, p. 22). Small, Hollenbeck and Haley (1982) also reported in their study that the "students' actual grades were significantly correlated with student ratings of instructors and courses, i.e., those receiving higher grades tended to rate the course and instructor more favorably" (p. 207). McKeachie (1997) stressed that "student ratings are not perfectly correlated with student learning, even in the validity studies carried out in large courses with multiple sections" (p. 1219). Abrami, d'Apollonia and Cohen (1990) describe the debate most eloquently suggesting that "despite many decades of research on the validity of student ratings, the thrust of our conclusion is that additional research lies ahead" (p. 231). From the viewpoint of the student, the challenge is compounded by first having to operationally define what it means to be an effective teacher before they can even

begin to consider accurately evaluating an instructor's performance.

Perception of Effective Teaching

Even though "students perceive the evaluation of teaching effectiveness to be important" (Wulff, Staton-Spicer, Hess & Nyquist, 1985, p. 42) "they are evenly split on whether students can appropriately evaluate instruction but seem to agree that student evaluations should carry some importance" (Gross & Small, 1979, p. 218). McKeachie (1979) reviewed multiple studies to find common instructor characteristics and determined that highly rated teachers possessed often contradictory scores of "extraversion, intuitiveness, and "feeling" on the Myer Briggs Type Indicator" and that "highly rated teachers were perceived to be dynamic, amicable, and highly intellectual" (p. 391).

Multiple studies report incongruency as specific personality traits and effective teaching characteristics vary between the types of courses and aspects of teaching being evaluated (Abrami, d'Apollonia & Cohen, 1990; Murray, Rushton & Paunonen, 1990). Marsh (1984) even suggests that out of the many observations he reviewed, there were none that could provide "an empirical demonstration of improvement of teaching effectiveness resulting from students' evaluation" (p. 746).

Raskin and Plante (1979) conclude that the most "controversial and most complex of all are those innumerable articles by psychologists which set out to demonstrate that an instructor's "personality" is a significant determining factor in the ratings - a factor which cannot be neutralized by even the most sophisticated of evaluation questionnaires" (p. 381-2). Murray, Rushton and Paunonen (1990) caution that despite the "abundance of research on the reliability, validity, and utility of student ratings, relatively little is known about characteristics of instructors that contribute to positive or negative evaluations from students" (p. 250).

Notwithstanding the vast number of often conflicting studies, McKeachie (1979) conducted his own review and research to come to the conclusion that relatively few student characteristics, such as "age, sex, and level of student" (p. 390) have any significant effects on student ratings of teaching and that the sex of the instructor or the instructor's "personality characteristics do not show consistent relationships to ratings of effectiveness" (p. 391). McKeachie (1979) stresses that with "respect to the general validity question, the cumulating evidence continues to support the conclusion that highly rated teachers tend to be those whose students achieve

well" (p. 384) with the caveat that "even though the data are now strongly supportive of the validity of student ratings for certain goals, this does not mean that they are impervious to influences by other factors" (p. 389). McKeachie (1979) concludes the most important concern that as professionals "we do not know whether these teachers were, or were not, effective in influencing student learning" (p. 391).

Detchen (1940) suggested that "properly collected and distributed to the offices and faculty concerned, such surveys of student opinion pay generous dividends in ideas, in student-teacher rapport, in student morale" (p. 147); whereas, Cohen (1981) concluded after reviewing the findings of a meta-analysis that "student ratings of instruction are a valid index of instructional effectiveness" and that "students do a pretty good job of distinguishing among teachers on the basis of how much they have learned" (p. 305).

Eric Arubayi (1987) reviewed the considerable available literature and felt confident in declaring that even "though some authors disagree, the balance of the findings in the literature agree that student ratings of instruction lead to teacher effectiveness and improvement of instruction provided appropriate feedback and expert advice are made available to instructors" (p.

274). Simpson (1995) also felt confident enough to state that "while there are many factors that substantially influence student ratings, it is safe to say that there is a reasonable positive relationship between student evaluation of teaching and academic achievement" (p. 3); while Greenwald (1997) suggests that "if grades correlate with ratings only or mainly because good teachers produce both high grades and high ratings, then all is well with the validity of student ratings" (p. 1184).

But as is often the case, many other researchers, like Marsh (1984), questioned the validity of student ratings by suggesting that "studies of the usefulness of student ratings are infrequent and often anecdotal" (p. 748). Crader and Butler (1996) examined the validity of students' evaluation of teaching and concluded that "teachers' behaviors and abilities are not the only - and probably not the most important - variables affecting students' ratings of teachers" and that for all practical purposes "teachers might have less control over their ratings than is commonly believed" (p. 313). Kemp & Kumar (1990) cautioned that "student ratings are inadequate as a standalone, or sole measure of an individual's teaching effectiveness; effectiveness can be adequately assessed only when multiple indicators of teaching effectiveness, including an evaluation of

class materials such as syllabi, examinations, handouts, and pedagogical innovation, are used" (p. 112).

Richardson (2005) explains that the even if the overall process may demonstrate a relationship between students' evaluations and academic performance, the "demands and the assessment criteria of different course units may vary, and so students' grades or examination marks cannot be taken as a simple measure of teaching effectiveness" (389-390). Pounder (2007) echoed this concern by suggesting that in regards to the Students-Evaluating-Teachers (SET) "process in its conventional form, its value is questionable as the sole measure of classroom performance since the quality, richness and diversity of what happens in the typical classroom cannot be captured by the SET process alone" (p. 186).

Raskin and Plante (1979) are remarkably forthright in suggesting that a "careful survey of the studies dealing with student evaluation of teachers supports the widely held suspicion that with little effort one could probably find a report containing evidence to confirm any contention regarding the reliability of such a tool in assessing the effectiveness of faculty in the classroom" (p. 381). With so many conflicting

studies, one must seriously consider what the effect of student ratings have on an instructor's overall performance.

Effect on Instructor's Performance

The conflict of students rating teachers has been ongoing for at least 60 years as Guthrie (1953) suggests there "will always be a number who oppose soliciting student opinions in spite of the fact that students are mature and are the only persons in a position to observe teachers directly" (p. 221). Guthrie (1953) goes so far as to suggest that the "opposition comes largely from "nonteachers" who have difficulty in accepting their failure as a fact and who rationalize it in terms of a low estimate of students' intelligence, industry, and purposes" (p. 221).

Raskin & Plante (1979) identify the limitation of student ratings suggesting that if student ratings were "capable of transforming a dull lecturer into a brilliant one and of inspiring the humorless to coin memorable witticisms, there would be cause to cheer behavior modifications. Unfortunately, what those ratings are able to modify and do modify is behavior quite remote from that related to the quality of a professor's mind and/or personality. The threat, if not the fact, that the results of these evaluations will be used to deny promotion, or tenure, or merit pay leads one to assume that a minority, at the

very least, will yield to the temptation of pleasing their judges by offering pop courses" (p. 382).

Marsh (1984) defines the role of student ratings as indicators of teaching effectiveness that will "provide a basis for informed administration decisions and thereby increase the likelihood that quality teaching will be recognized and rewarded, and that good teachers will be given tenure" (p. 746) because the "social reinforcement of getting favorable ratings will provide added incentive for the improvement of teaching, even at the tenured faculty level" (p. 746). Marsh (1984) cautions administrators that "any procedure used to evaluate teaching effectiveness would prove to be threatening and highly criticized" which is then "exacerbated by the realization that there are no clearly defined criteria of effective teaching" (p. 749). This is a real threat that forces faculty who often have had "little or no formal training in teaching" finding "themselves in a position where their salary or even their job may depend on their classroom teaching skills" (Marsh, 1984, p. 749).

McKeachie (1997) warns that "faculty members and administrators have stereotypes about what good teaching involves. In most meetings to make decisions about promotions or

merit salary increases, negative information is likely to be weighted more heavily than positive information" (p. 1219). Raskin and Plante (1979) conclude that although "universities do not exist for the purpose of maintaining the high morale of faculties" student ratings can create an environment where "morale sinks as a result of self-doubt and guilt, due to a system that requires heroic self-abnegation to escape the lowering of standards and fear of giving offense, corruption has wormed its way into the core of the enterprise" (p. 383).

In 2000, Simpson and Siguwaw conducted an explorative study of the literature in an attempt to develop an analytical framework of common categories and factors associated with student evaluation of teaching (SET). Simpson and Siguwaw (2000) found that faculty "do not perceive SETs as always providing valid measurements of their teaching abilities" but that "SETs encourage instructors to lower educational standards, serve as a tool for student revenge, encourage overreliance on ratings in performance evaluations, and are rife with measurement issues" (p. 209).

Trout (2000) concluded that the "the administrative use of these numerical evaluation forms creates an incentive for instructors to do the wrong thing: to please students instead of

teaching them” and that the “use of such evaluations to reward and punish instructors is doing more to dumb down college education than any other policy or practice on campus” (p. 10).

Ryan, Anderson and Birchler (1980) in a study of 193 faculty (63 percent of entire faculty who responded) found that 87 percent reported lowering their grading standards as a direct result of the mandatory student evaluations policy designed to improve instruction as a component of faculty performance evaluation. Ryan, Anderson and Birchler (1980) stressed that the mandatory student evaluations did result in specific changes in the faculty’s teaching practice but that the common element in instructional activities “was a reduction in coursework demands on students” (p. 329). Ryan, Anderson and Birchler (1980) concluded that “for a large proportion of this faculty, the policy seems clearly to have reduced the morale, job satisfaction, and personal confidence in the institutional administration” (p. 328).

The use of student ratings “has implications for reduced morale and lessened academic rigor” (Gross & Small, 1979, p. 219) when teachers are “pressured to change because of the evaluations made by students” (Gross & Small, 1979, p. 218). Trout (2000) suggests “to earn high scores, instructors must give students

what they want. And what a lot of students want nowadays are stress-free classes, "understanding" instructors, easy-to-get high grades, and undemanding workloads - in essence "education lite" (p. 10). Furthermore, Small, Hollenbeck and Haley (1982) suggest an even larger danger when the "instructor who has maintained high standards without inflated grades may be at a disadvantage when compared with peers who inflate grades" (p. 207). Although, "there is not convincing evidence that teachers use information available on student ratings to improve their courses or their course ratings" (Rodin & Rodin, 1974, p. 56) and that "it is evident that much of what is done by academics to influence student evaluations is of little or no educational value" (Pounder, 2007, p. 185) student ratings are still used extensively throughout public and private educational and training institutions.

With the sheer volume of often contradictory articles surrounding the use of student ratings it is difficult to grasp why nearly 90% of organizations limit their evaluation to Kirkpatrick's first level of evaluation (Lee & Pershing, 1999). It could be, as Faerman and Ban (1993) suggest, that due to the complexity and associated cost of trying to measure if the training has resulted in any actual learning or observable

behavior change that the simplest and often safest way to maintain a successful organizational image is by compiling the participant's own testimony as numerical 'proof' obtained from end-of-course, Likert-type questionnaires that minimally ask if the participants "liked the course and if they thought the material covered was relevant or useful" (p. 300).

What is most disturbing is the concern introduced by Swanson and Holton (2001) who expose the dangerous precedent and over-arching impact of student ratings on the field of training and development where the "idea of engaging learners with interesting activities has led to a perverted "fun-filled" training goal" because of the false belief that the "more participants like a program, the more effective it is" (p. 245). This leads to an examination of how student ratings can improve teaching which requires a literature review of reported ways to improve teaching.

Student Ratings on Improving Teaching

The following is a list of ways identified from student rating peer-reviewed articles on methods to improve instruction. Although a few of the recommendations are dated they were intentionally included to demonstrate the historical precedence of students rating teachers.

- Legitimate evaluation of training needs to “prove that the training solution fits the performance needs, and that the learning that training produces makes an effective and efficient contribution to the performance goal” (Brinkerhoff, 1995, p. 391).

- When students are “asked to state concrete deficiencies, useful suggestions have resulted” (Detchen, 1940, p. 148).

- The “only statistical significant predictor of class achievement was student’s ratings of teacher’s direction (“control”) of the discussion” (Braskamp, Caulley & Costin, 1979, p. 305).

- Listing “prominently the uses of the teaching evaluation on the evaluation instrument. This will inform the students of the uses of the evaluation” (Chen & Hoshower, 2003, p. 84).

- There were “two of the evaluation factors (class presentation and course organization and two summary items [overall instructor teaching course and overall teacher evaluation] correlated significantly ($p < .05$) with the performance criterion” (Marsh, Fleiner & Thomas, 1975, p. 836).

- One “dimension probably reflected student estimates of what they had learned in the course (relative to the objectives)

as well as the extent to which the course had been organized (e.g., the instructor was well-prepared for class)" (Centra, 1977, p. 22).

- For instructors "the use of student ratings is likely to result in improvement when: (a) the ratings provide new information; (b) the teacher is motivated to improve: (c) the teacher can use alternative methods of teaching effectively" (McKeachie, 1979, p. 394).

- McKeachie (1979) provides guidance for using student ratings of teaching which includes allowing space for comments because "comments give examples or incidents which clarify the meaning of ratings"; identifying who will read the comments, wording the items in terms of the individual student's perception, formatting items that encourages good feelings rather than discouragement, adding items that reflect "out-of-class educational functions such as course planning and advising and using teaching ratings to help development of effective teaching skills" (p. 396).

- "Vague personality traits should be avoided. Every trait the student is requested to rate should be clearly described on all points of the scale" (Detchen, 1940, p. 152).

- What “especially characterizes the highly rated teacher is verbal fluency and communication ability. The highly rated teacher seems cultured and sophisticated, expressive and enthusiastic” (Rodin & Rodin, 1974, p. 56).

- Students “are most satisfied with methods that result in instructor use of feedback to make changes during the current term when they can benefit from such implementation” (Wulff, Staton-Spicer, Hess & Nyquist, 1985, p. 42).

- Spooren & Mortelmans (2006), in a study of evaluation results of 222 students found that professional teachers – those “teachers who build up and organize their course in a professional and well-considered way” (p. 211) receive higher student evaluation ratings.

- Centra (2003) in a study of 55,000 2- to 4-year college students who participated in Student Instructional Report II between 1995 and 1999, concluded that “teachers will receive better evaluations when their courses are manageable for students” (p. 515).

- Students “will view instruction as most effective when it is at their level of preparation and ability rather than too difficult, when the course workload is close to what other

courses demand rather than much heavier, and when the pace at which material is covered is about right for them" (Centra, 2003, p. 515).

- Cohen (1981) suggests that what is "not as evident is the strong relationship between structure and achievement. Students of instructors who have everything going according to schedule, use class time well, explain course requirements, and in general have the class well organized tend to learn more than students of instructors who are not well organized" (p. 302).

It is clear there are a multitude of effective teaching methods and techniques to aid the professional trainer or evaluator but the overwhelming number can also serve to confuse and overwhelm the novice, or even the experienced, trainer. Shevlin, Banyard, Davies and Griffiths (2000) suggest that in their research "teaching is "shown to be multi-dimensional" (p. 403) which does little to clarify specific instructor characteristics that are effective.

It is even more confusing when researchers propose often polar opposite suggestions. Abrami and d'Apollania (1999) initially supported student ratings as a "diagnostic tool to encourage teaching improvement and as a source of information about teaching effectiveness" (p. 519) that were able to "reflect

whether faculty promote the learning of students" (p. 520), but later concluded that although there was an strong inference that "highly rated instructors positively affect instructional products" (p. 394) the evidence showed that "student ratings measure directly one product of instruction; namely student satisfaction with teaching" (Abrami, d'Apollania & Rosenfield, 2007, p. 393).

Marsh (1984) considers "students' evaluations of teaching effectiveness" the "most thoroughly studied of all forms of personnel evaluation, and one of the best in terms of being supported by empirical research" (p. 749). Marsh (1984) concludes by presenting contradictory evidence in his own extensive review of research that "the same findings also demonstrate that student ratings may have some halo effect, have at least some unreliability, have only modest agreement with some criteria of effective teaching, are probably affected by some potential sources of bias, and are viewed with some skepticism by faculty as a basis for personnel decisions" (p. 749). Simpson (1995) agrees suggesting that student evaluations appear to be the "most common, yet most controversial, source of information on which to judge effective teaching" (p. 4).

The ability of student ratings to provide beneficial feedback may be limited to how the instrument is developed and what is asked of the students. Therefore, this study will collect and compile questionnaire items currently employed in both public and private organizations to create an instrument capable of effectively and efficiently measuring student reactions within a military vocational training center.

Theoretical Framework

This section of chapter two reviews the overabundance of available research in order to better define a theoretical framework applicable to this specific population; i.e., U.S. Armed Forces entry-level occupational training.

The typical U.S. Armed Forces enlistee is 18 years old and nearly half (47%) of the 2.6 million service members are between the ages of 17 and 24 (Bureau of Labor Statistics, 2008; Congressional Budget Office, 2007). Therefore, the U.S. Armed Forces employ a considerable younger workforce than that found in the civilian workforce (Congressional Budget Office, 2007; National Research Council, 2003).

One of the most revealing differences between military service and other civilian vocational or educational institutions

(or even civilian hiring departments) is that high school transcripts, grade point average, letters of recommendation or even advanced placement courses are not considered or prerequisite in the qualification process to enlist in the U.S. Armed Forces (National Research Council, 2003). Although some previous occupational experience may be considered, assignment into a specific occupation is determined primarily by three factors: the applicant's interest, their ASVAB score and the needs of the specific branch of the U.S. Armed Forces (National Research Council, 2003). This means that the great majority of new enlistees have no experience whatsoever in the occupation for which they are being hired; therefore, the U.S. Armed Forces must somehow turn the young and inexperienced into the skilled and competent (Asch, 1994).

Herein lays the confusion over a theoretical framework - how to identify the exact educational theories that apply when the U.S. Armed Forces' organizational goals include the "acculturation of civilians into the military way of life" (Asch, 1994, p. 6) followed by specialized vocational training necessary to "perform wartime missions" (Wolfowitz, 2004, p. 10). For all practical purposes, enlistees are considered adults; therefore, humanistic theories focused specifically on adult learning must

be considered. Furthermore, the enculturation of enlistees to the U.S. Armed Forces certainly requires a considerable measure of social learning theory; although, mission performance combines elements of both behaviorist theories, related to skill development of competency-based performance objectives, and cognitivist theories, concerned with the development of mental capacity and processes that aid in learning. Therefore an exploration of key characteristics of the above concepts and theories - and their applicability to this study and population group - is necessary to reliably identify a theoretical framework.

Humanism

Humanists view "individuals as seeking self-actualization through learning and of being capable of controlling their own learning process" (Swanson & Holton, 2001, p. 155); therefore, humanistic educators are "concerned with the development of the whole person, the place of emotions and thoughts in personal development, and the emotional and affective dimensions of the personality" (Bertrand, 2003, p. 314). Jean-Jacques Rousseau (internal conflict and alienation), Abraham Maslow (hierarchy of needs and self-actualization) and Carl Rogers (self-esteem and

client-centered therapy) are a few of the primary contributors to humanistic, self-actualizing theories (Bertrand, 2003).

Humanism is a self-actualizing theory based on "congruence, or personal development according to one's own values" where learning is considered "non-directive education" and the learner becomes the "only judge of his or her development" (Bertrand, 2003, p. 311). This implies some measure of maturity often accorded only to adults who possess the capacity to make informed decisions. Therefore, adult education in the form of "adult learning theories, particularly andragogy" and "self-directed learning and much of career development" find their theoretical footings solidly "grounded in humanism" (Swanson & Holton, 2001, p. 155).

Undoubtedly, components of humanistic theory are of significant importance to service members for their progressive career advancement. For this study and this population group, acceptance of a humanistic theoretical framework, and particularly elements of adult education, requires further clarification.

Adult Education

In 1924, with the founding of the American Association for Adult Education (ASAE), adult education "became the total array of community facilities available to meet the needs of the adult learner" (Courtney, 1989, p. 16). This broad definition assumes to encompass the generalized needs of all adults which included such groups as "theater and art groups, libraries, museums, clubs, and voluntary associations" but which also recognized "the more typical kind of organization such as the school or college" (Courtney, 1989, p. 16).

To add to the confusion, adult education practitioners employ a variety of terms, such as "continuing education, lifelong learning, independent learning projects, community education, community development, adult learning, andragogy, adult basic education, animation, facilitation" (Courtney, 1989, p. 15), to describe their practice and the population they serve. Even today, the terms adult education, instruction, and training are often seen as ambiguous with the power to invoke a wide variety of mental images ranging from adult literacy for senior citizens, computer competency certification in office workers, apprentice-to-journeyman welders and non-traditional adults earning a college degree at night or online.

A relatively newcomer to the adult education field who has generated considerable attention (as well as related fields like Human Resource Development) was Malcolm Knowles. Knowles used the concept of andragogy to describe adult learning.

Andragogy

In 1968, Malcolm Knowles introduced the term andragogy to the United States to define core principles and processes that differentiate learning in adulthood from learning as a child (pedagogy) (Swanson & Holton, 2001). Knowles defines andragogy as the "art and science of helping adults learn by having them take charge of their own individual experience and personal characteristics" therefore, andragogy is "centered on personal learning and on development instead of traditional teaching and instruction" (Bertrand, 2003, p. 316).

Later in 1998, Knowles, Holton and Swanson generated a conceptual framework depicting andragogy in practice systematically to better apply Knowles' concept of andragogy across multiple educational adult learning domains (Knowles, Holton, & Swanson 1998). Andragogy in practice identifies three integrated dimensions called rings that influence adult learning, which "are: (1), Goals and Purposes for Learning (2) Individual

and Situational Differences, and (3) Andragogy: Core Adult Learning Principles" (Knowles, Holton, & Swanson 1998, p. 181).

The outer ring - Goals and Purposes for Learning - serves to shape and mold the learning experience and are portrayed as developmental outcomes which may fit into three general categories: individual growth, institutional growth, or societal growth (Swanson & Holton, 2001). The accepted view of adult learning is to "think exclusively of individual growth" although, adult learning is also concurrently vital to creating "better institutions" (Swanson & Holton, 2001, p. 163). The final category relates to societal growth as a direct benefit of adult learning that contributes to society as a whole; whereas, the more radical view of adult education is one that transforms society with the utopian view of creating a better world by developing and liberating the people (Swanson & Holton, 2001).

The second ring - Individual and Situational Differences - is composed of variables that reflect differences that affect adult learning grouped into three categories: individual learner differences, subject-matter differences, and situational differences (Swanson & Holton, 2001). Individual learner differences are those psychological factors (such as cognitive abilities and learning style preferences) that make the learner

unique and must be addressed within the context of the andragogical approach to adult learning (Swanson & Holton, 2001). Subject-matter differences imply that there is not a single method of teaching that will generalize to all teaching situations; simplistically, the type of the subject matter (complex, technical or novel) will dictate the training strategy (Swanson & Holton, 2001). Situational differences are those conditions that affect training such as the likelihood of distance learners to be more self-directed as opposed to traditional classroom settings that usually program all instructional activities (Swanson & Holton, 2001).

The center ring identifies the Core Adult Learning Principles initially introduced by Knowles in 1968 (which he later revised to six, in 1989). These "principles: (1) learners need to know, (2) self-concept of the learner, (3) prior experience of the learner, (4) readiness to learn, (5) orientation to learning, and (6) motivation to learn - are perspectives that come directly from the adult learner" (Knowles, Holton, & Swanson 1998, p. 181). If we explore andragogy's core principles in detail, as they apply to this study, we can better determine if the use of humanistic theories, particularly adult education, are appropriate.

There is little contention that informing learners 'why they need to learn' is relevant as "instruction must be seen as having personal utility (meeting an immediate need), instrumental utility (useful as a means to achieving a future goal), or cultural value (consistent with the values of peers, the organization, or society at large)" (Gagne & Medsker, 1996, p. 176), but this condition of learning is generic and therefore, not limited to adult learners. Knowles (1990) himself writes that if a "diagnosis of needs" is not derived mutually with the learner then the learning design is more an element of pedagogy than andragogy (p. 119). Additionally, the concept of a 'readiness to learn' is also not restricted to the adult learner as all "learners have needs (achievement, affiliation, and influence) that they may strive to meet" (Gagne & Medsker, 1996, p. 176). Knowles (1990) considers that if readiness is a component of "biological development" or due to "social pressure" instead of "developmental tasks of social roles" then the assumption for learning is one of pedagogy and not andragogy (p. 119).

Unlike a civilian employee who is hired based on previous knowledge and experience, the new U.S. Armed Forces enlistee must be trained in every aspect of the organization. For example, if

Delta Airlines sought to expand their aviation engine repair shop to meet a new product line and needed to hire additional hydraulic mechanics they would presuppose that the mechanic not only understood hydraulic principles but also the implied generic knowledge related to aviation such as a common mechanical, structural and electrical systems prevalent in most aircraft. In order to gain the greatest return on their investment, the organization must seek out employees who are already equipped to meet their immediate needs. In this case, Delta Airlines would seek to hire mechanics that already possess a detailed knowledge of aviation system's diagnostics and repair, identification and use of required safety equipment, tools, supply chain management, airfield operations, as well as common practices found in any professional organization.

In the U.S. Armed Forces, the new enlistee possesses minimal if any of the previous knowledge or experience necessary to perform within a specific occupation or organization; therefore, the overall benefit of the learner's prior experience may not be applicable to this study as the majority of learners completely lack any previous experience in their assigned vocations (Asch, 1994). Here again, Knowles (1990) considers when experience is "of little worth" vice a "rich resource for

learning” then the learning assumption is pedagogy and not andragogy (p. 119).

The other three core andragogical principles of self-direction, internal motivation, and self-actualization are addressed by Swanson and Holton (2001) who identify two challenges that may limit the practicality of applying humanistic theories to certain learning situations. They are the need for learners to be self-motivated to develop and improve (internal motivation and self-actualization) and the conflict over humanism’s inherent self-direction and the organization’s “performance paradigm” (Swanson & Holton, 2001, p. 156).

Undoubtedly, the U.S. Armed Forces is an organization comprised of a deeply structured hierarchy that demands collaborative efforts - over individual interests - to “perform wartime missions” (Wolfowitz, 2004, p. 10). Conversely, humanistic theories are focused almost exclusively on the individual in regards to self-directed learning and self-actualization, an individual’s motivation to develop and improve, and the learner’s role in forming, designing, implementing, and evaluating learning outcomes (Swanson & Holton, 2001). This performance paradigm is often criticized as being organizationally oppressive to the individual by “coercing and

demanding behaviors from individuals in return for compensation” (Swanson & Holton, 2001, p. 133). Knowles (1990) stresses that the “andragogical teacher (facilitator, consultant, change agent) prepares in advance a set of procedures for involving the learners (and other relevant parties) in the process” (p. 120). Conversely, in the U.S. Armed Forces, the learner, either current or future, are not involved in course design process beyond inclusion of combined data of previous “student reaction to the instruction, student performance while in the course and graduate performance on the job” (Hanlon, 2003, p. 9).

To the humanist, adult educator, and practitioner of andragogical learning design “the student’s role in the process of development is crucial” and “if you try to change their behavior, you are not humanistic” (Bertrand, 2003, p. 311). Therefore, in regards to this study, humanistic theories as well as adult education theories and specifically andragogy do not appear well suited.

Social Learning

Social learning considers the social context where learning occurs by focusing on “how people learn by interacting with and observing other people” (Swanson & Holton, 2001, p. 156). In 1962, Albert Bandura began to note the “we learn a lot by taking

others as our models, and that the media greatly influences our behaviors" (Bertrand, 2003, 157). Therefore, social learning relies heavily on what the learner witnesses (observes), their modeling of that observation, and then the subsequent consequences of that modeling "as the vicarious consequences of another's behavior will play an important part in strengthening the behavior of the observer" (Lundin, 1985, p. 230). These key components lead to the social learning practices of socialization and mentoring. Socialization is "the informal process that occurs through social interactions" "by which organizations pass on the culture of the organization to new employees to teach them how to be effective in the organization" (Swanson & Holton, 2001, p. 156). Mentoring is the process of teaching and coaching new employees and remains a "primary means of on-the-job training development in many organizations" (Swanson & Holton, 2001, p. 156).

Socialization and mentoring are both significant social learning components used in organizations today and are especially relevant to the U.S. Armed Forces which must inculcate enlistees not only to the military's organization structure and society but also to new unlearned vocations (Asch, 1994; Swanson & Holton, 2001). Social learning theories are considered

progressively relevant in the career development and progress of established employees. Whereas for the purpose of this study, our population group is composed of recent enlistees who receive very little informal socialization during their initial and vocational training and instead are subject to a carefully structured, programmed and systematic regimen (Asch, 1994; Bureau of Labor Statistics, 2008).

Behaviorism

Behaviorism is "primarily concerned with changes in behavior as a result of learning" (Swanson and Holton, 2001, p. 150). Many prominent figures like John Watson (conditioning and generalization), Ivan Pavlov (classical conditioning), Edward Thorndike ("law of effect" on reinforcement), Clark Hull ("reduction of drives theory") and B. F. Skinner (operant conditioning) all contributed models for explaining human learning (Lunden, 1985). Although, only Skinner, and his influential form of radical behaviorism, has earned the position as the modern era psychologist who has been most "misunderstood, misquoted, misjudged, and just plain maligned" (Burton, Moore & Magliaro, 1996, p. 6)

In 1968, Skinner introduced teaching machines that "provided programmed instruction that allowed students to proceed

through lessons by small steps, at their own pace, following an orderly sequence, and receiving immediate reinforcement for every correct response" (Bertrand, 2003, p. 185). Skinner focused on "the construction of carefully arranged sequences of contingencies leading to the terminal performance which are the object of education" (B. Skinner, quoted in Burton, Moore & Magliaro, 1996, p. 17). Therefore, Skinner's "concept of programmed instruction emphasized the need for a total educational plan" which included the criticality of "identifying objectives; arranging subject matter into logical sequences; preparing and testing instructional programs; and then implementing, testing, and revising them" (Bertrand, 2003, p. 185).

Behaviorism has provided the theoretical foundations for many learning innovations "such as teaching machines, computer-assisted instruction, competency-based education (mastery learning), instructional design, minimal competency testing, performance-based assessment, "educational accountability," situated cognition, and even social constructivism" (Burton, Moore, & Magliaro, 1996, p. 8). Furthermore, behaviorism still impacts training "in schools of engineering, on schools of management, and in training development departments" where the

"key words in these areas are "performance," "instructional design," and "instructional strategy development"" (Bertrand, 2003, p. 192). In regards to this study and this population group the behaviorist theoretical underpinnings - especially in regards to instructional design - have particular relevance.

Cognitivism

Cognitivism "arose as a direct response to the limits of behaviorism, particularly the "thoughtless" approach to human learning" (Swanson & Holton, 2001, p. 153). Cognitivists see learners "not as passive and shaped by their environment but as capable of actively shaping the environment" therefore the learning facilitator should be focused on "structuring the content and learning activity so learners can acquire information optimally" (Swanson & Holton, 2001, p. 153). Notable figures include Edward Tolman (Gestalt psychology), Kurt Lewin (organizational development), Robert Gagne (instructional design) and Jean Piaget (cognitive development) (Swanson & Holton, 2001). Cognitivists have significantly influenced adult learning through key contributions in the study of internal mental processes of "acquiring, understanding, and retaining learning" (Swanson & Holton, 2001, p. 153). Other key contributions include information processing model (sensory memory, short-term memory,

long-term memory, encoding, retrieval, and rehearsal), metacognition (control of cognitive processes), and cognitive development over the life span (Swanson & Holton, 2001).

Burton, Moore & Magliaro (1996) suggest that the increased demand for distance learning, has brought a change in the "theoretical landscape" as cognitive psychology has moved "away from its roots in information processing" as the "notion of the mind as a computer has fallen into disfavor" (p. 3). The need for asynchronous, scalable, value-laden learning opportunities, designed to maximize the learner's time, "has brought to the forefront behavioral paradigms that had fallen from favor in many circles" (Burton, Moore & Magliaro, 1996, p. 4). In regards to this study, and this population group, components of cognitivist theoretical underpinnings may not be as relevant as behaviorist theories as new U.S. Armed Forces' enlistees are immersed in entry-level vocational training with the goal of performing rudimentary tasks that support overall mission performance.

Gordon Lippitt and Leslie This (1979) caution that with the abundance of "conflicting theories and conflicting practices within the profession" that researchers must focus on the training objectives and "not become seduced by enchantment with the theories" (p. 17). While there is some confusion over which

theoretical school instructional design falls under - behaviorist (according to Swanson & Holton, 2001) or cognitivist (according to Bertrand, 2003) - they both are relevant to this study. Therefore, having reduced the viability of social learning, humanistic, and to some extent cognitivist theories, behaviorist theories - especially those concerned with systematic instructional design - have particular relevance to this study.

Systematic Instructional Design

Behaviorism is "prominent in the roots of the systems approach to the design of instruction" (Burton, Moore, & Magliaro, 1996, p. 18) but the systematic and structured design of learning appeared long before behavioral theories were introduced.

In 1886, Emerson E. White wrote *Elements of Pedagogy* for teachers and Normal Colleges under the belief that the time had "come for such a study of school education as will ascertain the limitations of its maxims and the coordination and harmonizing of its apparently conflicting methods" (p. iii.). Although, White (1886) does focus a great deal on the teaching of children and youth, he did not exclude the training and education of adults and describes the single "comprehensive end of education is to prepare man to fulfill the purposes of human existence: i.e., to live completely" which includes the "perfection of man's nature

for his highest well-being and happiness, and his preparation for the right discharge of all the obligations and duties which spring from his relations to his fellows, to society, to the state, and to God" (p. 97). White (1886) went further in defining the need for a systematically structured process by describing the appropriate method of teaching as a "series of teaching acts so arranged as to attain a definite end or result. Method is more than the manner or way of an act or several acts. It involves a systematic arrangement of a series of acts, an orderly and rational procedure to a given end" (p. 137). His thoughtful considerations were later echoed in the early 1900s, when Edward Thorndike also displayed "an interest in learning theory and testing" and this "interest greatly influenced the concept of instructional planning and the empirical approaches to the design of instruction" (Burton, Moore & Magliaro, 1996, p. 18).

Based on research conducted during the 1930's, Ralph Tyler employed a systematic process by describing the purpose of stating learning "objectives is to indicate the kinds of changes in the student to be brought about so that the instructional activities can be planned and developed in a way likely to attain

those objectives; that is, to bring about these changes in students" (Tyler, 1949, p. 45).

During World War II, researchers, psychologists and military trainers "based much of their work on instructional principles derived from research on human behavior and theories of instruction and learning" and began stating training outcomes in "terms of "performance" and found the need to identify specific "tasks" for a specific job" (Burton, Moore & Magliaro, 1996, p. 18). Burton, Moore and Magliaro (1996) effectively describe the role of the U.S. Armed Forces played in the development of learning theories related to training as:

"The role and importance of military research during World War II and immediately afterward cannot be underestimated either in terms of amount or results. Research studies on learning, training materials, and instruments took on a vital role when it became necessary to train millions of individuals in critical skills necessary for military defense. People were selected and trained to operate complex and complicated machine systems (i.e., radio detection, submarine control, communication, etc.). As a result, most of the focus of the research by the military during and after the war was on the devices for training, assessment, and troubleshooting complex equipment and instruments" (p. 15).

In 1959, Donald Kirkpatrick wrote a series of articles "designed to stimulate training directors to take a penetrating look at evaluation" (Kirkpatrick, 1979, p. 178). Later, Kirkpatrick (1994) moved beyond evaluation by adding "suggestions for planning and implementing the program to ensure its effectiveness" (p. 3) which included determining needs and subject content, setting objectives and selecting participants, appropriate facilities and instructors. These are all components of systematic design of instruction.

In 1962, R. M. Gagne used the term *learning hierarchy* to "refer to a set of specified intellectual capabilities having, according to theoretical considerations, an ordered relationship to each other" (Gagne, 1968, p. 93). Gagne (1968) suggested that it was possible by "beginning with a clear statement of some terminal objective of instruction, to analyze this final capability into subordinate skills in an order such that lower-level ones could be predicted to generate positive transfer to higher-order ones" (p. 93).

Rosenberg (1982) described the basic instructional systems design (ISD) model as critical for accountability in training as it "provides a procedure for systematically identifying and manipulating significant components which make up the

instructional process, the goals of which are increased learning and improved performance" (p. 44). The ISD model identifies the five processes of analyze, design, develop, implement and evaluate (which later became known as the ADDIE process) that serve as a flowchart to assist training designers in creating meaningful and effective training programs (Molenda, 2003; Rosenberg, 1982).

The Critical Events Model (CEM) is similar in many respects to the basic ISD models as it uses a systematic design process to "improve performance through more effective learning programs" (Nadler & Nadler, 1994, p. 5). Nadler and Nadler (1994) identify nine process events in CEM. They are, identifying the needs of the organization, specifying job performance, identifying learner needs, determining objectives, building curriculum, selecting instructional strategies, obtaining instructional resources, conducting training and evaluating outcomes. Although, these events are not uncommon in many other learning models they are unique in their placement. While many other models place evaluation in a linear sequence at the end of the design process CEM places evaluation and feedback in the center to provide a "hold-pattern" for constant interface of examining the "processes

and results" that occurred during "one or more of the events" (Nadler & Nadler, 1994, p. 17).

Swanson and Holton (2001) extend systems theory into human resource development (HRD) and personnel training and development (T&D) where T&D "is the process of systematically developing expertise in individuals for the purpose of improving performance" (p. 90). The Human Resource Model of Human Resource Development within the Organization and Environment (Swanson & Holton, 2001) is a five-phase model of inputs, processes and outputs enclosed within the organization and the organization's environment. Their model provides organizational leaders with a graphic portrayal of many of the internal and external forces that impact organization.

Dick, Carey and Carey (2005) describe the "elegance of a generic systematic instructional design process is its inherent ability to remain current by accommodating emerging technologies, theories, discoveries, or procedures" (p. xiii). Their model, called the Dick and Carey Systems Approach Model for Designing Instruction is distinctly different from the basic ISD, concentrating more practical effort on designing instruction than implementing instruction.

Bertrand (2003) suggests that with all "instructional design theories, teaching is built around four basic procedures: (1) some material is presented step by step to the student, followed by a question or an activity, (2) the student makes a response or does something, (3) the teacher or the computer gives feedback, and (4) everybody moves to the next step" (p. 194). Bertrand (2003) provides a comprehensive list of behaviorist associated instructional methodologies which include: "competency-based education (considered as a method and a concept), criterion-referenced instruction, programmed instruction, computer-based or computer-assisted instruction, mastery learning, teaching machines, contract learning, personalized system of instruction, individually guided education, and individually prescribed instruction" (p. 186-187). Furthermore, if "instruction is a set of events systematically organized to encourage the internal learning processes" and the goal of learning is to create an operational structure generic to multiple learning scenarios, then education becomes "nothing more than instruction" (Bertrand, 2003, p. 201).

Bertrand (2003) believes that educational systems design "reached a certain maturity in the 1970s, and its basic principles have changed little since then" and that "in short,

educational theories based on systems theory are all very much alike" (p. 186).

During World War II, Fred Keller, a self-described reinforcement theorist, describes his experience of training Morse-code signals to Signal Corps personnel. Keller recalls that instruction in a military training center was "highly individualized, in spite of large classes" with "clear specification of terminal skills for each course, together with carefully graded steps leading to this end" coupled with a "demand for perfection at every level of training and for every student" (Keller, 1968, p. 184). Keller mentions the "employment of classroom instructors who were little more than the successful graduates of earlier classes" and the "minimization of the lecture as a teaching device and the maximizing of student participation" (Keller, 1968, p. 184).

What Keller recalled concerning military instruction during World War II is still clearly evident in the U.S. Armed Forces today. Specifically the U.S. Marine Corps training programs use a form of instructional systems design called the Systems Approach to Training (SAT).

Summary of Theoretical Framework

Although there was some initial confusion assigning the study to a specific theoretical school, it seems evident the instructional systems design, as the product of the best attributes of behaviorist and cognitivist thought, is the only applicable theoretical framework sagaciously relevant and practical to this study.

Background Information - The U.S. Armed Forces

This section provides an overview of the U.S. Armed Forces to familiarize those readers who have not served and to assuage any misconceptions surrounding national military service. Also, this chapter provides a literature review of associated research.

In 1796, Thomas Jefferson wrote: "A debt of service is due from every man to his country proportioned to the bounties which nature and fortune have measured to him" (Foley, 1900, p. 269). Jefferson wrote those words to a new nation whose views of self-rule and freedom were immediate and intimately personal to every citizen. Jefferson's words have affected many Americans over the centuries including John Kennedy who echoed the sentiments in his 1961 inaugural "ask not" address that reminded Americans that "each generation of Americans has been summoned to give testimony

to its national loyalty" and that the "graves of young Americans who answered the call to service surround the globe" (Kennedy, 1961).

In 1988, during the midst of the Cold War, the Democratic Leadership Council (1988) reaffirmed the necessity of national service by suggesting that "No obligation is more fundamental to citizenship than that of preserving our free institutions" and also stressing that "while all wage-earning citizens contribute to the common defense by paying taxes, only a fraction assume the personal risk and sacrifice of military service" (p. 25). Therefore, America's national service is built upon a sense of patriotic duty and national pride where the "nature of the service - to put one's life on the line during wartime - is obviously a unique aspect of the military" (Asch, 1994, p. 4).

Ultimately, America's success and longevity depends on a defensive strategy that promotes our nation's ability "to assure allies and friends that the United States can fulfill its security commitments, to dissuade adversaries from undertaking operations that could threaten U.S. interests, to deter aggression by swiftly defeating attacks and imposing severe penalties for aggression, and to decisively defeat any adversary if deterrence fails" (National Research Council, 2003, p. 18-19).

The effectiveness of this strategy is greatly dependent on the quality and caliber of those responsible as the "education and aptitude of Service personnel has a direct relationship to mission performance" (National Research Council, 2003, p. 25).

A strong national defense requires individuals who can perform diverse tasks such as "run a hospital, command a tank, program computer system, operate a nuclear reactor, or repair and maintain a helicopter" (Bureau of Labor Statistics, 2008, p. 1). This requires the U.S. Armed Forces to be comprised of "dedicated, highly educated, and well-trained men and women capable of leveraging new ideas" (U.S. Office of Management and Budget [OMB], 2005b, section 1.1). Therefore, careful consideration of the potential candidates as well as the myriad training programs and systems is critical as the "men and women of our U.S. Armed Forces are the nation's most important strategic resource" (U.S. Office of Management and Budget [OMB], 2005b, section 1.1).

The U.S. Armed Forces is "often heralded as the nation's (if not the world's) largest educational and training institution" (Laurance, 2004, p. 14) while "the largest single provider of on-the-job training and experience in the U.S. economy is the U.S. military" (Goldberg & Warner, 1987, p. 62).

Unlike civilian organizations which hire an individual based on existing skills and education the "military provides the training and then hires the individual upon completion of the training" (Asch, 1994, p. 6).

Due to the nature of the U.S. Armed Forces where youth, strength, endurance and aptitude are critically essential to meeting operational goals service members are younger than their civilian cohort as nearly 50 percent of the "active-duty enlisted force is between the ages of 17 and 24, compared with just 19 percent of civilians of prime working age (17 to 49)" (Congressional Budget Office [CBO], 2007, p. 19). Additionally, of the over two million young men who turn 18 each year nearly 200,000 enter the U.S. Armed Forces (Congressional Budget Office [CBO], 2007; Kleykamp, 2006; Powers & Lawler, 2007). This makes the U.S. Armed Forces the "nation's largest employer of youth, with a successful record of imparting occupational skills to military-qualified youth" (Asch, 1994, p. ix) where "well planned and executed training and leadership shape the basic individual characteristics into an effective fighting (or peacekeeping) force" (Laurence, 2004, p. 14).

These young people enlist and are trained to "perform not just in combat jobs as infantrymen and tank crew members but as

mechanics, vehicle and other equipment operators, cooks, police officers and security guards, medical technicians, air traffic controllers, electronic and other equipment repairers, computer operators, dispatchers and clerks" (Laurence, 2004, p. 1). The reason an individual chooses to enlist are unique to each person but regardless of the motivation, the "military is the largest employer of youth" (Asch, 1994, p. 1).

Secondary Societal Purpose

While the U.S. Armed Forces exist to support the defense of the nation and those "who enlist are performing a patriotic duty by putting country ahead of self" (Asch, 1994, p. 4), military service also serves a secondary role of "preparing youth, especially disadvantaged youth, to become productive members of the workforce and society" (Asch, 1994, p. iii). What is interesting is that rather than being located at the opposite spectrum as college students, the U.S. Armed Forces also provides a path for the non-college population who for differing reasons were unable to go directly to college from high school. Kleykamp (2006) reports that young people who "aspire to attend college are more likely to join the military than work or pursue some other activity one year after high school graduation" (p. 286).

Myth of the Representative Population

It is important here to address the long-held and accepted belief that the all-volunteer force (created in 1973 after the Vietnam War) would be composed exclusively of minorities and the disadvantaged (lower-income) because the "young people from middle-and upper-income households would opt out of military service altogether" (Congressional Budget Office [CBO], 2007, p. 27). It was commonly accepted that the all-volunteer force would leave the "poor and underprivileged alone to defend us while our more fortunate sons and daughters take a free ride, forging ahead with their education and careers" (Democratic Leadership Council, 1988, p. 25). The belief that the U.S. Armed Forces would not be fairly representative of the American society was repeated by the Democratic Leadership Council (1988) when they warned that the military had become for "many low-income Americans, and particularly minorities, an employer of last resort" (p. 25).

The demographics of students who enlist in the U.S. Armed Forces have changed dramatically over the last decade when more individuals who tested well but were "in the bottom income" were "most likely to enlist" (Akerhielm, Berger, Hooker, & Wise, 1998, p. 17). A 2007 report issued by the Congressional Budget Office (2007) found that instead of being composed of mostly

disenfranchised Americans, the U.S. Armed Forces enlisted young people who were roughly represented from all income groups, but that "young people from the very highest- and lowest- income families may be somewhat less likely than others to join the military" (p. 27). In 2008, researchers Watkins and Sherk (2008) strongly suggested that the "facts do not support the belief that many American soldiers volunteer because society offers them few opportunities" and surprisingly that "Low-income neighborhoods were underrepresented among enlisted troops" (p. 2). Using data from the Defense Manpower Data Center, Watkins and Sherk (2008) examined the "recruits racial and ethnic background, their educational attainment when they enlisted, and information connecting recruits to their home census tracts" (p. 2) and discovered instead of coming from the bottom income levels that the "middle-class and high-income neighborhoods were overrepresented" (p. 2). One reason for this lack of representation in the lower income areas and associated home census tracks with fewer enlistments, could be because "many minorities are screened out by the military's stringent requirements" (Asch, 1994, p. 1). It is relevant to note that the "two most important qualifications for military service are aptitudes (as measured by the AFQT - the U.S. Armed Forces

Qualification Test) and a high school diploma. Other qualifications include good physical health and moral character (e.g., no criminal record)" (National Research Council, 2003, p. 70).

Therefore, lack of representation of this population group could be due to a larger educational issue facing our society where only "half of the high school students in the nation's 50 largest cities are graduating in four years, with a figure as low as 24% in Detroit" (Fields, 2008, p. A13). This is particularly relevant to lower incomes as high school dropouts are "disproportionately represented in the criminal justice system, where about 75% of state prison inmates did not attain high school diplomas" (Fields, 2008, p. A13). Rather than shutting doors on minorities and the disadvantaged it appears the U.S. Armed Forces is only interested in enlisting young people who can pass their entrance requirements which includes a "large fraction of the most qualified minority youth" (Asch, 1994, p. 1) and serves as a "source of social mobility for disadvantaged minorities during service because of the less discriminatory environment" (Kleykamp, 2006, p. 273).

Also, it appears that instead of being filled with the least desirable components of society the U.S. Armed Forces may

now be seen in a more positive light where quality is key and average “entrants into the military are superior in ability to those in their cohort in the general population” (Fredland & Little, 1980, p. 57).

The Office of the Assistant Secretary of Defense for Personnel and Readiness (2008) reports that 99.5% of enlisted service members (who make up about 84 percent of the U.S. Armed Forces) have either a high school diploma or equivalency compared to only 82.18% of those 18 to 23 year olds in the civilian population (p. 12). Table 2.1 demonstrates this comparison.

Table 2.1

Educational Tiers of Active Duty Enlistees to Civilian Comparison Group

Education Tier	Measure	2008 DoD Active Duty Accessions	Civilian Equivalent
1 or 2	High School Graduate or Alternate Credential Holder	99.49%	82.18%
3	Non-Graduate	.51%	17.82%

Note. Educational Tiers 1 and 2 were combined for both civilian and active duty accessions.

Additionally, nearly 95 percent of officers (who make up 16 percent of the force) possess a four-year bachelor's degree compared to only 25 percent of their civilian counterparts between 22 and 27 years of age (Watkins & Sherk, 2008; Bureau of Labor Statistics, 2008).

Enlistment Process

To maintain the 2.6 million service members who serve in the active duty forces (1.38 million - 1.16 million enlisted personnel, 206,000 commissioned officers and 18,000 warrant officers) and reserve (1.2 million) components, the U.S. Armed Forces must recruit 168,000 annually. The U.S. Army is the largest of the active-duty services with 540,000 enlisted personnel, commissioned and warrant officers while the U.S. Marine Corps is the smallest with around 200,000. The U.S. Reserve forces, comprised of the Army, Air Force, Navy and Marine Corps Reserve and the Army and Air National Guard, total 715,000 enlisted, and 122,000 commissioned and warrant officers (Office of the Assistant Secretary of Defense for Personnel and Readiness, 2008).

For those who elect to enlist, service in the all-volunteer force begins with the enlistment process. In 2008, 333,000 applicants sought to enlist in the U.S. Armed Forces of which 55

percent - or roughly 185,000 - men and women were accepted for enlistment for active duty (Office of the Assistant Secretary of Defense for Personnel and Readiness, 2008).

Enlistment in the U.S. Armed Forces is very different from entering the civilian work force, or even entering college, and should be considered carefully. It is a decision that should not be taken lightly. Besides accepting the rules, regulations, order and discipline inherent in military life, the enlistee should understand that the military also controls the location of employment, the hours of work, the type of housing and meals and even the leisure activities of service members. So while the contract does provide employment security, vocational training and health benefits there are certain associated constraints unique to the U.S. Armed Forces (National Research Council, 2003). The U.S. Armed Forces enlistment process begins when an individual, typically a high school senior or graduate between 17-21 years of age, contacts, or is contacted, by a military recruiter. The recruiter, after determining if the candidate is initially qualified and possesses a high school diploma (or equivalent), will schedule an enlistment physical exam and an aptitude exam (if not taken earlier) at the regionally located

military entrance processing station (Asch, 1994; Laurence, 2004; National Research Council, 2003; Powers & Lawler, 2007).

Qualifications

Minimum U.S. Armed Forces enlistment standards require that interested applicants must be at least 17 (with parental approval) but not more than 29 years of age for the Marine Corps, 35 years of age for the Navy and Air Force, and not more than 42 for the Army (Bureau of Labor Statistics, 2008). Additionally, applicants must be a U.S. citizen or legal permanent resident, if married must have no more than two dependents under the age of 18, and if unmarried cannot have custody of any dependent under 18 (National Research Council, 2003). U.S. Armed Forces physical and medical requirements are focused on the applicant being capable of completing required training and performing their duties, shows adaptability to military environments, and is free from any contagious disease that could endanger the health of other personnel (National Research Council, 2003).

Enlistment standards for the U.S. Armed Forces are also concerned with the applicant's moral character and uses specific moral standards to disqualify individuals with felony convictions, who are on probation, parole or imprisoned, individuals with prior military service who have separated under

less than honorable conditions, and individuals who exhibit antisocial or other character traits that would preclude them for serving in a military environment (National Research Council, 2003).

The U.S. Armed Forces do allow infrequent waivers for individuals who "despite factions such as youthful misconduct, are judged trustworthy and capable" (Gilmore, 2001, p. 1). About 20 percent of new recruits require some type of waiver with roughly two-thirds issued for minor infractions with the law (such as drug experimentation, traffic violations, or misconduct) and the other third concerned with health concerns (like too-high body fat), low aptitude scores, or other associated issues (Congressional Budget Office, 2007; Gilmore, 2001).

Attrition

The Department of Defense tracks recruiting and training costs necessary to enlist and train new acquisitions. In 2003, the cost of recruiting, screening and testing a prospective applicant was about \$14,206 dollars each (U.S. Army - \$16,536; U.S. Navy - \$13,394; U.S. Air Force - \$9,376; & U.S. Marine Corps - \$9,356), while the average cost of "all training costs related to selected occupations, including recruit training" was about \$10,600 dollars per individual (U.S. Navy & U.S. Marine Corps combined -

\$18,000, U.S. Air Force - \$7,400, & U.S. Army - \$6,400) (U.S. Government Accounting Office, 2005). Therefore, when an enlistee leaves - or attrites in military terms - before their obligation has been met, the U.S. Armed Forces will have allocated thousands of dollars with little to show for their investment.

To increase the likelihood of completing service obligations, which would reduce attrition, the Department of Defense developed a "mathematical model that links educational attainment, aptitude, and recruiting resources to job performance and used that model to establish recruit quality benchmarks" that "reflect the aptitude and education levels necessary to minimize personnel and training costs while maintaining an explicit level of performance" (Chu, 2006, p. 6). Concisely, the benchmarks identified high-caliber enlistees not only as those with higher scores on the aptitude test but also those who demonstrated a commitment to completing high school. It seems that when all other factors were controlled, possessing "a high school degree is probably the single most important indicator that a recruit will complete the first term of enlistment" (Moore and Reese, 2001, p. 7).

Therefore, high school graduates are considered the best prospects in the all-volunteer military as 80 percent complete

their initial obligation (U.S. Department of Education, 1998). Applicants with alternate credentials, like the General Education Development (GED) certificate, are seen as less promising as less than 60 percent complete their obligation and "low rates of persistence have been the major problem for GEDs in the military" (U.S. Department of Education, 1998, p. 2). The U.S. Department of Education (1998) found that although the "GED measures and certifies cognitive skills" it does not measure the "range of other attributes that contribute to persistence" and that "completion of high school demonstrates such attributes and that a high school diploma reflects them" (p. 2).

The Department of Defense set goals of enlisting only high school graduates and high school credential holders at 90 and 10 percent, respectfully (Chu, 2005, p. 2). Therefore with a 50 percent attrition level for high school dropouts they are rarely considered for enlistment (Congressional Budget Office, 2007, Chu, 2005).

ASVAB and AFQT

As the U.S. Armed Forces will invest approximately \$31,000 in the first six months to train a new enlistee (National Research Council, 2003, p. 38) the U.S. Armed Forces needs assurance that the applicant has the "mental aptitude to succeed at that job"

(Powers & Lawler, 2007, p. 17). To do this the U.S. Armed Forces uses the Armed Services Vocational Aptitude Battery or ASVAB as an entrance exam.

The ASVAB is the product of a 1948 Congressional requirement to develop a "uniform screening test to be used by all the services" (Powers & Lawler, 2007, p. 9) which, after over fifty years of Department of Defense revisions became the current ASVAB used today (Powers & Lawler, 2007). The ASVAB is also used as a placement exam that not only measures normal academic subjects like English and math areas, "but it also measures your knowledge of other areas, such as mechanics, electronics, science, and assembling objects" (Powers & Lawler, 2007, p. 1). To do this the ASVAB uses four of nine subtests - Work Knowledge, Paragraph Comprehension, Arithmetic Reasoning and Mathematics Knowledge - to compute a person's U.S. Armed Forces Qualification Test (AFQT) score which "determines whether you qualify to even enlist in the military" (Powers & Lawler, 2007, p. 9). Over the last fifty years, the Department of Defense has perfected the use of AFQT scores to determine higher-quality prospective candidates because there "exists a positive, statistically significant relationship between AFQT scores and performance" (Scribner, Smith, Baldwin, & Phillips, 1986, p. 198) especially since

aptitude scores have consistently "been shown to be an accurate predictor of personnel quality and ability" (Kavanagh, 2005, p. 27). Two studies support this decision. By using firing range success of rounds on target, researchers Scribner, Smith, Baldwin and Philips (1986) were able to demonstrate a 20 percent increase in productivity in personnel with average AFQT scores 60th percentile over personnel with average scores in 20th percentile regardless of years in service, education or race. Moore and Reese (2001) also concluded that when all other factors were equal "recruits who score in the upper 50th percentile of the AFQT have a small but reliable edge over those in the bottom half, and they are less likely to attrite" (p. 7).

To ensure that new enlistees possess the potential for greater production and therefore, the greatest and safest return on their investment, the Department of Defense established AFQT enlistment qualifiers. These qualifiers require that at least 60 percent of new enlistees are from Category I through IIIA (50th and above percentile) and 36 percent of new enlistees are from Category IIIB (31st to 49th percentile). No more than four percent of all applicants with an AFQT percentile score between 10 and 30 (Category IV) and no one with a score below 10 (Category V) may be enlisted (Chu, 2005, p. 2).

In 2008, the U.S. Armed forces exceeded their enlistment qualifiers as 68 percent of recruits scored at or above the 50th percentile on the AFQT. Only 51 percent of the civilian representative sample was able to do the same (see Table 2.2).

Table 2.2

Armed Forces Qualification Test (AFQT) Percentile Scores for Active Duty Enlistees and a Civilian Comparison Group

AFQT Category	Percentile Score	2008 DoD Active Duty Accessions	Civilian Comparison Group
I	93rd to 99th	5.87%	7.86%
II	65th to 92nd	37.67%	27.66%
IIIA	50th to 64th	26.63%	15.54%
IIIB	31ST to 49th	28.10%	18.78%
IV	10th to 30th	1.75%	20.75%
V	1st to 9th	0.0%	N/A
Unknown			9.41%

Note. The comparison group is composed of civilians aged 18 to 23 years of age who took the ASVAB in 1997.

A secondary benefit of setting high enlistment qualifiers was reported by Kavanagh (2005) who observed the effects of AFQT scores were “additive, meaning that each additional high-scoring team member increases the overall performance of the team.” (p. xi). Kavanagh (2005) stressed the relevance to a “military

context, given the number of group-centered tasks the armed forces are required to complete" (p. xi).

Occupational Specialty

Once an individual meets all requirements and scores well on the ASVAB they may be offered a specific military job or a specific job field such as infantry, electronics, law enforcement, aviation, logistics, or maintenance (National Research Council, 2003; Powers & Lawler, 2007). Each active duty service has differing enlistment contracts but all will guarantee a specific job or career area composed of "hundreds of enlisted job opportunities" (Powers & Lawler, 2007, p. 351) but selection is contingent upon formal school availability, the applicant's specific aptitudes measured on the ASVAB, any previous formal training, level of necessary waivers, medical exam results and the current needs of the service (Bureau of Labor Statistics, 2008).

Once the qualified applicant has finalized a career field they sign an enlistment contract that obligates the individual to serve for a specified length of service usually for 4 years but also 2-year, 3-year, or 6-years (Bureau of Labor Statistics, 2008). Applicants may not leave immediately though, as all of the branches of the U.S. Armed Forces provide a delayed entry

program which allows individuals to delay departure to initial training for up to a year - which allows high school seniors to enlist and depart after graduation (Bureau of Labor Statistics, 2008).

Training - Initial

After enlistment, individuals undergo a service specific initial training, better known as "boot camp" or "basic training" used to "develop civilians into military personnel that fit the mission of the Service" (U.S. Office of Management and Budget, 2005a, section 1.3). Although, each service provides "specialized individual and group skills for that Service" (U.S. Office of Management and Budget [OMB], 2005a, section 1.3) the "acculturation of civilians into the military way of life is a key part of basic training" (Asch, 1994, p. 6).

Basic training provides a 6- to 13-week introduction to military life through courses in military skills and protocol, personal health and hygiene, first aid, preparation of uniforms and self-discipline concurrently while developing a "sense of responsibility for their own actions and for those of others in their group" (Asch, 1994, p. 6). Basic training is "carefully structured and include rigorous physical exercise designed to improve strength and endurance and build each unit's cohesion"

(Bureau of Labor Statistics, 2008). Additionally, supervision during basic training is constant and ever-present and strict adherence to military rules and standards is absolute (Asch, 1994).

Training - Vocational

Upon successful completion of basic training, most recruits receive vocational training at formal schools to prepare them for their specific military occupational specialty; although, a few recruits may receive on-the-job training at the first duty station. This training usually lasts from 10 to 20 weeks but some vocational programs - such as advanced electronics, foreign language or nuclear science - could take as long as a year (Asch, 1994; Bureau of Labor Statistics, 2008).

Unlike civilian employers who consider a person's education and experience before hiring, the U.S. Armed Forces does not require a college degree in "computer science before you're hired to become a computer programmer" (Powers & Lawler, 2007, p. 17). The U.S. Armed Forces role is unique as it "trains people primarily in vocational skills directly tied to the job" (Asch, 1994, p. ix) where "individuals earn money and develop additional skills during their service and in some cases military training itself can be counted toward a college degree" (Kleykamp, 2006,

p. 275). It is interesting to note that in the U.S. Armed Forces, only 31 percent of the enlisted vocational training fell within "occupations such as infantry, craftsmen, and service and supply handling" while 43 percent served in mid-level "medical and dental, functional support and administration, and electrical/mechanical equipment repair" vocations, and 21 percent served in high-skill vocations like "electronic equipment repair, communications and intelligence" (Office of the Under Secretary of Defense, 2006, p. 9-10).

Formal School Vocational Training

U.S. Armed Forces Formal Schools are responsible for "specialized skill training, which provides military members with initial job qualification skills or new or high levels of skill in their current military specialty or functional areas" but unlike civilian organizations the "focus is on "occupational" training that relates to skills that individual military members require to perform wartime missions" (Wolfowitz, 2004, p. 10).

The great majority of this specialized skill training is provided by military Formal Schools in traditional classroom or simulated settings (Asch, 1994; U.S. Office of Management and Budget, 2005b). The Department of Defense allocates resources, such as instructor and support staff, facilities, equipment and

operating budget, to meet this goal (Department of Defense, 2008). Every service member that graduates from a Formal School will join a military unit in either the regular or reserve forces; therefore, each branch of service is required to justify their resources against the effectiveness of producing technically trained graduates (U.S. Office of Management and Budget, 2005b).

Military Evaluation of Training

The U.S. Armed Forces, and particularly the U.S. Marine Corps, is also focused on how to "measure academic success and legitimately determine if its students have acquired the requisite knowledge to make them proficient in their professions?" (Stein, 2000, p. 55). In a rare collaborative effort, the U.S. Armed Forces directed policy to standardize training across all branches of the service to ensure that service members from one service would receive the same quality training while attending another service school. The inter-service training directive requires all U.S. Armed Forces' Formal Schools to employ a systematic approach to training in creating and reviewing curriculum with the intent to "reduce or avoid costs, standardize instruction, and increase efficiency" through a series of evaluation milestones and

measures, such as instructor evaluations, for the "purpose of achieving continuous product improvement" (Holcomb, 1998, p. 7).

The Formal Schools measure efficiency through internal and external evaluations which are focused on successfully meeting operational goals and mission objectives (U.S. Office of Management and Budget, 2005b). While external evaluations include submitting scheduled reports to higher authority or independent evaluations from inspection offices, internal evaluations stress continuous review through "instructional systems design and course content review boards" (U.S. Office of Management and Budget, 2005b, Section 2.6).

The Department of Defense uses one performance measure that is very similar to civilian organizations - comparing student graduates against the planned number to determine a completion percentage. The underlying assumption is that "students that successfully graduate from training have met the high standards in their fields and are considered ready to perform missions and join their units in the field" (U.S. Office of Management and Budget, 2005b, Section 2.1).

In the Marine Corps evaluation within the Formal Schools is mandatory and directs managers to "conduct evaluations on an ongoing basis by surveying students, graduates, and their

supervisors" (Hanlon, 2003, p. 9) and to report course attendance, attrition, and graduate numbers. Using established training standards as measures of effectiveness, managers will evaluate instructor's performance by gathering and analyzing data on "student reaction to the instruction, student performance in the course, and graduate performance on the job" (Hanlon, 2003, p. 9).

The Systems Approach to Training (SAT) is a "comprehensive methodology for analyzing, designing, developing, implementing, and evaluating the total process of learning and teaching in terms of specific objectives written to support task performance" to be used as the "methodology for all training and education conducted by operating forces and training institutions" (Gray, 1991, p. 2-3). Like civilian organizations, the U.S. Armed Forces' Formal Schools use the same reaction-type instructional rating forms, compare graduates to enrollees, and employ scheduled reports to determine if their training programs were successful.

The Role of the Military Vocational Instructor

In regards to training providers, the U.S. Armed Forces is quite different from their civilian equivalents that may be more inclined to hire employees who already possess the necessary

skills, or who contract with external agencies to meet any specific training needs that may arise. In the U.S. Armed Forces, military instruction is most often provided by current, usually senior, service members within their occupational specialties who are selected based on their "current qualifications and ability to fill a valid requirement" while other factors such as time of service, availability, volunteer status are secondary considerations (Wolfowitz, 2007, p. 2). Rather than being an institutional burden, senior personnel are critical to the Formal School's role because "training contributes so significantly to performance and productivity, the effectiveness of military performance, as well as overall readiness" (Kavanagh, 2005, p. 25). During a press interview in 1998, former Secretary of Defense William Cohen stressed the importance of utilizing senior military personnel as the "crucial task of training can be accomplished only through the assignment of the most highly skilled and motivated non-commissioned officers who are led by experienced quality officers" (Cohen & Shelton, 1998, March 16).

In the U. S. Marine Corps, personnel are selected for trainer assignments based on their availability to fill a vacant position. Once assigned, U. S. Marines are required to attend a

one-week Formal School instructor course and complete two online courses: the Systems Approach to Training and Operational Risk Assessment (Formal School Instructor Course, 2008, p. 2). Upon completion of these minor requirements, the instructor is given a blank lesson plan to personalize and is then assigned a class to teach. These militarily proficient subject matter experts are filling instructor roles with minimal preparatory time where student reactions to training are often used to rank order instructors against their peers which could ultimately affect their promotion potential and career longevity.

Summary of Literature Review

This chapter discussed the multiple challenges encountered identifying and selecting a practical theoretical framework and creating a functioning conceptual framework. A literature review was conducted and discussed to rationalize the research methodology. A brief description of the U.S. Armed Forces' enlistment process and requirements, initial training procedures, and the vocational training process was introduced as pertinent to this study.

This chapter provided a review of the extensive analysis of student evaluation models and processes. Effective teaching was also reviewed to identify underlying similarities, albeit oft-

conflicting, that could potentially affect the student rating process specific to a military vocational training center.

A brief summary of potential educational theoretic frameworks was discussed which include humanism, adult education, andragogy, social learning, behaviorism, and cognitivism were reviewed. The motivation to utilize the theoretical underpinnings of Instructional Systems Design (ISD), as a merging of the positive elements of behaviorist and cognitivist theories, was discussed to establish a link between theory and this study's bedrock theoretical perspective.

Lastly, this chapter described the process for accession and training within the U.S. Armed Forces. Enlistment parameters, qualifiers, and initial and vocational training, describes a detailed and complex process focused mainly on effectively and efficiently training personnel to fill critical vocational roles in support of the U.S. Marine Corps warfighter. An appropriate comparison of U.S. Armed Forces' new accessions to their civilian counterparts shows an organization comprised of quality and capable young people possessing higher levels of educational accomplishment, testing ability, and physical fitness along with fewer incidents of criminal activity.

CHAPTER 3

Methodology

The purpose of chapter three is to describe the research design, methods, participants, sample, instrumentation, data collection and data analysis. Chapter three discusses the motivation and appropriateness of the research design used for the study as well as the population, sample and sample size determination.

Furthermore, this chapter addresses the development of the instrument, to include selection and creation of statements to assess relevant constructs, subsequent review of statements, and data collection procedures. The data analysis techniques, an explanation of the assumptions for data analysis, missing and outlier data considerations, factor analysis, and reliability as also addressed. The chapter concludes with a review of the data analysis conducted during the initial instrument development.

Research Design

The study employs a quantitative, nonexperimental, nonrandom survey research design with a purposive sample. The study is a nonexperimental survey design because it does control who, what and when to measure but does not manipulate the independent variables or contain a control group. Sproull (2002) validates

the use of a nonexperimental design as appropriate when "the researcher has control over who or what to measure, when the measurement takes place and what to ask or observe" especially "when the researcher is interested in reactions to a specific event" (p. 153). This is appropriate for this study to examine students' ratings of instructor characteristics directly related to the specific event of training that took place at a U.S. Marine Corps vocational training center.

This descriptive study measured variables of interest and analyzed collected data that "describe certain aspects of a group of people" (Holton & Burnett, 2005, p.38). Furthermore, this study can be classified as correlational because it seeks to "determine whether, and to what degree, a relationship exists between two or more quantifiable variables" (Gay, Mills, & Airasian, 2009, p 196). Since the specific focus of the study is the U.S. Marine Corps, participants are limited to U.S. Marine Corps personnel. The study's uniqueness warrants the specific criterion of U.S. Marines currently enrolled in entry-level formalized military occupational (vocational) specialty training; therefore, a nonrandom purposive sampling consisted of paper-based reaction survey results and post-lesson test scores was used. Nonrandom purposive sampling was appropriate for use as

the "sample is arbitrarily selected because characteristics which they possess are deemed important for the research" (Sproull, 2002, p. 119). The study's use of a survey is also warranted as "the survey is now recognized as the most frequently used data collection method in organizational research for assessing phenomena that are not directly observable" (Bartlett, 2005, p. 98). The use of a questionnaire was considered appropriate as an information gathering tool because of its flexibility and applicability to multiple settings (Jonassen, Hannum, & Tessmer, 1989).

The study created an instrument to report reactions to instructor characteristics by collecting data from the participants in the form of responses to statements that are summated into Likert-type scales. The research design supported the collection of reactions employing the Likert technique. According to Gable and Wolf (1993) "the Likert scales have been frequently used because they are relatively easy to construct, can be highly reliable, and have been successfully adapted to measuring many types of affective characteristics" (p. 50); therefore, their use was deemed appropriate for this study.

Research Methods

Data was collected for this study using a nonexperimental research design. The survey instrument collected individuals' reaction to statements related to instructor characteristics. Post-lesson test score data were gathered and matched with each participant's rated instrument responses.

The primary goal of this study was to examine the relationship between instructor characteristics and student perceptions of instructor quality and performance. To achieve the primary goal of the study, it was necessary to develop an instrument that could be implemented to address the research questions and specifically sought to measure the instructor characteristics.

The overarching method for this study is a correlational descriptive design. To create a valid and reliable instrument to implement this method, exploratory factor analysis was used to identify and examine specific instructor characteristics that consistently predict student ratings of perceived instructor quality (on an ordinal scale measured by the instrument) and student learning (on a ratio scale as indicated by post-lesson test scores) recorded at the end of the training iteration.

Because of the multiple and often conflicting theoretical perspectives, exploratory factor analysis greatly aided the creation of a theoretically sound instrument. Yang (2005) describes the "fundamental principle of factor analysis is to explain correlations among a large number of observable variables by identifying or confirming underlying factors that explain these correlations" (p. 183). Therefore, the results help define a practical conceptual model and contribute to a new theoretical framework or at a minimum support existing systematic design theory.

Research Questions

The study sought to examine U.S. Marine Corps vocational student's reaction to instructor characteristics using a quantitative research design. There were six research questions considered in the study:

Research Question 1 - *What are the instructor characteristics that can be used to describe instructors and explain student's reaction toward instructor quality and student learning of lesson material?*

Research Question 2 - *How do students perceive instructor characteristics as measured by the developed instrument?*

Research Question 3 - *What are the student reactions to perceived instructor quality as measured by the instrument and student learning as reported by post-lesson test scores?*

Research Question 4 - *Are there relationships between the instructor characteristics and student's reaction to perceived instructor quality as measured by the instrument and student learning as reported in post-lesson test scores?*

Research Question 5 - *Do instructor characteristics explain a significant amount of variance in perceived instructor quality?*

Research Question 6 - *Do instructor characteristics explain a significant amount of variance in post-lesson test scores?*

Research Variables Used in the Study

The conceptual framework illustrated in chapter one provides a representation of the variables used in the study. The variables used in each of the research questions will be described in detail in the following paragraphs.

The first research question sought to identify the instructor characteristics that could be used to describe instructors and explain student's reaction toward perceived instructor quality and student learning of lesson material. Table 3.1 lists the independent variables (instructor

characteristics) and dependent variables (student's reaction to perceived instructor's quality and student learning) found in the six research questions.

The second research question identifies how the students perceive instructor characteristics as rated on the instrument. Means and standard deviations report the student's perception of the instructor characteristics identified by the instrument's nine factors.

Research question three examined the student's reaction to the perceived instructor's quality as measured by the instrument and student's learning as reflected by post-lesson test scores.

Research question four sought to report the relationship between the student's reaction to perceived instructor quality as measured by the instrument and student learning as reflected by post-lesson test scores.

Research question five sought to explore if instructor characteristics explain a significant amount of variance in perceived instructor quality.

Finally, research question six sought to explore if instructor characteristics explain a significant amount of variance in post-lesson test scores.

Table 3.1

Variables Used to Describe Instructors and Explain Student Reactions of Instructor Quality and Learning

	Variables
Independent Variables	Instructor Characteristics
	Effective Delivery
	Learning Environment
	Enjoys Teaching
	Clear Communication
	Subject Expertise
	Provides Feedback
	Concern for Learning
	Fair and Respectful
	Available and Relevant
Dependent Variables	Instructor's Quality
	Post-Lesson Test Scores

Population

The population of the study was adult, entry-level service members currently serving in the U.S. Marine Corps and attending post-secondary, initial entry-level vocational skills training at a U.S. Marine Corps Formal School. The population includes male

and female high school graduates or equivalent, between 18 and 30 years of age. All participants are considered qualified, by a vocational aptitude exam, and capable of successfully completing the training as well as successfully performing the vocational duties assigned upon graduation of training. Participants not meeting these criteria were considered ineligible and were not solicited for participation. Participants may or may not have selected the specific occupational field enrolled. Participants were solicited from available training sessions during May 2010 to July 2010 of survey collection.

Sample and Sample Size

Nonrandom purposive sampling was utilized for the study. Using a nonrandom sampling method has the potential for large sampling errors and bias (Sproull, 2002) but for the purpose of this study we were more concerned with the instrument's internal validity - how well it measures what it's supposed to - than to be able to generalize the findings. Additionally, with the homogeneous group it was more important the participants met specific criterion than be randomly sampled.

Nine independent variables and two dependent variables were used to determine the study's sample size. Hair, Black, Babin, Anderson & Tatham (2006) recommend that for successful factor

analysis, researchers try to “obtain the highest cases-per-variable ratio to minimize the chances of overfitting the data (i.e., deriving factors that are sample-specific with little generalizability)” (p. 112). A ratio of 5 to 20 cases per variable with no fewer than 50 cases, but preferably 100 or more, observations is considered appropriate to reduce the chances of overfitting (Hair et al., 2006, p. 112). With nine independent variables and two dependent variables an appropriate sample size for this study would range from 55 to 220 cases. To determine the appropriate sample size to conduct the multiple regression analysis, Bartlett, Kotrlik & Higgins (2001) suggest a ratio of 5 cases per each independent variable, but a more conservative 10:1 ratio is desired. For this study there were nine independent variables and the sample size was between 45 to 90 cases. However, others such as Hair et al. suggest a more conservative sample size is preferable, such as a 15:1 or 20:1 ratio, or for this study 135 to 180 cases should be obtained. The field test first round collected data from 319 cases and the second round collected data from 139 cases.

Instrument Development - Item Generation

The instrument used in this study was developed in the form of a survey questionnaire. After the first round of data collection,

responses to the original instrument were analyzed and a revised version was produced. The instrument's collection of demographic data and post-lesson test scores was not changed in the second version.

The instrument was initially developed using an exhaustive review of literature. Items were selected from survey instruments already in use to evaluate instructor performance from a wide range of fields and academic settings. A comprehensive search of databases was conducted using terms appropriate to teacher evaluations, instructional effectiveness, attributes and techniques, and evaluation. Multiple disciplines were examined. Instrument statements were drawn from the review of student rating instruments found within multiple peer-reviewed studies (Abrami, d'Apollonia & Rosenfield, 2007; Aleamoni, 1981; Bearor, 2004; Bell, 1962; Braskamp, Caulley & Costin, 1979; Centra, 1973; Crader & Butler, 1996; Davidovitch & Soen, 2006; Feldman, 1996; Feldman, 2007; Fenker, 1975; Gable & Wolf, 1993; Guthrie, 1953; Jackson, Teal, Raines, Nansel, Force & Burdsal, 1999; Kember, Leung & Kwan, 2002; Leach, 1996; Lee & Pershing, 2002; Mandebvu, 1998; Marsh & Bailey, 1993; Marsh, 1984; Marsh, Fleiner & Thomas, 1975; McBean & Al-Nassri, 1982; Milano & Ullius, 1998; Mowrer-Reynolds, 2008; Murray, 2007; Nadler &

Nadler, 1994; Onwuegbuzie, Witcher, Collins, Filer, Wiedmaier & Moore, 2007; Patrick & Smart, 1998; Powers, 1992; Pregent, 2000; Rae, 1993; Richmond, Gorham & McCroskey, 1987; Ryan, Anderson & Birchler, 1980; Shevlin, Banyard, Davies & Griffiths, 2000; Spooren & Mortelmans, 2006; Thompson, 2001; Wilson, 2006; Wortruba & Wright, 1975; Young & Shaw, 1999) based on the instrument's 'closed-end' response design; i.e., Likert-type scale, multiple choice, etc. This review produced 1514 initial items specifically related to the study's focus of student's reaction of the instructor characteristics.

With the overwhelming amount of available research focused on education, training, teaching, instructing, and learning, it quickly became necessary to apply concise parameters to discern the appropriateness of individual items. Since the focus of the study remained specific to instructor characteristics only those items implicitly within the instructor's sphere of influence were considered appropriate.

The considerable pool of items required the development of a relevance matrix to standardize the selection and removal process. Initially, all items were considered and assigned into similar broad categories. The goal was to develop factors composed of multiple items that were "mathematically grouped

together to measure a variable" (Holton & Burnett, 2005, p. 36). Items were collected and converted into statements that were clearly stated, easily understood and worded without ambiguity, and structured to provide a controlled response (Sanders, 1998, p. 188).

Further refinement was conducted to remove items from the instrument that dealt with facilities, the entire course or curriculum design thought outside the instructor's control as well as duplicate items or duplicative authors. Statements that were "double-barreled"; i.e., that failed to identify a single action, were discarded (Hinkin, 2005, p. 166). Incomplete, redundant or vague statements where the participant was likely to lack sufficient information to accurately rate, and items considered incongruent or irrelevant to an entry-level military training population, were also discarded. The 20 reasons for deleting items and statements are listed in Appendix A.

Subsequent screening of the remaining items identified multiple negative-language items. Negative-language items were often the polar opposite of positive-language items retained in the survey; therefore, as the content was usually quite similar less the dichotomous nature, items were combined with the retained item regardless of positive- or negative-language

affect. Those items that were identical or grossly similar were combined and all authors were identified on an author-to-item external spreadsheet to ensure proper referencing occurred. Finally, due to the magnitude of the item pool, no more than five statements were placed within each of the sub-factors.

The researcher, and an expert from the field and chair of this study, conducted two separate face-validity content analyses. The statements were reviewed and a 98-item instrument was created to negate any possible response burden on participants. It was important to retain items that adequately assessed participant reactions to specific instructor characteristics as identified in the emergent themes.

After combination, conversion, screening for appropriateness, removal of polarized items, and two review sessions to assess content and face-validity, 1416 items were removed for this study. Of the original 1514, 98 items within 9 factors and 35 sub-factors remained.

Once applicable items were identified, additional standardization was employed during instrument creation. As all items were initially transcribed verbatim from their source documents it was necessary to reduce confusing jargon; therefore, any "presenter" to include teacher, trainer, professor, lecturer,

speaker and/or staff member was changed to "instructor" to meet this specific population group. The plural "instructor's" were changed to singular "instructor". Participants and trainees were standardized to "students" while training session, session, course and class were standardized to "lesson". Grammatical tense was changed from present to past tense; i.e., "uses" to "used", "explains" to "explained", etc. All items began with the heading "The instructor:" followed by individual statements and response scale requiring a measure of agreement.

The initial instrument employed a Likert-type scale to record the participant's level of agreement to statements related to instructor characteristics. The response scale was built using a six-option Likert-type scale similar to those found in Gable & Wolf (1993). A six-option scale was used purposefully to force a positive or negative response; i.e., there were no "non-applicable" or "no opinion" options. Although, Hinkin (2005) suggests that participants must be provided an opportunity to "opt out of answering a question if it does not apply to his or her situation" (p. 167) for the purposes of this study, every participant response was considered applicable to their situation and relevant enough to remove this neutral "neither agree nor disagree" midpoint scale. Scale choices for this study were:

strongly disagree (would disagree most of the time), moderately disagree (would frequently disagree), slightly disagree (would occasionally disagree), slightly agree (would occasionally agree), moderately agree (would frequently agree), and strongly agree (would agree most of the time).

Reverse wording (or coding) is a procedure used not only to stimulate the respondent into concentrating on their responses but also to reduce the potential likelihood of acquiescence bias (where the participant responds in singularly positive way regardless of the item or question) or response set (when participants respond based on their general feelings about the topic instead of the specific item) (Woods, 2006).

Two items required the participants to select a different measurable scale related to instructor quality. The scale to assess instructor quality was: one of the best, very good, above average, below average, very bad, and one of the worst. The second instructor quality statement was assessed with participants responding: strongly agree, moderately agree, slightly disagree, slightly agree, moderately disagree, and strongly disagree.

The field-test instrument also collected participant demographics including birth year, gender (male or female), state

enlisted, ethnicity (American Indian/Alaskan Native; Asian; Black; Hispanic; Pacific/Hawaiian Islander; White and other), and highest education level gained (general equivalency diploma, high school diploma, 2 years college, 4 years college, or 6+ years of college).

Data was collected with a paper based instrument due primarily to the vocational nature of the training environment. The lack of computer access made paper-based instruments the best available option; even though more time was needed for data entry. Student post-lesson test scores were collected and recorded on each instrument after the student completed the survey.

Factor Naming

The initial instrument was constructed around nine themes. A deliberate item content review was undertaken to assign individual items into appropriate factors, *Effective Delivery*, *Learning Environment*, *Enjoys Teaching*, *Clear Communication*, *Subject Expertise*, *Provides Feedback*, *Concern for Learning*, *Fair and Respectful*, and *Available and Relevant*.

Factor one was named *Effective Delivery* and included items to rate the student reactions to how effective the instructor

delivered the lesson. Each sub-factor and item described the instructor's ability to successfully present the lesson material and stimulate and hold student's interest during well-prepared lessons designed to enhance student learning. The 8 sub-factors and 25 items introduced in this factor are shown in Appendix B. One of the most repeated topics found in multiple studies concerning student rating of instructor's effectiveness was a communicative, enthusiastic delivery (Lowman, 1994), regardless of whether the lesson content was good or poor (Polk, 2006) or even when the content was inaccurate (Madsen, 2003). Instructors that implement well-designed, organized and prepared lessons and who follow outlines closely in a well-managed classroom positively contribute to the student's sense of belonging (Freeman, Anderman & Jensen, 2007). Instructors rated highly on this factor are perceived by students as employing successful training techniques, free from distracting mannerisms that stimulated and enhanced the lesson material.

Factor two was named *Learning Environment* and included items to rate the student reactions to the instructor's management of the learning environment that was conducive to learning. Each sub-factor and item reflected the instructor's time management, punctuality and classroom management abilities.

The two sub-factors and six items in this factor are shown in Appendix B. The ability of the instructor to manage the learning environment is strongly associated with effective teaching performance (Madsen, 2003; Opdenakker & Van Damme, 2006; Plax, Kearney, & Downs, 1986; Polk, 2006, Turman & Schrodt, 2006) while instructors who are unable to effectively monitor and control students in overcrowded classrooms "report declining levels of satisfaction" (Plax, Kearney, & Downs, 1986, p. 380). Students rating instructors highly in this factor perceived the instructor's competency at pacing the lesson, staying on topic within lesson time - all while maintaining an environment that was safe, orderly and conducive to learning.

Factor Three was named *Enjoys Teaching* and included items to rate the student reactions to the instructor's apparent enjoyment of teaching and preparedness. Each item and sub-factor centered on the instructor's enthusiasm, use of humor, and interest in training students. The two sub-factors and four items in this factor are shown in Appendix B. An instructor's enthusiasm and appropriate use of humor makes the class interesting while improving the learning climate (Wanzer, Frymier, Wojtaszczyk, & Smith, 2006) because the "passion for teaching and learning does not go unnoticed by students" (Polk,

2006, p. 24). Instructors rated highly on this factor are perceived by students as showing interest, enthusiasm and genuine enjoyment in training students.

Factor Four was named *Clear Communication* and included items to rate the student reactions to the instructor's communication skills. Each item and sub-factor described the instructor's ability to effectively and clearly communicate in a manner that clarified lesson material by using language that students understand. The 4 sub-factors and 15 items in this factor are shown in Appendix B. Students recognize the importance of instructor's "getting the point across clearly" and "giving directions clearly," (Polk, 2006, p. 25). As an integral part of teaching, verbal ability is a general indication of academic ability; therefore, it is not surprising that the "weakest teachers, on average, have the lowest verbal scores" (Andrew, Cobb & Giampietro, 2005, p. 352). Instructors rated highly on this factor are perceived by students as capable of effectively and clearly communicating student goals and assignments, of demonstrating how work should be completed and of employing a method of teaching that aided in the learning process.

Factor Five was named *Subject Expertise* and included items to rate the student reactions to the instructor's demonstrated subject matter expertise. Each of the 8 items and 3 sub-factors (shown in Appendix B) describe the instructor's use of real world application of subject material to the lesson, their overall subject matter expertise, and their ability to integrate student knowledge into the lesson. While the learning environment is "influenced by the integration of students in the class group" (Opdenakker & Van Dammea, 2006. p. 11), an instructor's well-developed, accurate and current subject-matter content knowledge is a critical element of effective instruction (Freeman, Anderman & Jensen, 2007; Madsen, 2003; Nathan & Petrosino, 2003; Polk, 2006). Instructors rated highly on this factor are perceived by students as capable of demonstrating the importance of the subject matter, the relationship of the subject matter to real-life situations and other subject matter known by the student.

Factor Six was named *Provides Feedback* and included items to rate the student reactions to how well the instructor provided feedback. Each item and sub-factor described the instructor's ability to provide feedback, and the instructor's questioning techniques, in regards to student comprehension. The six sub-factors and 18 items in this factor are shown in Appendix B.

Freeman, Anderman & Jensen (2007) provide a list of feedback characteristics that student's felt were relevant that includes the instructor openness to other viewpoints and ideas, the instructor praising student performance, the promptness that tests and papers were graded and returned, the extent that student's are kept informed concerning their progress, and if students felt safe enough to volunteer their opinions and to even disagree with the instructor (p. 212). Instructors rated highly on this factor are perceived by students as thorough and timely in providing feedback and as competent in questioning by maintaining eye contact and actively listening.

Factor Seven was titled *Concern for Learning* and included items to rate the student reactions to the instructor's level of concern for the student's progress and learning. Each item and sub-factor described the instructor's ability to motivate, encourage and improve the performance and progress of the student. The four sub-factors and nine items in this factor are shown in Appendix B. Students seem to respond to both conscious and unconscious behaviors; therefore, sincere enthusiasm, a passion for teaching and learning, and genuine concern for a student's success can serve as a powerful emotional base capable of influencing student ratings (Benson, Cohen & Buskist, 2005;

Freeman, Anderman & Jensen, 2007; Madsen, 2003, Polk, 2006). Instructors who establish rapport with their classes are likely to have students who engage in pro-academic behaviors by attending class, paying attention and enjoying the subject matter (Benson, Cohen & Buskist, 2005; Freeman, Anderman & Jensen, 2007). Borke & Livingston (1989) found that expert teachers were more capable than new teachers in predicting when students were likely to have problems, and were "more successful in relating their demonstrations to specific learning difficulties of students" (van Driel, Beijaard, & Verloop, 2001, p. 148). Instructors rated highly on this factor are perceived by students as concerned with their progress and learning, as capable of motivating students to learn, of knowing when students don't understand the lesson material, and in providing specific study skills to improve student performance.

Factor Eight was named *Fair and Respectful* and included items to rate the student reactions to the instructor's positive, respectful, and fair attitude toward students. Each sub-factor and item described the instructor's ability to demonstrate a positive attitude toward students in a fair and impartial manner. The three sub-factors and seven items in this factor are shown in Appendix B. Freeman, Anderman and Jensen (2007) reported that a

student's "sense of belonging" was associated with academic motivation particularly when the instructor knew individual student's names, encouraged student participation, and was perceived as enthusiastic and friendly. Instructors rated highly on this factor are perceived by students as being positive, genuinely respectful of the students as individuals, and of being fair in their treatment and evaluation of student performance.

Factor Nine was titled *Available and Relevant* and included items to rate the student reactions to the instructor's availability outside of class and the relevance of the lesson material. Each sub-factor and item described the instructor's availability outside the classroom and the overall value of the lesson as taught by the instructor. The three sub-factors and five items in this factor are shown in Appendix B. Students report effective teacher qualities as approachable and accessible (Benson, Cohen & Buskist, 2005). Instructors rated highly on this factor were perceived by students as available, instrumental in their gaining knowledge and capable of making the lesson valuable.

Data Collection

Coordination with course managers and instructors allowed access to lessons prior to the exam/evaluation phase. There were 68

suitable lessons in session during data collection that met the study's parameters of a single instructor of an entire lesson with subsequent lesson performance exam measuring only lesson material taught therein. Fourteen instructor-led real-time classroom learning events were surveyed from varying vocational courses including: motor transport operations and organizational maintenance, financial management, personal administration, and supply management. Each of the fourteen lessons surveyed had a different instructor for that lesson. Since instructor characteristics were in part the focus of the study, only lessons taught by a single instructor throughout to lesson performance examination were considered suitable to survey students to participate in this study. No attempt was made to align or restrict lesson topic, number of trainees, or instructor's age, tenure or seniority. One researcher collected data for all fourteen lessons and post-lesson test scores.

Survey Administration

The researcher used an approved script (See Appendix C) to solicit voluntary subjects for participation in the study. Those students who declined to participate were requested to leave the classroom during the survey phase in order to prevent potential distractions. All subjects were informed of the study's purpose

and their role as a research participant in accordance with all North Carolina State University Institutional Review Board (IRB) for Use of Human Subjects in Research guidelines and approvals (Appendix D). All participants signed a Participant Informed Consent to Release Form for Research (Appendix E) prior to data collection and were provided a copy upon request. The students' four-digit identification number was used to code the participants for confidentiality.

The survey package was a single page informed consent to release form with a coded survey number (requiring their signature and student four-digit identification number) and a second two-sided survey instrument with the same survey number (Appendix F). By assigning matching survey numbers to both sheets and then immediately separating signed student consent forms from the survey instrument (with demographic information and scaled responses), the investigator was able to protect the identities of participants.

Data from the survey packages and individual lesson scores were transferred onto a master reference spreadsheet. The master reference spreadsheet used the survey number (located on the informed consent form and the survey instrument) to link the student number to the post-lesson test scores - which were

collected once the lesson was complete. Aligned data was then transferred to a working spreadsheet for data analysis showing only survey number, lesson scores, demographics and scaled item responses; therefore, there was no way to connect students to cases.

Statistical Data Analysis

SPSS was used for all statistical operations. Table 3.2 summarizes the statistical analysis employed for each of the six research questions described in the following paragraphs. Means and standard deviations are used to describe participant age (ratio) and post-lesson test scores (interval). Frequency and percents were used to describe gender (dichotomous), highest education level (ordinal), and ethnicity (nominal).

Research Question 1 - *What are the instructor characteristics that can be used to describe instructors and explain student's reaction toward instructor quality and student learning of lesson material?*

A comprehensive literature review of previously published and peer-reviewed instruments was conducted to create an initial instrument. Independent variables were the instructor characteristics identified by items within the instrument.

Instructor quality (the last two opinion-style items, combined) and post-lesson test scores served as the dependent variables.

Table 3.2

Statistical Analysis

Question	Statistical Procedure
1	Factor Analysis
2	Descriptive (Mean, Standard Deviation, Frequency, & Percents)
3	Descriptive (Mean, Standard Deviation, Frequency, & Percents)
4	Correlation Pearson's Product Moment
5	Step-wise Multiple Regression
6	Step-wise Multiple Regression

Exploratory factor analysis was used to identify unique factors and to calculate correlations within all items to reduce the overall number of items. Factor pattern matrixes were used to report the factor loadings of each measured item with each measured factor, when the other factors are not included.

Cronbach's alpha was used to assess the internal consistency of all scaled factors. As factors contained differing numbers of items, averages were calculated and presented for comparison.

Post-lesson test scores were initially reported by frequency and percent but were averaged for step-wise multiple regression analysis.

Research Question 2 - *How do students perceive instructor characteristics as measured by the developed instrument?*

Means and standard deviations, frequency and percent were used to report perceived instructor characteristics identified by the instrument's original nine factors discussed previously. Each factor had a different number of items; therefore, averages were calculated by first summing - then dividing by the number of items within that factor. Mean, standard deviation, frequency and percent were used to report the averages.

Research Question 3 - *What are the student reactions to perceived instructor quality as measured by the instrument and student learning as reported by post-lesson test scores?*

Means, standard deviations, frequency and percent were used to report perceived instructor quality (Items 45 and 46, combined, in the revised instrument) and post-lesson test scores.

Research Question 4 - *Are there relationships between the instructor characteristics and student's reaction to perceived*

instructor quality as measured by the instrument and student learning as reported in post-lesson test scores?

Pearson's Product Moment correlations were used to identify relationships between instructor characteristics, instructor quality and post-lesson test scores. Instructor characteristics, instructor quality and post-lesson test scores were coded as interval variables.

Research Question 5 - *Do instructor characteristics explain a significant amount of variance in perceived instructor quality?*

Step-wise multiple regression analysis was used to explain student's perceptions of instructor quality. The step-wise regression was initially set to .05 as the probability of F for entry and .10 as the probability of F for removal; although, a forced entry model was used. The independent variables were the instructor characteristics identified by the three factors (*Effective Delivery, Negativity, and Concern for Learning*) while *Instructor Quality* (Item 45 and 46 combined) served as the dependent variable. All variables were placed into the model by the three factors.

Research Question 6 - *Do instructor characteristics explain a significant amount of variance in post-lesson test scores?*

Step-wise multiple regression analysis was used to explore if perceived instructor characteristics explain post-lesson test scores. The step-wise regression was set to .05 as the probability of F for entry and .10 as the probability of F for removal. The independent variables were the instructor characteristics identified by three factors (*Effective Delivery*, *Negativity*, and *Concern for Learning*) while post-lesson test scores served as the dependent variables. Variables were placed into the model by the amount of variance the variable explained.

First Round Field Test

The initial instrument was administered to 314 first round participants. None of the first round participants requested their data not be used in the study.

Demographics

The study collected demographic data from participants regarding their gender, age, highest educational level attained, ethnicity, and state where the participant enlisted; although, the state where enlisted is not reported in this study (but may be considered for future studies).

All of the first round participants ($n = 306$) reported gender, education level and ethnicity and all but one participant

reported age (.3%). Appendix G reports the frequencies and percents of the first round participant demographics.

Male respondents accounted for 93.1 percent ($n = 285$) of the participants while female respondents accounted for slightly less than seven percent ($n = 21$; 6.9%). The age of respondents ranged from 18 to 30 years of age with an average age of 20.93 years ($SD = 2.147$). Three ages (19 years, $n = 89$, 29.1%; 20 years, $n = 84$, 27.5%; and 21 years $n = 42$, 13.7%) were reported most frequently and accounted for 70.5 percent of participants.

All first round respondents reported obtaining a General Equivalency Degree (GED) ($n = 21$, 6.9%) or high school diploma ($n = 264$, 86.3%) as reflected in Table 4.1. Twenty-one participants (6.9%) reported completing two- ($n = 17$, 5.6%) or four-years ($n = 4$, 1.3%) of college.

The majority of participants selected White ($n = 168$, 54.9%) as their ethnic group followed by Hispanic ($n = 55$, 18.0%), Black ($n = 29$, 9.5%), "other" ($n = 24$, 7.8%), Pacific/Hawaiian Islander ($n = 19$, 6.2%), Asian ($n = 7$, 2.3%) and American Indian/Alaskan Native ($n = 4$, 1.3%). Participants were provided the opportunity to either select more than one ethnic category or to write-in their ethnic preference under "Other".

Addressing Missing Responses

Not all participants responded to all statements. The missing responses were examined. In all cases, missing data was replaced with the statistical mean. In those cases where the number of missing responses exceeded three standard deviations from the average number of cases missing those case responses were considered statistical outliers and thus were not used in the data set analysis.

Identification of Case Outliers

Participant responses were surveyed during two rounds of data collection. This section will discuss participant responses and non-responses by data collection round.

The first round field test collected participant responses using the initial 98-item instrument. Means and standard deviations were collected to determine case outliers from the initial 314 cases. During the initial stage, any case with three or more non-responses ($M = .3981$, $SD = .9206$ (3 standard deviations from the mean is = 3.159) was discarded.

During the validation stage, any case exceeding two non-responses ($M = .3052$, $SD = .6129$, 3 standard deviations from the mean is = 2.144) was discarded. Eight cases were discarded due

to the number of non-responses exceeding three standard deviations from the norm during both stages. Non-response cases identified in the first round are reported in Table 3.3.

Table 3.3

Cases Discarded as Outliers Due to Non-Response

Case Number	Number of Non-Response (Missing)	
	<i>f</i>	%
Stage ^a		
A15	7	.309
A46	7	.309
E3	5	.309
A36	4	.309
A23	4	.309
C58	4	.309
Stage ^b		
D17	3	.309
D31	3	.309

Note. ^aStage, $M = .3981$, $SD = .9206$ (3 standard deviations from the mean is = 3.159). ^bStage, $M = .3052$, $SD = .6129$ (3 standard deviations from the mean is = 2.144).

First round data was then transposed to analyze the non-response frequency of the 98 items. During the initial stage,

any item with 8 or more non-responses ($M = 1.816$, $SD = 2.335$, 3 standard deviations from the mean is = 8.820) was discarded. During the validation stage, any item exceeding 7 non-responses ($M = 1.604$, $SD = 1.803$ (3 standard deviations from the mean is = 7.014) was discarded. Items 94, 95 and 31 were discarded as the total number of non-responses for that item exceeded three standard deviations from the mean. Non-Responses items are reported in Table 3.4.

Table 3.4

Items Discarded as Outliers Due to Non-Response

Item Number	Number of Non-Response (Missing)	
	<i>f</i>	<i>%</i>
Stage ^a		
Item 94	10	1.0
Item 95	14	1.0
Stage ^b		
Item 31	8	1.0

Note. ^aStage, $M = 1.816$, $SD = 2.335$ (3 standard deviations from the mean is = 8.820). ^bStage, $M = 1.604$, $SD = 1.803$ (3 standard deviations from the mean is = 7.014).

Exploratory Factor Analysis

Factor analysis using a non-rotated, principle axis factoring, based on initial eigenvalues greater than one, calculated 21 components, accounting for a cumulative variance of 66.57% that was used to exclude items that correlated poorly from further analysis. Principle axis factoring, also called common factor analysis, was used in this study "because it is more appropriate than principal components analysis when the objective is to identify latent structures, rather than for pure prediction" (Holton, Bates & Ruona, 2000, p. 342). Factor analysis loadings were considered significant at .325 ($p < .05$) for a sample size of 300 participants (Hair et al., 2006, p. 128).

Table 3.5 reflects the exploratory factor analysis of the initial instrument. Seven items were discarded due to poor factor loading: Item 16 (-.312), Item 23 (.303), Item 11 (.299), Item 49 (.203), Item 13 (.235), Item 15, (.208), and Item 12 (.102). No factors exceed .325 beyond factor eight; therefore, only those factors are shown.

A non-orthogonal oblique rotation, (oblimin rotation in SPSS), was employed in hopes of gaining more realistic findings, presuming all factors have "some degree of correlations" which "generally result in superior simple and interpretable solutions"

(Yang, 2005, p. 193). Oblique rotations produce both factor pattern and factor structure matrices. While a factor structure matrix represents correlations between variables, this study employed a factor pattern matrix to display a more linear combination of the variables. Table 3.6 reflects the initial oblimin factor analysis using an oblimin rotation, primary axis factoring with an eigenvalue of one.

Table 3.5

Exploratory Factor Analysis of Initial Instrument

Item	Factor							
	1	2	3	4	5	6	7	8
45	.779							
87	.768							
81	.737							
67	.729							
58	.721							
36	.712							
83	.707							
41	.703							
82	.700							
42	.700							
46	.700							
44	.695							
89	.693							
43	.685							
30	.684						.334	
68	.682							
73	.680							
52	.667							
57	.665							
32	.662							.330
26	.657							
69	.655							
34	.650							
29	.650							
47	.642							
84	-.639							
51	.636							
37	.632							
3	.612							
79	.612							
53	.610							
24	.608				-.336			
17	.608							
92	.607							
55	.603							
28	.597							
54	.589							
5	-.582							
80	-.576							
27	.572							
62	.571							
35	.570							
70	-.568	.413						
75	.566							

Table 3.5

Exploratory Factor Analysis of Initial Instrument (Continued)

Item	Factor							
	1	2	3	4	5	6	7	8
6	.561							
19	.559							
40	.558							
78	.558							
66	-.546	.475						
2	.545		-.387					
76	.539		.352					
61	-.526							
8	.520							
10	.518					.345		
90	-.514	.416						
25	.513				-.351			
71	-.511	.452						
1	.511		-.416					
50	-.510							
20	.503							
7	.502							
21	.502							
18	-.496							
64	.488							
96	.482							
48	-.481							
33	-.470							
74	-.470	.431						
60	.452							
93	-.451	.341						
59	.428							
65	.426							
38	-.421							
77	.416							
3	-.413							
85	.411							
88	-.411	.391						
72	.400							
14	-.370	.330						
22	.367							
39	-.359							
16	-.312							
23	.303							
11	.299							
49	.203							
91	-.476	.491						
63		.370	.351					
13				.235				
15				.208				
9	.390					-.407		
12						.102		

Table 3.6

Exploratory Factor Analysis Oblimin Rotation

Item	Factor																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
40	.598																		
41	.569																		
42	.402																		
52	.339																		
28	.288																		
44	.277																		
43	.275																		
45	.193																		
88		.553																	
66		.544																	
80		.506																	
90		.497																	
84		.453																	
74		.396																	
71		.367																	
93		.339																	
61		.268																	
2			-.735																
1			-.684																
4			-.562																
81			-.424																
5			.349																
46			-.347																
47			-.292																
22			-.285																
86				.484															
85				.449															
77				.394															
79				.384															
78				.326															
59				.316															
6				-.224															
24					-.624														
8					-.385														
96					-.337														
62					-.334														
83					-.245														
3					.245														.200
33						-.682													
32						.451													
14						-.269													
54																			
57																			
35																			

Table 3.6

Exploratory Factor Analysis Oblimin Rotation (Continued)

Item	Factor																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
58							-.358												
51							-.357												
20								-.593											
21								-.512											
38									.476										
39									.427										
48									.366										
53							-.308		-.314										
91										.549									
50										.494									
70										.330									
69										-.329									
36											-.500								
37											-.475								
27											-.405								
30	.293										-.302								
87											-.287								
60											-.252								
26											-.234								
18												-.768							
17												.659							
9													-.766						
10													-.477						
29					.249								-.304						
76														.650					
75														.407					
67														.333					
73											-.235			.260					
89															.411				
68															.358				
92															.351				
65																.558			
63																.548			
64																.232			
72																		-.574	
56																		-.311	
34										-.260								-.279	
25					-.232													-.259	
55							-.227											-.231	
82																			.414
19								-.277											-.285
7				-.228															.251

Factor loadings in the table were sorted by size with the largest values on the top. Twenty-three items correlated poorly and were therefore discarded. They were: Item 3 (.245), Item 5 (.310), Item 6 (-.224), Item 14 (-.269), Item 22 (-.285), Item 26 (-.234), Item 28 (.288), Item 36 (.280), Item 44 (.277), Item 45 (.193), Item 47 (-.292), Item 50 (.325), Item 52 (.288), Item 56 (-.311), Item 59 (.316), Item 60 (-.252), Item 62 (-.287), Item 64 (.232), Item 68 (.260), Item 72 (-.312), Item 78 (.326), Item 87 (-.184), and Item 92 (-.308). Appendix H shows the statement each item represents as they appear in the instrument.

In an attempt to optimize the number of factors to be extracted establishing a "cutoff is most reliable when the number of variables is between 20 and 50 items" (Hair et al, p. 120); therefore, extraction was reduced to eight to match the scree plot of eight factors (Figure 3.1). A scree test plots the eigenvalues on a curve that links the largest to the smallest with the scree being the point where the curve turns upward from the horizontal. Those "factors above the scree are considered to be real factors. All residual factors below the scree are viewed as error factors" (Yang, 2005, p. 191).

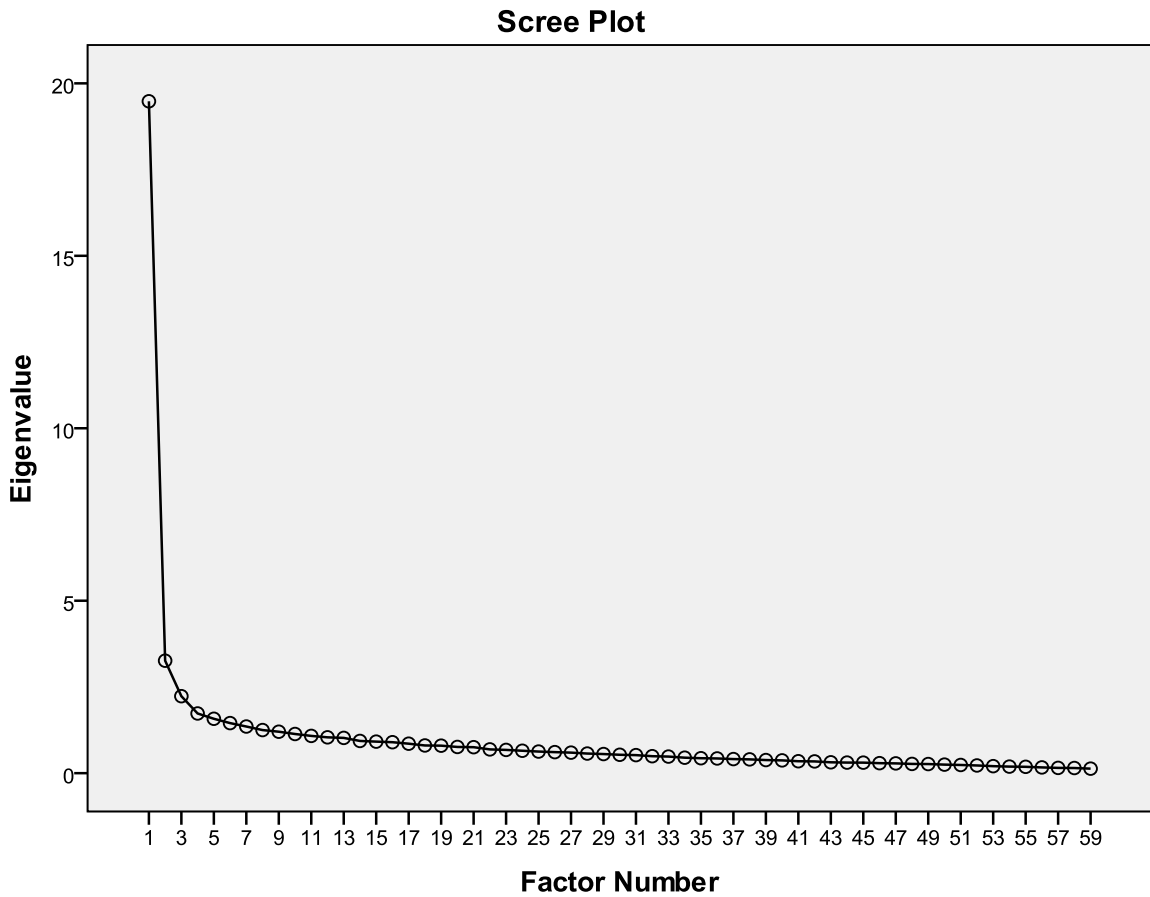


Figure 3.1.

Un-Rotated Scree Plot with Eigenvalue of One

The majority of the items had substantial factor loading structure coefficients within the eight fixed factors, shown in Table 3.7, accounted for 56.45% of the variance.

Seven items were discarded as poorly correlating within the eight fixed factors: Item 37 (.308), Item 38 (.281), Item 35 (-.235), Item 48 (.264), Item 57 (.313), Item 58 (-.225), and Item

67 (-.313). Table 3.8 demonstrates the pattern factor matrix reduced to eight fixed factors, principal axis factoring, and oblimin rotation. Appendix H shows the statement each item represents as they appear in the instrument.

Table 3.7

Factor Loading Coefficients within Eight Fixed Factors

Factor	Eigenvalue	Explained Variance
Factor 1	16.645	32.638%
Factor 2	3.064	6.009%
Factor 3	2.083	4.084%
Factor 4	1.600	3.686%
Factor 5	1.525	3.136%
Factor 6	1.403	2.751%
Factor 7	1.298	2.546%
Factor 8	1.170	2.294%

Table 3.8

Exploratory Factor Analysis with Eight Fixed Factors

Item	Factor							
	1	2	3	4	5	6	7	8
76	.573							
27	.468							
75	.451							
54	.429							
51	.428							
57	.313							
69	.311							-.246
37	.308							
91		.709						
90		.676						
66		.659						
88		.611						
84		.567						
93		.542						
80		.539						
70		.539						
74		.529						
71		.517						
96		-.358						
1			-.744					
2			-.678					
4			-.551					
81			-.526					
46			-.410					
35			-.235					
86				.636				
79				.548				
85				.516				
77				.471				
82				.420				
89				.373				
39					-.572			
33					-.521			
10					.466			
32					.439			
9					.399			
63						.545		
65						.459		
38						.382		
48						.264		
42							-.499	
8							-.498	
40							-.452	
21							-.420	
20							-.409	
41							-.401	
24							-.376	
58							-.225	
18								.706
17								-.609
67								-.313

Data was further factored to five-fixed extractions in an attempt to again reduce the number of instrument variables in line with the scree plot shown in Figure 3.2.

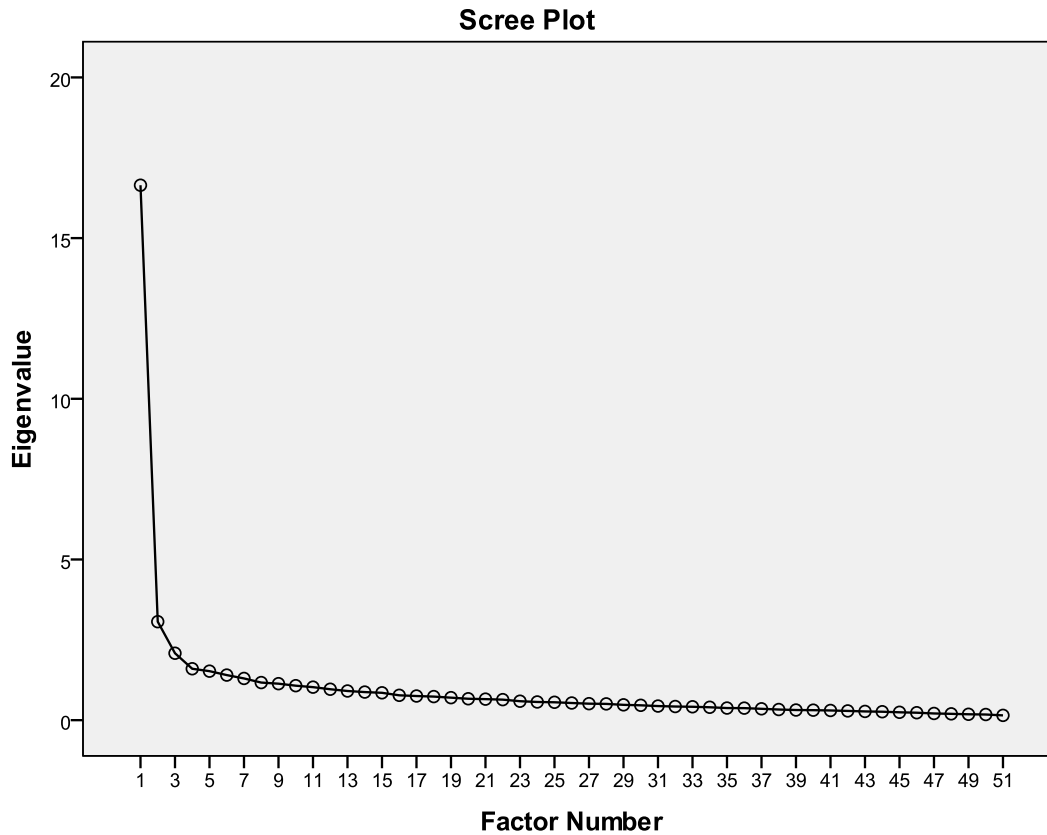


Figure 3.2.

Oblimin Rotation Eigenvalue Scree Plot with Five Fixed Factors

Factor loadings in the five fixed factor table were sorted by size with the largest values on the top. The majority of the items had substantial factor loading coefficients while the first

five factors, shown in Table 3.9, accounted for 50.033% of the variance.

Table 3.9

Factor Loading Coefficients within Five Fixed Factors

Factor	Eigenvalue	Explained Variance
Factor 1	13.647	31.737%
Factor 2	2.983	6.937%
Factor 3	1.861	4.327%
Factor 4	1.578	3.670%
Factor 5	1.446	3.363%

Eight items were discarded due to poor correlation and subsequent recalculations: Item 17 (-.212), Item 18 (-.299), Item 27 (.291), Item 63 (.323), Item 65 (.282), Item 75 (.321), Item 76 (.310), and Item 96 (.302). Table 3.10 displays the exploratory factor analysis with five fixed factors. Appendix H shows the statement each item represents as they appear in the instrument.

Table 3.10

Exploratory Factor Analysis with Five Fixed Factors

Item	Factor				
	1	2	3	4	5
41	.656				
42	.634				
2	.567		-.387		
40	.489				
4	.472				
1	.442		-.406		
21	.436				
46	.426			.384	
20	.394				
54	.377				
51	.359				
75	.346				
17	.344				
18	-.299				
27	.291	-.272	.272		
66		.744			
71		.720			
91		.689			
90		.639			
70		.619			
74		.581			
80		.555			
88		.553			
93		.524			
84		.522			
63			.414		
76			.340		
65			.282		
86				.665	
85				.606	
79				.584	
82				.524	
77				.512	
81	.363			.455	
89				.398	
96				.302	
10					-.595
9					-.555
33					.509
24					-.507
39					.492
32					-.475
8					-.352

Data was further factored to three extractions to again reduce the number of instrument variables. Figure 3.3 reports the three fixed factor scree plot validating the extraction to three factors. Extraction was then reduced to match the scree plot of three factors.

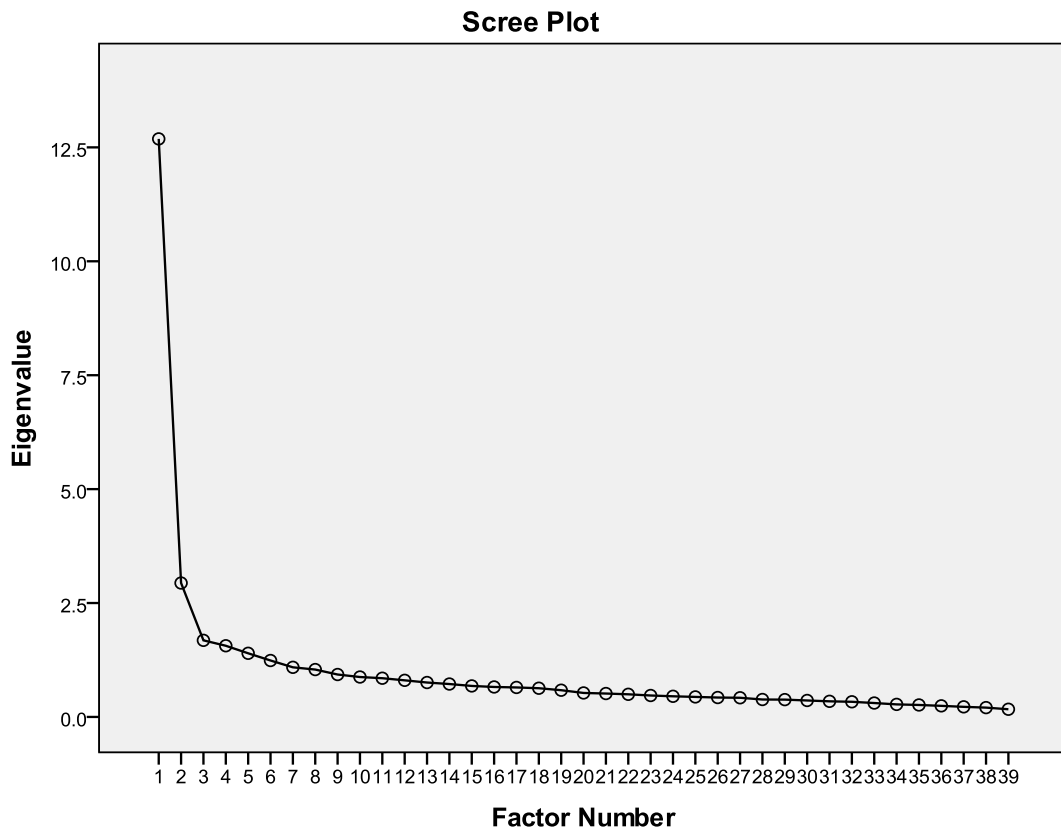


Figure 3.3.

Oblimin Rotation Eigenvalue Scree Plot with Three Fixed Factors

The majority of the items had substantial factor loading structure coefficients while the first three factors accounted for 46.702% of the variance, as shown in Table 3.11.

Table 3.11

Factor Loading Coefficients within Three Fixed Factors

Factor	Eigenvalue	Explained Variance
Factor 1	11.493	33.802%
Factor 2	2.830	8.324%
Factor 3	1.556	4.576%

There were two items discarded due to poor correlation; Item 89 (-.309) was discarded due to initial poor correlation and Item 39 (-.299) was discarded upon subsequent recalculation. Table 3.12 shows the extracted factors.

Upon face validity review, items that correlated poorly but also presented high-cross loadings in multiple factors were reentered into the final three fixed factors. They were: Item 7 (.251, -.228), Item 8 (.225, -.313), Item 19 (-.277, -.285), Item 25 (.232, -.259), Item 29 (-.249, -.304), Item 30 (.293, -.302), Item 34 (-.260, -.279), Item 43 (.275, -.190), Item 53 (-.308, -

.314), Item 55 (-.227, -.231), Item 61 (.268, .207), Item 69 (.311, -.246), Item 73 (-.235, .260), and Item 83 (-.245, -.198).

Table 3.12

Extracted Factors Using Oblimin Rotation and Three Fixed Factors

Item	Factor		
	1	2	3
82	.679		
81	.645		
85	.586		
46	.585		
79	.573		
86	.572		
77	.480		
2	.479		
89	.462		
1	.443		
66		.782	
91		.725	
71		.697	
90		.677	
70		.640	
80		.628	
74		.610	
88		.596	
84		.576	
93		.515	
10			.569
24			.550
41			.542
32			.526
51			.519
42			.517
9			.504
33			-.440
54			.422
20			.411
40			.391
4			.391
21			.386
39			-.299

There were four items discarded due to item similarities between other previously retained items identified during face validity analysis conducted by the researcher. Those items were: Item 2 (The instructor maintained student attention as the lesson progressed) was discarded as too similar to Item 33 (Was not interested in training students); Item 4 (The instructor introduced stimulating ideas about the subject material) was discarded as it matched Item 53 (Related lesson material to real life situations); Item 21 (The instructor used computerized electronic media that enhanced the learning experience) was discarded due to similarity with Item 20 (The instructor used technology comfortably while teaching); and Item 46 (The instructor had a method of teaching that helped me understand the lesson material) was discarded in lieu of Item 55 (The Instructor related content of this lesson to other subject matter that students already understand). The final instrument contains 46-items and is presented in Appendix H.

Revised Factor Naming

Factors were reassessed based on resultant responses and analysis. Three factors became evident and are named below:

Factor 1 - *Effective Delivery*: The student perceived the instructor clearly and effectively delivered the lesson material.

This factor was made up of six sub-factors and twenty-five items. All items are related to specific instructor characteristics that students react to in normal training situations. Polk (2006) stresses that "delivery must be effective or else the information and its quality cannot be consumed" (p. 24). Therefore, the instructor characteristics found in this factor include effective use of time, clear communication of lesson objectives, relevant use of examples and expertise, summarizing and signaling transitions throughout the lesson, and being attentive to, and enthusiastic with, the class. Additionally, instructor's management of the learning environment, training equipment and technology were key elements found in this factor. Instructors rated highly on this factor are perceived by students as employing successful training techniques, being responsive and dutiful to the student's learning needs, effectively communicating lesson goals and managing the learning situation. The six sub-factors and twenty-five items in this factor are shown in Appendix I.

Factor 2 - *Negativity*: The student perceived the instructor as disinterested in, and uncooperative toward, student's learning. There were three sub-factors and eleven items within this factor related to negative instructor characteristics.

Students described the "worst" instructors as having unfair evaluation or grading procedures (Brown & Tomlin, 1996) while conversely "best" instructor qualities included fairness (Benson, Cohen & Buskist, 2005). Therefore, instructors rated highly did not review properly or evaluate student work fairly. Students also rated instructors highly in this factor when they appeared to lack interest, failed to inspire student performance or were perceived as discourteous and uncaring. They also became angry when questions were asked, answered questions in incomplete or inaccurate ways, interrupted and criticized student responses and openly argued with students. The three sub-factors and eleven items in this factor are shown in Appendix I.

Factor 3 - *Concern for Learning*: The student perceived the instructor appeared to be concerned about the student's progress and motivation to learn. There were two sub-factors and eight items within this factor. Students rated instructors highly when they felt that the instructor is genuinely interested in them and even liked them (Benson, Cohen & Buskist, 2005). Instructor characteristics rated low in this factor reflected an inability to motivate or encourage student performance coupled with a lack of concern for student achievement and understanding. Instructors rated highly in this factor presented the lesson in

an interesting way, referred to student's by name, knew when students didn't comprehend the material, and was able to provide specific study skills to improve individual performance. The two sub-factors and eight items that make up this factor are shown in Appendix I.

Reliability of Instrument

Cronbach's alpha was used to determine reliability of each summated factor. Summated scores are multivariate measurements where joined individual variables measure a composite construct to gain a multi-faceted perspective. Cronbach's alpha uses internal consistency to determine an estimate of reliability. Internal consistency is calculated for summated scales by calculating the coefficient alpha for indication of strong covariance between variables. Spector (1992) states "Internal-consistency reliability is an indicator of how well the individual items of a scale reflect a common, underlying construct" (p. 65). Hair et al., (2006), suggests a normative coefficient alpha of .70 but states it maybe be lower, such as .60 for exploratory research. To improve reliability, it is suggested that if there is a sufficient number of retained items, the researcher can discard those items that - if removed - will improve the overall scale reliability (Hinkin, 2005, p. 173).

Table 3.13 shows the estimated reliability using Cronbach's alpha of the instrument's three factors. Estimated alphas ranged from .846 to .919. Reliability for factor 1 ($\alpha = .92$), factor 2 ($\alpha = .89$), and factor 3 ($\alpha = .85$) indicate that all scales are above the minimum for reliability, as show in Table 3.13.

Table 3.13

Reliability of Three Factor Instrument

Factor Item	Factor 1	Factor 2	Factor 3	Reliability Alpha α
Factor 1				.92
Item 7	.522			
Item 8	.523			
Item 9	.430			
Item 10	.518			
Item 19	.541			
Item 20	.487			
Item 24	.627			
Item 25	.525			
Item 27	.555			
Item 29	.647			
Item 30	.687			
Item 32	.595			
Item 33	-.423			
Item 34	.612			
Item 40	.558			
Item 41	.684			
Item 42	.706			
Item 43	.674			
Item 51	.603			
Item 53	.610			
Item 54	.572			
Item 55	.605			
Item 69	.631			
Item 73	.648			
Item 76	.519			
Factor 2				.89
Item 61		.517		
Item 66		.699		
Item 70		.666		
Item 71		.649		
Item 74		.617		
Item 80		.644		
Item 84		.676		
Item 88		.565		
Item 90		.662		
Item 91		.659		
Item 93		.557		
Factor 3				.85
Item 1			.469	
Item 82			.734	
Item 81			.713	
Item 85			.512	
Item 79			.607	
Item 86			.575	
Item 77			.464	
Item 83			.672	

Summary of Methodology

The research design and methodology was determined appropriate to identify and validate student's perception of instructor characteristics and associated military instructor quality, purposively selected from current entry-level U.S. Marines attending military-orientated vocational training. The steps taken to conduct this study are as follows:

1. A comprehensive literature review was conducted to identify existing student rating instruments already in use from a broad range of fields and academic settings.
2. Variables (items) were selected based on terminology appropriate to teacher evaluations, instructional effectiveness, attributes and techniques, and evaluation.
3. Items were filtered for appropriateness and content validity and clustered into themes, factors and sub-factors based on similarity.
4. Rating scales were designed based on current practices identified in current educational literature.
5. A survey instrument was created using the items and rating scales which included individual demographics.

6. A North Carolina State University IRB application was completed and submitted to the IRB committee for approval, and approval was subsequently given to conduct the study.
7. Researcher met with school officials to gain access to participant classrooms.
8. A purposive sample was identified from courses currently in session at the time of data collection. Sample size was established for two rounds of data collection.
9. The researcher met with available courses in session. No pattern or designation was instituted to restrict participation.
10. A master reference sheet (Microsoft Excel spreadsheet) was used to compile participant responses and demographics.
11. The first round of data were collected $n = 314$. Eight cases were later discarded as outliers resulting in $n = 306$ valid participants (97.4%).
12. Initial data analysis was conducted employing exploratory factor analysis and examination of content validity to reduce variables and revise the instrument. Data were analyzed using *Statistical Package for the Social Sciences (SPSS), version 18.0*.
13. The second round of data collected ($n = 139$) using the same steps as round one. Three cases (3.0%) were discarded as

outliers resulting in 136 valid participants using the revised instrument.

14. Both rounds had a total of 442 (97.6%) useable participants.

15. Final round data were analyzed using SPSS 18.0 and appropriate statistical analysis was conducted to answer the research questions.

CHAPTER 4

Analysis of Data

This chapter provides the analysis of collected data concerning the instrument's ability to assess instructor characteristics based on student reactions. This chapter reports the findings regarding the study's research questions.

The chapter is divided into sections, a section reporting participant demographics, six sections reporting the findings concerning the research questions, and a summary to discuss overall data findings. Data were also collected concerning participant's post-lesson test scores. Data were analyzed using statistical procedures provided by *Statistical Package for the Social Sciences (SPSS)*, version 18.0.

Demographics

The study collected demographic data from participants regarding their gender, age, highest educational level attained, ethnicity, and state where the participant enlisted; although, the state where enlisted is not reported in this study (but may be considered for future studies). This section reports participant demographics for the study.

All of the participants ($n = 136$), reported age, gender, education level and ethnicity. Table 4.1 reports the frequencies and percents for gender, age, education level attained, and ethnicity of participants.

Table 4.1

Frequencies and Percents of Age, Gender, Education Level, and Ethnicity of Participants

	<i>N</i>	<i>n</i> 136	% 100
Gender			
Male		122	89.7
Female		14	10.3
Age			
19 yrs		30	22.1
20 yrs		35	25.7
21 yrs		22	16.2
22 yrs		12	8.8
23 yrs		10	7.4
24 yrs		7	5.1
25 yrs		6	4.4
26 yrs		9	6.6
27 yrs		2	1.5
28 yrs		2	1.5
30 yrs		1	.7
Highest Education Level			
GED		3	2.2
High School Diploma		120	88.2
2yr College		9	6.6
4yr College		4	2.9
Ethnicity			
Asian		1	.7
Black		5	3.7
Hispanic		20	14.7
Pacific/Hawaiian Islander		10	7.4
White		82	60.3
Other		18	13.2

Male respondents accounted for nearly 90 percent ($n = 122$, 89.7%) of the participants while female respondents accounted for slightly more than ten percent ($n = 14$; 10.3%). The age of respondents ranged from 19 to 30 years of age with an average age of 21.46 years ($SD = 2.455$). Three ages (20 years, $n = 35$, 25.7%; 19 years, $n = 30$, 22.1%; and 21 years $n = 22$, 16.2%) were reported most frequently and accounted for 64% of the participants.

All respondents reported obtaining a high school diploma ($n = 120$, 88.2%) or General Equivalency Degree (GED) ($n = 3$, 2.2%). Additionally, in addition to completing high school or equivalency, nine participants (6.6%) reported completing two years of college and four participants (2.9%) reported completing four years of college.

The majority of participants selected White ($n = 82$, 60.3%) as their ethnic group followed by Hispanic ($n = 20$, 14.7%), "other" ($n = 18$, 13.2%), Pacific/Hawaiian Islander ($n = 10$, 7.4%), and Black ($n = 5$, 3.71%). One participant (.7%) identified themselves as Asian. As with the previous round, participants were provided the opportunity to either select more

than one ethnic category or to write-in their ethnic preference under "Other".

Summary of Participant Demographics

The average participant in the study was white ($n = 250$, 56.4%), male ($n = 408$, 92.1%), with a high school diploma ($n = 385$, 86.9%) and an average age of 21.10 years ($SD = 2.252$).

Research Questions

This section will discuss the research questions, data analysis and finding relevant to the study. The analysis of the data was focused on the following research questions:

Research Question 1 - *What are the instructor characteristics that can be used to describe instructors and explain student's reaction toward instructor quality and student learning of lesson material?*

Research Question 2 - *How do students perceive instructor characteristics as measured by the developed instrument?*

Research Question 3 - *What are the student reactions to perceived instructor quality as measured by the instrument and student learning as reported by post-lesson test scores?*

Research Question 4 - *Are there relationships between the instructor characteristics and student's reaction to perceived*

instructor quality as measured by the instrument and student learning as reported in post-lesson test scores?

Research Question 5 - *Do instructor characteristics explain a significant amount of variance in perceived instructor quality?*

Research Question 6 - *Do instructor characteristics explain a significant amount of variance in post-lesson test scores?*

Data were collected using the instrument, as detailed in Chapter 3, and from post-lesson test scores. The instrument collected demographic data as indicated previously. The revised instrument contained 46 items distributed into three factors: Effective Delivery, Negativity, and Concern for Learning. A six-point Likert scale was used to measure participant agreement to item statements, with 1 (strongly disagree), 2 (moderately disagree), 3 (slightly disagree), 4 (slightly agree), 5 (moderately agree), and 6 (strongly agree).

Item 45 was reverse-scaled (1=6, 2=5, 3=4, 4=3, 5=2, and 6=1) and reverse worded as: 1 (one of the best), 2 (very good), 3 (above average), 4 (below average), 5 (very bad), and 6 (one of the worst) while Item 46 was reverse scaled but retained the same verbiage, with 1 (strongly agree), 2 (moderately agree), 3 (slightly disagree), 4 (slightly agree), 5 (moderately disagree),

and 6 (strongly disagree). Both items were transformed prior to data analysis.

Using a nonrandom purposive sampling technique, the researcher coordinated survey times with available courses in session at the time of data gathering. Fourteen sessions were conducted with 459 participants; although 17 cases (3.7%) (participant responses) were discarded as non-response outliers. No participant requested their responses to be removed from the study.

Research Question 1 - *What are the instructor characteristics that can be used to describe instructors and explain student's reaction toward instructor quality and student learning of lesson material?*

This section will present the findings and analysis of the data for research question one. This question was focused on instructor characteristics (as factors) and the student's perception of instructor quality and post-lesson scores. A comprehensive literature review of previously published and peer-reviewed instruments was conducted to create an initial instrument. Chapter 3 discussed the exploratory factor analysis that was used to identify the three factors based on the strength of the factor loadings and Cronbach's alpha reliability.

Additional analysis was required to validate the revised instrument.

Revised Instrument Analysis

Using a nonrandom purposive sampling technique, the revised 46-item survey instrument was administered to 139 participants to collect rated item responses and demographics. None of the 139 participants requested their data not be used in the study.

The instrument utilized a 6-point Likert scale ranging from 1 to 6 to rate participant reactions with 1 = strongly disagree, 2 = moderately disagree, 3 = slightly disagree, 4 = slightly agree, 5 = moderately agree, and 6 = strongly agree on 44 items related to perceived instructor characteristics. Two items in the revised instrument were reverse-scaled to reduce the potential likelihood of response set or acquiescence bias. Item 45 (Compared with all instructors I have had, both in high school and since, this instructor was: 1 = one of the best, 2 = very good, 3 = above average, 4 = below average, 5 = very bad, and 6 = one of the worst) required the participants to select an alternate measurable scale related to instructor quality. Item 46 (Compared with all instructors I have had, both in high school and since, this instructor was not qualified to be an instructor) was also reverse scaled as (6 = 1, 5 = 2, 4 = 3, 3 = 4, 2 = 5, 1

= 6). All reverse scaled items were transformed to match all other responses prior to data analysis.

Revised Instrument Item Outliers

Three cases were discarded from the 139 participants surveyed due to the number of non-responses exceeding three standard deviations from the norm. Table 4.2 reports the non-response cases.

Table 4.2

Revised Instrument Cases Discarded Due to Non-Response

Case Number	Number of Non-Response (Missing)	
	<i>f</i>	%
M42	3	.719
C14	2	.719
A52	2	.719

Note. $M = .1942$, $SD = .4797$ (3 standard deviations from the mean is = 1.633)

There were no items discarded as the number of non-response items fell within three standard deviations of the mean ($M = .8696$, $SD = 1.24023$ (3 standard deviations from the mean is 6.6088)).

Exploratory Factor Analysis

Factor analysis was conducted on the revised instrument to reduce the large number of variables into a smaller number of variables. Initial factor analysis using principle axis factoring extraction and varimax rotation, based on an eigenvalue greater than one, calculated eight components (factors) and was used to exclude items that correlated poorly from further analysis. Principle axis factoring was employed as "more appropriate than principal components analysis when the objective is to identify latent structures" (Holton, Bates & Ruona, 2000, p. 342).

With an orthogonal rotation, such as the varimax, the factors are not permitted to be correlated (they are orthogonal to one another). Varimax was used as the rotation method in an attempt to "maximize the variance of squared loadings on a factor in order to produce some high loadings and some low loadings for each factor" (Yang, 2005, p. 192). For varimax, the factor structure and the factor pattern matrices are the same.

Factor analysis loadings were considered significant at .425 ($p < .05$) for a sample size of 175 participants (Hair et al., 2006, p. 128). Table 4.3 demonstrates the initial varimax rotated factor matrix with an eigenvalue of one. Factor loadings

in the table were sorted by size with the largest values on the top.

Factor analysis using principle axis factoring and varimax rotation, based on an eigenvalue of one, revealed four items that correlated poorly: Item 22 (.427), Item 11 (.425), Item 38(.359), Item 19 (.386), & Item 26 (-.339). A majority of the items had substantial factor loading structure coefficients in the first factor. Subsequent factors gradually decreased proportions of the remaining variance; although, all items possessed substantial high cross-loadings; therefore, extraction was reduced to match the scree plot of three factors shown in Figure 4.1.

Table 4.3

Varimax Rotated Factor Matrix for Instructor Characteristics

Item	Factor							
	1	2	3	4	5	6	7	8
30	.854							
29	.850							
35	.842							
33	.781							
36	.762							
34	.761							
31	.726	-.453						
28	.726							
32	.649							
27	.629							
13	.490				-.426			
25	-.486							
24	-.432							
41		.771						
43		.703						
40		.680						
39		.554						
37		.506						
44		.504						
22		.427						
18			.615					
16			.589					
15			.589					
23			.487					
17			.471					
10			.466					
11			.425					
38			.359					
21				.707				
20				.667				
8				.544		.452		
7				.497				
19				.386				
12					.692			
9					.569			
14					.553			
42		.485			.496			
3						.746		
4						.564		
6							.763	
5					.440		.447	
26							-.339	
2								.611
1								.452

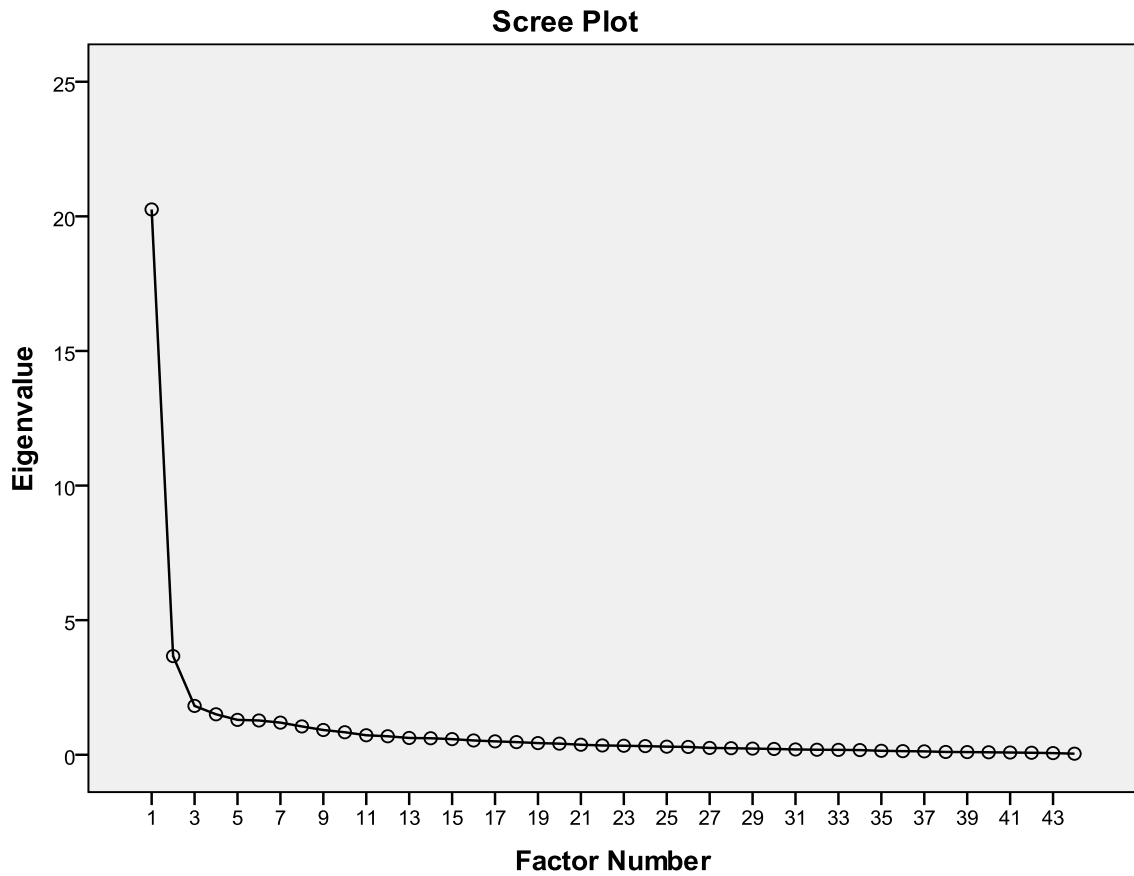


Figure 4.1.

Varimax Rotation Scree Plot for Instructor Characteristics

Factor analysis using principle axis factoring and varimax rotation, based on a fixed number of three factors revealed three items that correlated poorly: Item 4 (.429), Item 2 (.411) and Item 1 (.328). The data were recalculated using a varimax rotation and principle axis factoring with three fixed factors.

All items loaded above .425 ($p < .05$) within three fixed factors as shown in Table 4.4.

Table 4.4

Varimax Rotated with Three Fixed Factors for Instructor Characteristics

Item	Factor		
	1	2	3
42	.780		
41	.772		
20	.743		
21	.723		
40	.713		
14	.643		
7	.633		
37	.629		
43	.621		
12	.602		
44	.601		
22	.595		
39	.584		
8	.584		
9	.495		
29		.851	
35		.848	
30		.842	
33		.791	
34		.772	
36		.769	
31	-.425	.716	
28		.707	
32		.666	
27		.620	
13	-.455	.492	
25		-.485	
18			.727
16			.660
17			.610
15	.433		.584
23			.562
5			.550
24	.456	-.429	.513
10			.511
6			.506
3			.431

The majority of the items had substantial factor loading structure coefficients while the first three factors accounted for 62.50% of the variance, as shown in Table 4.5.

Table 4.5

Factor Loading Coefficients within Three Fixed Factors

Factor	Eigenvalue	Explained Variance
Factor 1	17.965	48.55%
Factor 2	3.543	9.58%
Factor 3	1.618	4.37%

Final Instrument Factor Naming

The three factors that comprise instructor characteristics were analyzed to rename into appropriate and meaningful titles. Content review of the individual items resulted in the three factors being named Effective Delivery (15 items), Negativity (12 items), and Concern for Learning (10 items). Specific item statements are listed in Appendix J.

Research Question 2 - *How do students perceive instructor characteristics as measured by the developed instrument?*

This section reports the data analysis and findings concerning the second research question of the study. The

revised instrument created in response to research question one, resulted in 46 items within three factors; although, the last two items were excluded and combined into the dependent variable - Instructor Quality. The three factors were named Effective Delivery (15 items), Negativity (12 items), and Concern for Learning (10 items). Participants had the option to rate statements from 1 to 6 (1 = strongly disagree, 2 = moderately disagree, 3 = slightly disagree, 4 = slightly agree, 5 = moderately agree, and 6 = strongly agree). Table 4.6 reports the number of items, mean, standard deviation, minimum, and maximum factor item response ratings.

Table 4.6

Number of Factor Items, Mean, Standard Deviation, Minimum and Maximum Response Ratings for Instructor Characteristics

Factors	Number of Items	<i>M</i>	<i>SD</i>	Min	Max
Effective Delivery	15	73.86	12.85	25.00	90.00
Negativity	12	23.08	9.90	12.00	66.00
Concern for Learning	10	51.99	7.39	11.00	60.00

Note. Each factor had a different number of items. The scale ratings were 1 = strongly disagree, 2 = moderately disagree, 3 = slightly disagree, 4 = slightly agree, 5 = moderately agree, and 6 = strongly agree.

Factor 1 - Effective Delivery had the greatest minimum scale based on the number of rating 1 (strongly disagree) available responses (25.00), followed by Factor 2 - Negativity (12.00) and Factor 3 - Concern for Learning (10.00). Participants rated items higher in Effective Delivery ($M = 73.86$, $SD = 12.85$) which suggests a large number of positive agreement ratings, followed by Concern for Learning ($M = 51.99$, $SD = 7.39$), and Negativity ($M = 23.08$, $SD = 9.90$).

Factor 2 - Negativity showed that participants did not rate items in the greatest negative scale (Maximum = 66.00). Strongly agree ratings in Negativity equate to high levels of disinterest and an inability to inspire student performance. This indicates that respondents elected not to rate instructor characteristics at the extremely negative (72.00) while an even less critical rating is noted within the low average rating (strongly disagree) compared to the other two factors.

To gain a better understanding of the data, a comparison is made by adding the total number of rated item responses and then divided by the total number of available responses. The same rating scale used previously is used again here (1 = strongly disagree, 2 = moderately disagree, 3 = slightly disagree, 4 = slightly agree, 5 = moderately agree, and 6 = strongly agree).

On average, participants moderately-to-strongly agreed to items within the Concern for Learning factor ($M = 5.20$, $SD = 26.82$), strongly-to-moderately disagreed with Negativity ($M = 1.92$, $SD = 32.38$) items and moderately agreed with Effective Delivery ($M = 4.92$, $SD = 21.71$) items. Table 4.7 reports the averaged, minimum and maximum scores for each factor.

Table 4.7

Factor Mean, Standard Deviation, Minimum and Maximum Response Ratings of Averaged Scales for Perceived Instructor Characteristics

Factors	<i>M</i>	<i>SD</i>	Min	Max
Effective Delivery	4.92	.86	1.67	6.00
Negativity	1.92	.83	1.00	5.50
Concern for Learning	5.20	.74	1.10	6.00

Note. Each factor had differing number of items (Effective Delivery = 15 items; Negativity = 12 items; & Concern for Learning = 10 items). Rating scale utilized allowed for a 1-6 rating (1 = strongly disagree, 2 = moderately disagree, 3 = slightly disagree, 4 = slightly agree, 5 = moderately agree, and 6 = strongly agree).

Research Question 3 - *What are the student reactions to perceived instructor quality as measured by the instrument and student learning as reported by post-lesson test scores?*

This section presents the findings and data analysis for the third research question. The revised instrument created in response to research question one, resulted in 44 items within three factors and two items making up the dependent variable, Instructor Quality. This research question explored the student perception of Instructor Quality, comprised of two items, as reported in Table 4.8.

Table 4.8

Instructor Quality Mean, Standard Deviation and Response by Rating Scale

Items			1		2		3		4		5		6	
	<i>M</i>	<i>SD</i>	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
Instructor Quality ^a	4.73	1.07	32	23.5	57	41.9	33	24.3	8	5.9	4	2.9	2	1.5
Instructor Not Qualified ^b	1.74	1.56	11	8.1	4	2.9	5	3.7	2	1.5	10	7.4	104	76.5
Instructor Quality	6.46	1.65	43	15.8	61	22.4	38	14	10	3.7	14	5.2	106	39

Note. ^aScale ratings were 1 = one of the best, 2 = very good, 3 = above average, 4 = below average, 5 = very bad, and 6 = one of the worst. ^bScale ratings were 1 = strongly agree, 2 = moderately agree, 3 = slightly agree, 4 = slightly disagree, 5 = moderately disagree, and 6 = strongly disagree.

On average, participants moderately ($n = 57, 42\%$), slightly ($n = 33, 24\%$) and strongly ($n = 32, 24\%$) agreed that the instructor they were rating was one of the best of all instructors they have even had, including in high school. Additionally, the majority of participants strongly disagreed ($n = 104, 76\%$) with the statement that their instructor was not qualified to be an instructor.

Post-Lesson Test Scores

Post-lesson test scores were collected to report participant's mastery of the lesson material taught by the instructor participants rated using the instrument developed. Post-lesson test scores are shown in Table 4.9.

Table 4.9

Frequency and Percent of Post-Lesson Test Scores by Percentile

	<i>n</i>	<i>%</i>
Test Scores	136	100
100%	31	22.88
99-90%	50	36.72
80-89%	38	27.90
<79%	17	12.50

The average post-lesson test scores were 90.23 percent ($SD = 8.03$) with a range of 65 to 100 percent. Of the participants, 31 (22.9%) received perfect mastery scores of 100 percent and 17 participants (12.5%) demonstrated non-mastery of the lesson objectives by scoring less than 79 percent on the post-lesson test.

Research Question 4 - Are there relationships between the instructor characteristics and student's reaction to perceived instructor quality as measured by the instrument and student learning as reported in post-lesson test scores?

This section presents the findings and analysis of the data for research question four. The fourth research question explored the relationships between the instructor characteristic factors, student's perceived instructor quality, and post-lesson test scores. Correlations were measured for each factor (Effective Delivery, Negativity, and Concern for Learning) instructor quality and summated scores. Table 4.10 presents the relationships.

None of the instructor characteristic factors had a statistically significant relationship with the perceptions of the instructor quality. Only Factor 1 - Effective Delivery ($r = .130$, $p = .005$) was statistically significant with post-lesson

test scores; although, it would be interpreted as a low relationship according to Davis (1971). Interestingly, when the ratings of effective delivery increased the test scores decreased and when ratings of effective delivery decreased, test scores increased. Neither of the remaining factors (Negativity and Concern for Learning) had significant relationships with test scores.

Table 4.10

Relationship between Perceived Instructor Quality, Post-Lesson Test Scores and Instructor Characteristics as Measured by Pearson's Product Moment Correlation

	Perception of Instructor Quality			Test Scores		
	<i>r</i>	Interpretation	<i>p</i>	<i>r</i>	Interpretation	<i>p</i>
Effective Delivery	.130	Low	.133	-.240	Low	.005
Negativity	-.156	Low	.070	.059	Negligible	.496
Concern for Learning	.028	Negligible	.750	-.024	Negligible	.781

Note. Davis (1971) describes relationships with: .70 or higher = very strong association, .50 to .69 = substantial association, .30 to .49 = moderate association, .10 to .29 = low association and .01 to .09 = negligible association. The correlation between instructor quality and the test scores was $-.213$ ($p = .013$).

Research Question 5 - *Do instructor characteristics explain a significant amount of variance in perceived instructor quality?*

This section presents the analysis of data and findings for the fifth research question. Research question five sought to determine if instructor characteristics (comprised of three factors) explained a significant amount of variance in perceived instructor quality. Step-wise regression was employed to determine the magnitude of relationships between multiple variables in order to identify the best combination of predictors (independent variables) to dependent variables. Initially, a step-wise regression using the probability of F for entry into the model at .05 and removal from the model at .10 was calculated. No variables entered into the model. To further explore the relationship of the variables, a model was calculated using a full-entry model. The multiple regression ($F = 2.653$, $p = .051$) revealed that 5.7 percent of the variance of perceived instructor quality was explained by the three instructor characteristic factors. However, according to Bartlett, Kotrlik and Higgins (2001) an R^2 less than .13 or 13 percent would be considered a small effect size and not practically significant. Table 4.11 reports the results of step-wise regression analysis

used to examine the amount of variance instructor characteristics explain in perceived instructor quality.

Table 4.11

Step-wise Multiple Regression Analysis Using Forced Entry to Explore if Instructor Characteristics Explain a Significant Amount of Variance in Perceived Instructor Quality

Source of Variation	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Regression	20.921	3	6.974	2.653	.051
Residual	347.032	132	2.629		
Total	367.952	135			

	Overall <i>R</i> ²	<i>B</i>	<i>Beta</i>	<i>t</i>	<i>p</i>
Constant (8.667)	5.7%			5.095	<.001
Effective Delivery		.030	-.568	1.721	.088
Negativity		-.037	-.225	-1.910	.058
Concern for Learning		-.069	-.326	-2.071	.040

Research Question 6 - Do instructor characteristics explain a significant amount of variance in post-lesson test scores?

This section presents the analysis of data and findings for the final research question. The sixth research question sought to determine if instructor characteristics (comprised of three factors) explained a significant amount of variance in perceived post-lesson test scores. The step-wise regression used the

probability of F for entry into the model at .05 and removal from the model at .10.

Table 4.12 presents the results of the step-wise regression analysis used to examine if perceived instructor characteristics explain a significant amount of variance in post-lesson test scores.

The step-wise multiple regression ($F = 3.330, p = .070$) revealed that 2.4% of the variance in post-lesson test score is explained by students' perceived instructor negativity ($B = -.156$). However, according to Bartlett, Kotrlik and Higgins (2001) an R^2 less than .13 or 13 percent would be considered a small effect size and not be practically significant.

The student's perception of Factor 1 - Effective Delivery ($t = .519, p = .604$) and Factor 3 - Concern for Learning ($t = -1.256, p = .211$) did not enter the model.

Table 4.12

Step-wise Multiple Regression Analysis to Explore if Perceived Instructor Characteristics Explain a Significant Amount of Variance in Post-Lesson Test Scores

Source of Variation	SS	df	MS	F	p	
Regression	8.921	1	8.921	3.330	.070	
Residual	359.031	134	2.679			
Total	367.952	135				
	Overall R ²	B	Beta	t	p	Collinearity Tolerance
	2.4%					
Constant (7.060)				19.772	<.001	
Negativity		-.026	-.156	-1.825	.070	
Excluded Variables						
Effective Delivery				.519	.604	.635
Concern for Learning				-1.256	.211	.537

Summary of Analysis

Chapter four presented the data analysis results and findings for the study. Each research question will be summarized below.

Research question one employed exploratory factor analysis to identify perceived instructor characteristics (within three factors) in order to create a survey instrument. Instructor

quality (the last two opinion-style items, combined) and post-lesson test scores were reported by frequency and percent. Descriptions and frequency of rated responses were used to address research questions two and three. Correlations were calculated and presented for research question four. Step-wise multiple regression analysis was calculated for research questions five and six. The study also produced findings demonstrating the current U.S. Marine entry-level vocational trainee. Participant demographics included age, gender, education and ethnicity.

The findings from research question one allowed for the creation of an instrument by identifying instructor characteristics. Three factors resulted from the second round of exploratory factor analysis, Effective Delivery, Negativity, and Concern for Learning. Findings from research question two indicate that participants positively rated their instructors on Factor 1 - Effective Delivery and Factor 3 - Concern for Learning, while rating instructors poorly on Factor 2 - Negativity items.

Research question three demonstrated the majority of students rated their instructor high in instructor quality while post-lesson test scores indicate over 90% of the participants

mastered the lesson material without remediation. Research question four used correlation analysis to show that none of the instructor characteristic factors had significant relationships with the student's perception of instructor quality and only Factor 1 - Effective Delivery showed a low relationship with post-lesson test scores. Research question five showed that no amount of variance was explained by instructor characteristics in perceived instructor quality. Only when forced into full-entry model do instructor characteristic factors explain a non-practical 5.7% of variance found within perceived instructor quality. Research question six revealed instructor characteristic factors explained a non-practical 2.4% of variance found in post-lesson test scores.

CHAPTER 5

Summary, Conclusions, And Recommendations

"Man cannot survive except by gaining knowledge, and reason is his only means to gain it. Reason is the faculty that perceives, identifies and integrates the material provided by his senses. The task of his senses is to give him the evidence of existence, but the task of identifying it belongs to his reason, his senses tell him only that something is, but what it is must be learned by his mind."

~ Ayn Rand

This chapter provides a summary of the study beginning by reaffirming the purpose and goal of the study, data collection procedures, a description of the participants, and a detailed examination of the rated instructors. Each of the study's chapters is summarized; whereas, chapter one reiterates the theoretical and conceptual frameworks used for the study. A review of the research objectives and the findings and conclusions of each research question is provided next. The chapter concludes with a review of the study's overall conclusions and limitations, and recommendations for future research.

Purpose of the Study

The purpose of the study was to examine instructor characteristics for application in the indoctrination, monitoring and evaluation processes undertaken by formal school vocational training instructors within U.S. Marine Corps entry-level occupational specialty courses. Training effectively in an efficient manner allows for superior management of resources; therefore, a great majority of leaders require, collect and tabulate vast quantities of raw data to determine if 1) the instructor is qualified, 2) did the trainee master the learning material, and 3) what are the instructor characteristics that would lead to continued success.

The Goal of the Study

This study sought to identify instructor characteristics capable of predicting student's reaction to instructor quality and student's performance on post-lesson performance testing. The overall goal of the study was to identify specific instructor characteristics that could be first factored and then used to create an instrument that would benefit administrators (who could empirically assess their instructors), instructors (to improve their training skills), and finally the student (who could rate

their instructors with the expectation of their reactions actually improving future training sessions).

Data Collection

The study examined perceived instructor characteristics, instructor quality, and student learning. Data were collected from U.S. Marine Corps Combat Service Support Schools located in Jacksonville, North Carolina during the months of May 2010 to July 2010. Participants were solicited from courses in session at the time of data collection and were selected by availability through purposive nonrandom sampling. Scaled responses were collected using a survey instrument designed to capture student reactions to perceived instructor characteristics and instructor quality. Post-lesson test scores were collected once lessons were finalized.

Summary of Participant Demographics

The average participant in the study (both field test and final round) was white ($n = 250, 56.4\%$), male ($n = 408, 92.1\%$), with a high school diploma ($n = 385, 86.9\%$) and an average age of 21.10 years ($SD = 2.252$).

Summary of U.S. Marine Corps Vocational Instructors

It is relevant to the study to discuss the U.S. Marine Corps vocational instructors whom the participants rated. As mentioned previously, all instructors are selected based on availability with limited, if any, screening prior to arrival.

The overwhelming majority of instructors are mid-career enlisted personnel usually possessing only a high school diploma; although, a few do possess post-secondary and even graduate level degrees. All instructors, both male and female, are in superb physical shape with superior leadership skills having served between 8- to 12- years. Every instructor is considered a subject expert in their particular field with equitable AFQT scores required to serve within U.S. Marine Corps combat service support (CSS). Some of the vocational fields that comprise CSS include: vehicle maintenance, administration, finance, supply, and distribution. Instructor qualification, rank, level of responsibility and time-to-advancement are similar across all CSS fields. Therefore, because of the homogeneity of the CSS occupations, the uniformity of U.S. Marine Corps leadership, vocational, and instructor training as well as the service-wide standardization of consistent rules, regulations and traditions,

CSS vocational instructors are very comparable at an organizational level.

Consequently, the instructors that the participants rated were very similar in ability, education, vocational competency, age, and time-in-service. Their knowledge, experience, and leadership are very analogous; therefore, it is difficult, if not meaningless, to measure instructor differences as they relate to this population. These differences would be more a product of personality or preference, than any uniform, interchangeable instructor characteristics assigned solely to a particular CSS field. If participants and instructors had been selected from alternate vocational fields; e.g., aviation, nuclear submarine, infantry, cryptology, or language, then it would be reasonable to consider potential differences.

Research Objectives

The study had three primary research objectives. They were:

The first research objective was to determine and align instructor characteristics into distinct factors that could be used to predict student's reaction to instructor quality and student's performance on post-lesson examination. This objective was met with an extensive literature review to first identify

statements that could be used during instrument development followed by exploratory factor analysis to reduce those statements based on factor loading strength. Three factors were defined: Effective Delivery, Negativity, and Concern for Learning.

Research objective two focused on exploring the instructor characteristics (as factors) that explain student's scaled rating response of satisfaction in regards to perceived instructor quality. This research objective was met through Pearson's Product Moment correlation and step-wise multiple regression analysis that examined the three instructor characteristic factors for relationships with the student's scale ratings of perceived instructor quality.

The final research objective was to explore the factors that comprise instructor characteristics that explain student learning evidenced by post-lesson test scores. This research objective was met through Pearson's Product Moment correlation and step-wise multiple regression analysis that examined the three instructor characteristic factors for relationships with the student's post-lesson test scores.

Summary of Chapters

The study was built around five chapters. To better summarize the entire study, each chapter is described in the following paragraphs.

Chapter One

Chapter one introduced the study's research objectives and the associated research questions that guided the study. The first two research questions dealt with 1) what students consider to be instructor characteristics, and 2) how students rate those items on an evaluation instrument using their current instructors aboard a U.S. military training installation. The third research question sought to capture student's rated responses of perceived instructor quality and their post-lesson test scores. Research question four sought to identify any relationships that might exist between scaled instructor characteristics, perceived instructor quality and post-lesson test scores. The fifth and six research questions sought to explain the amount of variance between instructor characteristics and perceived instructor quality and post-lesson test scores, respectively. The research design and methodology are introduced with the study's variables.

The theoretical framework underlying the study is derived from instructional systems design (ISD). As ISD contains elements of behaviorism and cognitivism, it was found to be used effectively in military performance based training where little collegiate level educational theory is introduced or required. Students are taught with well-structured lessons comprised of linked learning objectives which lead lock-step to performance mastery of a specific task.

The conceptual framework shows an interaction between the independent variables (instructor characteristics) and the dependent variables (perceived instructor quality and post-lesson test scores). Although the original concept included nine factors this study's findings reduced that number down to three practical factors using exploratory factor analysis and multiple regression analysis.

The purpose and goals of the study, delimitations, and significance in terms of a greater understanding of the relationship between student reactions to perceived instructor quality and student learning were introduced. Chapter one concludes with a brief description of the study's organization.

Chapter Two

Chapter two describes the comprehensive examination of literature related to student evaluation and satisfaction, the consumer model, effective teaching, and instructor performance. The chapter examines multiple learning theories (humanism, adult education, andragogy, social learning, behaviorism, cognitivism and systematic instructional design) in order to explain the study's theoretical framework. As a blend of two theoretical schools of thought - behaviorism and cognitivism - into a single construct called Instructional Systems Design (ISD), U.S. Marine Corps vocational training is effective at systematically identifying the procedures and components necessary to create effective learning scenarios with little or no experience in the field of education or training. The chapter concludes with an overview of the U.S. Armed Forces for the benefit of non-military readers. Specific areas discussed include the enlistment process, qualifications, assignment to occupational specialties, training both initial and vocational, military evaluation of training and the role of the military instructors specific to the U.S. Marine Corps.

Chapter Three

The third chapter describes in detail the research methodology used for this study. Item screening and selecting, IRB approval, survey construction and facilitation were discussed. Sample size to required items, missing and outlier data procedures, and overall data collection was described in detail. Field testing of the initial instrument, preliminary data analysis, and findings were provided. Exploratory factor analysis was conducted on the instrument using SPSS 18.0. Reliability tests were conducted on the factors using Cronbach's alpha. Factors were reduced as items were discarded due to low loading coefficients which resulted in a revised instrument containing roughly half of the initial field test instrument.

Chapter Four

Chapter four presents the analyzed findings of the collected data. Findings were presented that address the research questions. A description of the participant's demographics, correlations between the study's variables, and regression analysis used to explore relationships between perceived instructor characteristic factors, instructor quality and post-lesson test scores is provided.

Chapter Five

The fifth and final chapter provides a detailed summary of the study's findings and conclusions by research question. Lastly, overall study conclusions, recommendations for future research, and limitations of the study are provided.

Findings and Conclusions by Research Question

There were six research questions guiding the study. The first question described instructor characteristics that could explain student's rating of instructor quality and subsequent post-test lesson scores. Research question two sought to describe how students perceive instructor characteristics as ratings on a scaled instrument. Research question three sought to identify student's reactions to rated perceived instructor quality and student learning. The fourth question sought to determine if a relationship existed between instructor characteristics and student's reaction to perceived instructor quality and student learning. Questions five and six sought to discover if instructor characteristics explained significant amounts of variance in perceived instructor quality and post-lesson test scores. Each research question is described in detail, with the researcher's conclusions, in the following paragraphs.

The findings from research question one allowed for the creation of a practical instrument. This was done by identifying those instructor characteristics that were rated and correlated highly by students.

Conclusion One: By reviewing literature related to student evaluation of trainer effectiveness, concurrent to collecting individual items from peer reviewed instruments, we were able to examine and identify consistent themes within the literature. Factors, sub-factors and items were conceptualized from these themes that led to the development of an instrument. The instrument was piloted and revised using exploratory factor analysis to identify those items with high factor loadings for application in U.S. Marine Corps vocational training. From the original 1514 items identified, 37 items remained within three factors, effective delivery (15 items), negativity (12 items), and concern for learning (10 items), capable of assessing student perceptions of instructor characteristics rated within this population group. Organizational guidance may be required to establish protocol procedures to direct the use of these rated responses; i.e., used for trend analysis, instructor monitoring, or climate surveys reported from the student's vantage point.

Research question two examined how students rated their instructor on the instructor characteristics found in the created instrument. Analyzing the participant responses to items within the three factors, effective delivery, negativity, and concern for learning, the results of the study suggest that students were capable of rating their instructors using the study's instrument 1-6 scale (with a 1 representing least positive and 6 representing most positive).

Items within effective delivery were related to the instructor's proper use of examples and control of the learning environment, the level of interest student's perceived in the training sessions, and the student's perception of the instructor's motivation, enthusiasm, and projected enjoyment gained from training students. The items found in the effective delivery factor were rated highly ($M = 4.92$, $SD = .86$) implying that the instructor's rated during this study were perceived as highly effective at delivering lesson material.

The negativity factor included items that demonstrated perceived lack of respect or blatant contempt for the students. Publically criticizing students, lacking interest in student learning, cynical attitudes, angry, argumentative, incomplete or factually wrong responses to student questions, and unfair

evaluation are all items found within this factor. In this study, student's perceived and rated instructors low in this factor ($M = 1.92$, $SD = .83$) which indicates that instructors did not exhibit these characteristics to a high degree.

Items found in factor three, concern for learning, were associated with the instructor's ability to clearly define requirements, assignments, objectives and demonstrations, their effective lesson pace, questioning and listening techniques, and their proper use of technology and training equipment. Students rated their instructors high in this factor ($M = 5.20$, $SD = .74$) suggesting that these instructor characteristics were highly evident in this study's sample.

Conclusion Two: The results of the study suggest that students were capable of rating their instructors using the study's instrument. In this study, participants perceived and rated their U.S. Marine Corps vocational instructors high on Factor 1 - Effective Delivery and Factor 3 - Concern of Learning, and low on items within Factor 3 - Negativity.

Research question three sought to explore student perception of instructor quality. The study's findings demonstrate that the majority of U.S. Marine Corps entry-level students ($n = 122$, 89.7%) considered their instructors to be

among the best instructors they have ever had - including high school. Only 14 participants (10.3%) rated their instructor negatively as below average ($n = 8, 5.9\%$), very bad ($n = 4, 2.9\%$), or one of the worst ($n = 2, 1.5\%$).

In regard to their instructor being unqualified to instruct, the majority of participants ($n = 116, 85.4\%$) disagreed with the statement; although, 24 participants (14.7%) did agree with the statement. This implies that the majority of students (85.4%) perceive their instructors as qualified to instruct.

Post-lesson test scores were collected to report participant's mastery of the lesson material taught by the instructor participants rated using the instrument developed. The average post-lesson test scores were 90.23 percent ($SD = 8.03$) with a range of 65 to 100 percent. Of the participants, 31 (22.9%) received perfect mastery scores, another 64.62 mastered all learning objectives initially while 17 participants (12.5%) failed to demonstrate mastery by scoring less than 79 percent on the post-lesson test.

Conclusion Three: The majority of instructors were perceived by respondents as qualified and overall 87.5% of the participants demonstrated initial mastery of learning objectives, as reflected by post-lesson test scores. Together, this gives

the impression that students consider their instructors qualified and are mastering the learning objectives.

The fourth research question explored the relationships between the instrument's instructor characteristic factors and student's perceived instructor quality and post-lesson test scores. Instructor quality was comprised of two instrument items, the student's opinion of the instructor, compared with all other instructors from which the student has been taught, and the student's perception of the instructor's qualification to teach. Correlations were calculated for each factor (effective delivery, negativity, and concern for learning), summated scores, and instructor quality. In this study, none of the instructor characteristic factors identified in the previous research questions had a statistically significant relationship with the student's perceptions of the instructor quality.

In regards to instructor characteristic factor's relationship to post-lesson test scores only Factor 1 - Effective Delivery was statistically significant with post-lesson test scores; although, the significance would be interpreted as a low relationship. Neither of the remaining factors had significant relationships with test scores.

It was also curious to note an inverse relationship between effective delivery and post-lesson test scores; i.e., when the ratings of effective delivery increased the test scores decreased and conversely when effective deliver decreased the test scores increased.

Conclusion Four: Respondents were able to identify instructor characteristics and rate those same characteristics within their instructors (as shown in research question two). Therefore, the participants witnessed and validated instructor characteristics, shown as items on the instrument, which were drawn from other peer-reviewed literature.

Secondly, the majority of participants reported high ratings of satisfaction with their instructors and their instructor's qualifications to teach. This demonstrates their ability to appreciate, and discriminate, between quality and poor instruction.

Lastly, the majority of the respondents mastered all learning objectives without remediation. This shows that somehow, either by instructor characteristics, course design, situation, simplicity, or some other unknown variable, the students are learning the lesson material.

This could be due to the vocational nature of the training participants. Knowles (1995) describes a "strong relationship between achieving technical mastery and being self-directed" (p. 216). This implication is that the learning concept changes from being externally directed to the "internal discipline provided by the learner" (Knowles, 1995, p. 216). U.S. Marine Corps entry-level students are adult learners; therefore, they possess some level of andragogy especially in regard to self-directed learning. Also, as self-directed learners they are responsible for their own learning and apparently succeed regardless of perceived instructor characteristics of their trainers.

In regards to the divergence between ratings of effective delivery and test score, Perry, Hall, & Ruthig (2007) describe the concept of academic control where students who feel in control of their own learning "work harder, feel better about their studies, obtain better grades, and have more productive academic careers than their low-control counterparts" (p. 480). In this situation a high control student may thrive in autonomous situations but become irritated in highly structured environments; although, they would still perform at a level sufficient to master the learning objectives. Conversely, a low control student may thrive within a low control environment;

i.e., an instructional method that uses highly structured and predictable lecture methods - which effectively describe U.S. Marine Corps vocational training.

Research question five sought to determine if instructor characteristics explained a significant amount of variance in perceived instructor quality. No variables entered into the model. Only when forced into a full-entry model does a non-practical 5.7% of instructor characteristic variance explain perceived instructor quality.

Conclusion Five: Although respondents were able to recognize and rate instructor characteristics within their instructors this ability had little to do with explaining their perception of instructor quality. When forced, instructor characteristics do account for a little less than 6% of variance in instructor quality, which is impractical for use in the classroom or instructor development program.

The sixth research question sought to determine if instructor characteristics explained a significant amount of variance in perceived post-lesson test scores. Results revealed that 2.4% of the variance in post-lesson test scores was explained by student's perceived instructor negativity; although,

with a small effect size. The effective delivery and concern for learning factors did not enter the model.

Conclusion Six: Respondents post-lesson test scores were affected by items found within the negativity factor; although, at a non-practical significance. Effective delivery and concern for learning factors had no significant affect on learning. This demonstrates the respondent's ability to note and rate instructor characteristics, whether positive or negative, but that those same ratings do not affect their performance.

Overall Conclusions and Recommendations

The current study created an instrument to assess instructor characteristics within U.S. Marine Corps vocational instructors. The study's findings indicate that U.S. Marine Corps entry-level vocational students are able to discern and rate instructor characteristics in their instructors, and that their instructors were considered fully qualified and some of best they have ever experienced.

Furthermore, the study found that regardless of how the student perceived and rated instructor characteristics, there was little relationship to their performance. With little significance between instructor characteristics, instructor

quality and post-lesson test scores it is difficult to identify specific factors or elements capable of predicting student ratings or post-lesson test scores. Students seem able to discern instructor quality and to rate instructor characteristics but regardless neither action seemed to predict student performance.

This illustrates a critical area of concern. If student reactions, exemplified by instrument satisfaction rating scales, fail to demonstrate significant relationships to perceived instructor quality and performance; why are nearly 90% of organizations still using customer satisfaction ratings? If the reaction rating forms are incapable of predicting performance, then it is reasonable to question their purpose and associated resource expense.

Evaluation is an expensive element of the training process regardless of the organization. Most organizations still employ the most basic of evaluation tools, customer reaction forms, as the initial component of Kirkpatrick's evaluation model. Faerman & Ban (1993) suggest that organizations generally use end-of-course reaction forms because it is too difficult and too resource intensive to actually measure learning; i.e., a change in a student's behavior. Consequently, it appears that a great

many organizations are using reaction forms as a simplistic, cost-efficient, check-in-the-box regardless of whether the data is relevant or accurate. By simply collecting customer satisfaction forms from all students, organizational leaders are able to quantify their program's efficiency by the sheer volume of positive responses. But this is one of the weaknesses in Kirkpatrick's evaluation model.

Kirkpatrick's evaluation model describes four stages, reaction, learning, behavior, and results. This study examined the first two stages, reaction and learning. Student reactions were captured by the instrument items and post-lesson test scores demonstrate that learning has occurred; although the findings indicate little relationship between the two.

The results of this study support Nancy Dixon's (1990) conclusions that Kirkpatrick's model failed to support any significant relationship between "trainee perceptions of enjoyment and their posttest scores" and no significant relationship between "perceptions of instruction skillfulness and their posttest scores" (p. 137). Kirkpatrick (1979) himself warned that there was "no assurance that any learning has taken place" (p. 45); therefore, what benefit does the reaction form serve?

One erroneous benefit that organizations often accept as legitimate is using the quantified student satisfaction ratings of an individual instructor's performance to rank order their entire training staff. But there are inherent dangers to using these rating forms for this purpose. Rank ordering your instructor staff based on the student's satisfaction of the instructor, could lead to organizational strife and higher employee turnover. Even though information can be quantified, results are still subjective - especially in regards to the participant's opined reaction to training.

Another potential danger could be that instructors would intentionally "dumb-down" the lesson material, or become classroom "edu-tainers" (educator-entertainer), simply to gain higher ratings. The end-state is a degradation of the quality of education where students are provided with an "easy-A" course instead of a challenging, thought-provoking, learning experience. In this situation, instructor popularity replaces subject expertise and training innovation, especially if those ratings lead to salary increases or tenure.

Lastly, the student-instructor relationship could turn adversarial if students come to believe that the instructor really isn't an authority because they are not trusted by the

front office to train the students. This could easily happen when the instructor's supervisor reaches down from on high, around the instructor, to gather the student's opinion of their level of satisfaction of the instructor's performance. Experienced leaders understand that this only serves to undermine the trust and confidence of the subordinate's capabilities. Instead of the supervisor telling the students "Here is the best of the best, a most qualified, experienced and knowledgeable expert; learn everything you can from them," the use of student rating forms tell the student "We brought this instructor in to teach you, but we are not even sure if they can teach or if they have taught you anything at all. Please tell us if we made a mistake and need to find someone else." Also, in many cases, the instructor is forbidden from even touching the student rating forms; a student is tasked to collect and deliver the forms to the administration office. This not-so-subtle signal degrades the instructor's authority telling the students that the instructor can't be trusted and is likely to sneak and cheat by changing the forms - if only given the chance.

With so many potentially harmful outcomes of using student reaction forms, organizations must ask: Should we define "training effectiveness" simply in terms of student satisfaction?

By extension, if student ratings on a survey instrument only report the level of satisfaction, how does an organization legitimately assess training effectiveness?

If we adhere to the theoretical Instructional Systems Design (ISD) model, and subsequently the Analyze, Design, Develop, Implement, and Evaluation (ADDIE) process, then organizations are bound to evaluate their training processes. The question remains: How can organizations effectively and efficiently evaluate their training programs?

The literature provides various alternatives to student reaction forms that allow for organizations to measure training effectiveness. First and foremost, this study recommends that program effectiveness should be considered as a product of student performance shown by post-lesson test scores, attrition rate, and required remediation. Reduction in training time without sacrificing quality, as a result of training innovation or methodology, would also demonstrate effectiveness and encourage instructor initiative.

Other recommended methods to evaluate training effectiveness include instructor-peer evaluation, lesson accreditation plans, and training portfolios. All of these recommendations remove the student from the subjective evaluation

process and lays it where it belongs - with the instructor's peers, mentors and trained evaluators. This process would require judgments from other training members and would likely increase evaluation time and effort, but the end state would be instructors rating, developing, mentoring and certifying instructors vice students reporting their level of satisfaction. These methods also stay true to the ISD and ADDIE theoretical models.

Limitations

Population and Sample. A significant limitation was the non-random, purposive sample composed entirely of entry-level U.S. Marines. Confounding factors could include the fact that all participants elected to enlist in the U.S. Armed Forces, specifically the U.S. Marines, which must attest to some similarity in thought, capabilities, drive, and motivation. Secondly, all participants received roughly the same ASVAB vocational scores which were used to assign them into combat service support. Lastly, all instructors were mid- and senior-level U.S. Marines which could lead to a type of "brand-loyalty" biased ratings from entry-level U.S. Marines.

Instrument Design. The nature of vocational study limited the survey questionnaire to a paper-based instrument. Results could

vary if participants were in a comfortable, relaxed environment vice restricted to the classroom during the collection process. Facilitator. The study employed a single facilitator to administer and collect all survey instruments. The facilitator was not a military person dressed in military garb, clean shaven with a military haircut. Rather the facilitator was a civilian dressed in professional business attire introduced as a NCSU graduate student and former U.S. Marine. By using a civilian facilitator the study may have added an unrecognized barrier to the exclusively military population. Student responses may have varied if the facilitator was a military person in a perceived position of respected authority.

Recommendations for Future Research

Our literature review demonstrates an overabundance of articles related to instructor characteristics and student evaluation of effectiveness, most either strongly supporting or opposed to their use. For the novice instructor it is difficult to learn exactly what works for their particular settings. It is also equally difficult for the administrator or instructional designer to try to assess instructor effectiveness with so many available options and references.

The study's findings illustrate key concerns with the use of student reaction forms to measure program success if student performance is the desired end-state. It is recommended that administrators consider the overall resources required to develop and validate relevant instruments, the time and personnel it takes to administer, collect, tally, and decipher the raw data into useable information, and the effect the results will have on the instructor staff when used. If the resource cost is too high, alternate methods of program effectiveness are recommended to include, instructor accreditation plans, instructor peer evaluations and training portfolios. Potential future research could include creating a peer-review evaluation instrument, based on this study's final factored items, to identify if significant relationships exist between those instructor characteristics and the peer's perception of instructor quality and associated student post-lesson test scores. The result would be an evaluation instrument capable of rating perceived instructor quality with the goal of developing instructor competencies without subjective student involvement.

Additional research specific to this population is suggested using the final instrument in conjunction with a round of personality indexing tools, like the Myers-Briggs Personality

Inventory, prior to instructor characteristic rating; therefore, the researcher could align personalities to instructor characteristics to perceived instructor quality and performance scores. This could support the high-low academic control concepts described by Perry, Hall & Ruthig (2007).

Lastly, follow-on research could be conducted by directly manipulating the variables; i.e., programmed instructor characteristics with assigned participants into control settings. Assignment to experimental or control settings could demonstrate empirical causation which may lead to stronger confirmation or prediction.

REFERENCES

- Abrami, P. C., & d'Apollonia, S. (1999). Current concerns are past concerns. *American Psychologist*, 54(7), 519-520.
- Abrami, P. C., d'Apollonia, S., & Rosenfield, S. (2007). Dimensionality of student ratings of instruction: What we know and what we do not. In R. P. Perry, & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*, (pp. 385-445). The Netherlands: Springer.
- Abrami, P. C., d'Appollonia, S., & Cohen, P. A. (1990). Validity of student ratings of instruction: What we know and what we do not. *Journal of Educational Psychology*, 12(2), 219-231.
- Akerhielm, K., Berger, J., Hooker, M., & Wise, D. (1998). *Factors related to college enrollment*. Washington, DC: U.S. Department of Education.
- Aleamoni, L. M. (1981). Student ratings of instruction. In J. Millman (ed.), *Handbook of teacher evaluation*, (pp. 110-145). Newbury Park, CA: Sage.
- Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology*, 42, 331-342.

- Andrew, M. D., Cobb, C. D., & Giampietro, P. J. (2005). Verbal ability and teacher effectiveness. *Journal of Teacher Education, 56*(4), 343-354.
- Arubayi, E. A. (1987). Improvement of instruction and teacher effectiveness: Are student ratings reliable and valid? *Higher Education, 16*(3), 267-278.
- Asch, B. J. (1994). *Military support for youth development: An exploratory analysis*. Santa Monica, CA: RAND Corporation.
- Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal, 19*(1), 43-50.
- Bartlett, K. R. (2005). Survey research in organizations. In R. A. Swanson & E. F. Holton (Ed.) *Research in organizations: Foundations and methods of inquiry*, (pp. 97-113). San Francisco, CA: Berrett-Koehler.
- Bearor, J. W. (2004). *The systems approach to training*. Quantico, VA: U.S. Marine Corps.
- Bell, T. H. (1962). *Effective teaching: How to recognize and reward competencies*. New York: Exposition Press.

- Benson, T. A., Cohen, A. L., Buskist, W. (2005). Rapport: Its relation to student attitudes and behaviors toward teachers and classes. *Teaching of Psychology, 32*(4), 237-270.
- Bertrand, Y. (2003). *Contemporary theories and practice in education*. Madison, WI: Atwood Publishing.
- Borko, H., & Livingston, C. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Educational Research Journal, 26*(4), 473-498.
- Braskamp, L. A., Caulley, D., & Costin, F. (1979). Student ratings and instructor self-ratings and their relationship to student achievement. *American Educational Research Journal, 16*(3), 295-306.
- Brinkerhoff, R. O. (1995). Using evaluation to improve the quality of technical training. In L. Kelly (Ed.), *The ASTD Technical Skills Training Handbook* (pp. 385-409). New York: McGraw-Hill.
- Brown, W., & Tomlin, J. (1996). Best and worst university teachers: The opinions of undergraduate students. *College Student Journal, 30*(4), 431.

Bureau of Labor Statistics (2008). *Occupational outlook handbook, 2008-09 edition: Job opportunities in the U.S. Armed Forces*. Retrieved September 22, 2008, from <http://www.bls.gov/oco/ocos249.htm>.

Bureau of Labor Statistics (2010a). *College enrollment and work activity of 2009 high school graduates*. Retrieved 1 October 2010 from <http://stats.bls.gov/news.release/hsgec.toc.htm>.

Bureau of Labor Statistics (2010b). *Occupational outlook handbook, 2010-11 edition*. Retrieved 1 October 2010 from <http://www.bls.gov/oco/ocos249.htm#outlook>.

Burton, J., Moore, D., & Magliaro, S. (1996). Behaviorism and instructional technology. In D. H. Jonassen (Ed.), *Handbook of Research for Educational Communications and Technology* (pp. 3-36). New York: Macmillan.

Centra, J. A. (1973). Self-rating of college teachers: a comparison with student ratings. *Journal of Educational Measurement, 10*(4), 287-295.

Centra, J. A. (1977). Student ratings of instruction and their relationship to student learning. *American Educational Research Journal, 14*(1), 17-24.

- Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education, 44*(4), 495-518.
- Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education, 28*(1), 71-88.
- Chu, D. S. (2005). *Department of Defense Instruction 1145.01, Qualitative distribution of military manpower*. Washington, DC: Department of Defense.
- Chu, D. S. (2006). *Department of Defense Instruction 1304.30, Enlisted Personnel Management Plan (EPMP) procedures*. Department of Defense: Washington, DC.
- Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection. *Review of Educational Research, 51*(3), 281-309.
- Cohen, W., & Shelton, H. O. (1998, March 16). *Response to the Kassebaum Baker report on gender-integrated training*. Retrieved October 21, 2008, from U.S. Department of Defense News Transcript:
<http://www.defenselink.mil/transcripts/transcript.aspx?transcriptid=770>.

- Congressional Budget Office (2007). *All volunteer military*. Retrieved October 13, 2008, from <http://purl.access.gpo.gov/GPO/LPS83766>.
- Courtney, S. (1989). Defining adult and continuing education. In S. B. Merriam and P. M. Cunningham (Eds.), *Handbook of adult and continuing education* (pp. 15-25). San Francisco, CA: Jossey-Bass, Inc.
- Crader, K. W., & Butler, J. K. (1996). Validity of students' teaching evaluation scores: The Wimberly-Faulkner-Moxley questionnaire. *Educational and Psychological Measurement*, 56(2), 304-314.
- Davidovitch, N., & Soen, D. (2006). Using students' assessments to improve instructors' quality of teaching. *Journal of Further and Higher Education*, 30(4), 351-376.
- Davis, J. A. (1971). *Elementary survey analysis*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Democratic Leadership Council. (1988). *Citizenship and National Service*. Retrieved October 23, 2008, from http://www.ndol.org/ndol_ci.cfm?contentid=250403&kaid=115&subid=145.

- Department of Defense. (2008). *Fiscal Year Budget: Military personnel programs*. Washington, DC: Office of the Under Secretary of Defense.
- Detchen, L. (1940). Shall the student rate the professor? *Journal of Higher Education, 11*(3), 146-154.
- Dick, W., Carey, L., & Carey, J. O. (2005). *The systematic design of instruction (6 ed.)*. Boston: Pearson.
- Dixon, N. M. (1990). The relationship between trainee responses on participant reaction forms and posttest scores. *Human Resource Development Quarterly, 1*(2), 129-137.
- Faerman, S. R., & Ban, C. (1993). Trainee satisfaction and training impact: Issues in training evaluation. *Public Productivity & Management Review, 16*(3), 229-314.
- Feldman, K. A. (1996). Identifying exemplary teaching: Using data from course and teacher evaluation. *New Directions for Teaching & Learning, 65*, 41-50.
- Feldman, K. A. (2007). Identifying exemplary teachers and teaching: Evidence from student ratings. In R. P. Perry and J. C. Smart (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*. The Netherlands: Springer.

- Fenker, R. M. (1975). The evaluation of university faculty and administrators: A case study. *The Journal of Higher Education, 46*(6), 665-686.
- Fields, G. (2008, October 21). The high school dropout's economic ripple effect. *The Wall Street Journal*, p. A13.
- Fitzpatrick, J. J. (2004). Evaluating teaching effectiveness. *Nursing Education Perspective, 25*(3), 109.
- Foley, J. P. (1900). *The Jeffersonian cyclopedia: A comprehensive collection of the views of Thomas Jefferson*. New York: Funk & Wagnalls.
- Formal School Instructor Course. (2008). *Course descriptive data*. Quantico, VA: Department of the Navy.
- Fredland, J. E., & Little, R. D. (1980). Long-term returns to vocational training: Evidence from military sources. *The Journal of Human Resources, 15*(1), 49-66.
- Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education, 75*(3), 203-220.
- Gable, R. K., Wolf, M. B. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings*. Boston, MA: Kluwer.

- Gagne, R. M. (1968). Learning hierarchies. In D. P. Ely, & T. Plomp (Eds.), *Classical writings on instructional technology* (pp. 93-107). Westport, CT : Libraries Unlimited.
- Gagne, R., & Medsker, K. L. (1996). *The conditions of learning: Training applications*. Belmont, CA: Wadsworth Group.
- Gay, L. R., Mills, G. E., & Airasian, R. (2009). *Educational research*. Saddle River, NJ: Pearson Prentice.
- Gilmore, G. J. (2001, August 13). *Recruit attrition rates fall across the services*. Retrieved October 25, 2008, from American Forces Press Service:
<http://www.defenselink.mil/news/newsarticle.aspx?id=44782>.
- Goldberg, M. S., & Warner, J. T. (1987). Military experience, civilian experience, and the earnings of veterans. *The Journal of Human Resources*, 22(1), 62-81.
- Gray, A. M. (1991). *Marine Corps Order 1553.1B, The Marine Corps training and education system*. Washington, DC: Department of the Navy.
- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, 52(11), 1182-1186.

- Gross, R. B., & Small, A. C. (1979). A survey of faculty opinions about student evaluations of instructors. *Teaching of Psychology, 6*(4), 216-219.
- Guthrie, E. R. (1953). The evaluation of teaching. *The American Journal of Nursing, 53*(2), 220-221.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multiple data analysis* (6 ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Hanlon, E. (2003). *Marine Corps Order 1553.2, Management of Marine Corps formal schools and training detachments*. Washington, DC: Department of the Navy.
- Hinkin, T. R. (2005). Scale development principles and practices. In R. A. Swanson and E. F. Holton (Eds.), *Research in organizations: Foundations and methods of inquiry* (pp. 161-179). San Francisco, CA: Berrett-Koehler.
- Holcomb, K. T. (1998). *Marine Corps Order 1580.7D, School interservice training*. Washington, DC: Department of the Navy.
- Holton, E. F. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly, 7*(1), 5-21.

- Holton, E. F., & Burnett, M. F. (2005). The basics of quantitative research. In R. A. Swanson and E. F. Holton (Eds.), *Research in organizations: Foundations and methods of inquiry* (pp. 29-44). San Francisco, CA: Berrett-Koehler.
- Holton, E. F., Bates, R. A., & Ruona, W. E. (2000). Development of a generalized learning transfer inventory. *Human Resource Development Quarterly*, 77(4), 333-360.
- Industry Report (2007). *Training*, 11, 8-24.
- Jackson, D. L., Teal, C. R., Raines, S. J., Nansel, T. R., Force, R. C., & Burdsal, C. A. (1999). The dimensions of students' perceptions of teaching effectiveness. *Educational & Psychological Measurement*, 59(4), 580-597.
- Jonassen, D. H., Hannum, W. H., and Tessmer, M. (1989). *Bloom's taxonomy of educational objectives: Handbook of task analysis procedures*. New York: Praeger.
- Kavanagh, J. (2005). *Determinants of productivity for military personnel: A review of findings on the contribution of experience, training, and aptitude to military performance*. Santa Monica, CA: RAND Corporation.
- Keller, F. (1968). "Good-bye, teacher...". In D. Ely, & T. Plomp (Eds.), *Classical writings on instructional technology* (pp. 183-197). Westport, CT: Libraries Unlimited.

- Kember, D., Leung, D., & Kwan, K. P. (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment & Evaluation in Higher Education*, 27(5), 411-425.
- Kemp, B. W., & Kumar, G. S. (1990). Student evaluations: Are we using them correctly? *Journal of Education for Business*, 66(2), 106-111.
- Kennedy, J. F. (1961). *Presidential inaugural address*. Retrieved October 25, 2008, from Great Speeches of the Twentieth Century:
<http://www.guardian.co.uk/theguardian/2007/apr/22/greatspeeches>
- Kirkpatrick, D. (1979). Techniques for evaluating training programs. In D. P. Ely, & T. Plomp (Eds.), *Classic writings on instructional technology* (pp. 119-142). Westport, CT: Greenwood Publishing.
- Kirkpatrick, D. (1994). *Evaluating training programs: The four levels*. San Francisco: Berrett-Koehler.
- Kleykamp, M. A. (2006). College, jobs, or the military? Enlistment during a time of war. *Social Science Quarterly*, 87(2), 272-290.

- Knapper, C. (2001). Broadening our approach to teaching. *New Directions for Teaching and Learning*, 88, 3-9.
- Knowles, M. (1990). *The adult learner: A neglected species* (4 ed.). Houston, TX: Gulf Publishing.
- Knowles, M. (1995). The adult learner in the technical environment. In L. Kelly (Ed.) *The ASTD technical and skills training handbook* (pp. 211-241). New York: McGraw-Hill.
- Knowles, M., Holton, E., & Swanson, R. (1998). *The adult learner*. Houston, TX: Gulf Publishing.
- Laurance, J. H. (2004). *Performance of the all-volunteer force*. Retrieved October 1, 2008, from Office of the Under Secretary of Defense (Force Management Policy): <http://www.rand.org/pubs/monographs/MG265/images/webS0838.pdf>.
- Leach, J. A. (1996). Distinguishing characteristics among exemplary trainers in business and industry. *Journal of Vocational and Technical Education*, 12(2) (1996).
- Lee, S. H., & Pershing, J. A. (1999). Effective reaction evaluation in evaluating training programs. *Performance Improvement*, 38(8), 32-38.

- Lee, S. H., & Pershing, J. A. (2002). Dimensions and design criteria for developing training reaction evaluations. *Human Resource Development International*, 5(2), 175-197.
- Lippitt, G. L., & This, L. E. (1979). Learning theories and training: Part 1. *Training and Development*, 36(9), 5-17.
- Lowman, J. (1994). Professors as performers and motivators. *College Teaching*, 42(4), 137-142.
- Lundin, R. (1985). *Theories and systems of psychology*. Lexington, MA: D.C. Heath and Company.
- Madsen, K. (2003). The effect of accuracy of instruction, teacher delivery, and student attentiveness on musicians' evaluation of teacher effectiveness. *Journal of Research in Music Education*, 51(1), 38-50.
- Mandebvu, O. W. (1998). *An investigation of the distinguishing characteristics of exemplary trainers in business and industry in Zimbabwe* (Doctoral dissertation). Iowa State University: Iowa. Retrieved from UMI Microform. (UMI Number: 9841067).
- Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76(5), 707-754.

- Marsh, H. W., & Bailey, M. (1993). Multidimensional students' evaluations of teaching effectiveness: A profile analysis. *The Journal of Higher Education, 64*(1), 1-18.
- Marsh, H. W., Fleiner, H., & Thomas, C. S. (1975). Validity and usefulness of student evaluations of instructional quality. *Journal of Educational Psychology, 67*(6), 833-839.
- Mathieu, J. E., Tannenbaum, S. I., & Salas, E. (1992). Influences of individual and situational characteristics of training effectiveness. *Academy of Management Journal, 35*(4), 828-847.
- McBean, E. A., & Sabah Al-Nassri, S. (1982). Questionnaire design for student measurement of teaching effectiveness. *Higher Education, 11*(3), 273-288.
- McDaniel, T. R. (2006). Student evaluations of instructors: A good thing? *Academic Leader, 22*(8), 8.
- McKeachie, W. J. (1979). Student ratings of faculty: A reprise. *Academe: Bulletin of the American Association of University Professors, 65*(6), 384-397.
- McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 51*(11), 1218-1225.
- Milano, M., & Ullius, D. (1998). *Designing powerful training: The sequential-iterative model*. San Francisco, CA: Jossey-Bass.

- Molenda, M. (2003). In search of the elusive ADDIE model. *Performance Improvement, 42*(5), 34-36.
- Moore, C. S., & Reese, D. (2001). *Attrition of Navy recruits from initial skills training*. Alexandria, VA: Center for Naval Analysis.
- Mowrer-Reynolds, E. (2008). Pre-service educator's perceptions of exemplary teachers. *College Student Journal, 42*(1), 214-224.
- Murray, H. G. (2007). Low-inference teaching behaviors and college teaching effectiveness: Recent developments and controversies. In R. P. Perry and J. C. Smart (eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 145-183). The Netherlands: Springer.
- Murray, H. G., Rushton, J. P., & Paunonen, S. V. (1990). Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Educational Psychology, 82*(2), 250-261.
- Nadler, L., & Nader, Z. (1994). HRD in perspective. In W. R. Tracey (Ed.), *Human resources management & development handbook* (2 ed., pp. 19-26). New York, NY: AMACOM.

Nathan, M. J., & Petrosino, A. (2003). Expert blind spot among preservice teachers. *American Educational Research Journal*, 40(4), 905-928.

National Research Council (2003). *Attitudes, aptitudes, and aspirations of American youth: Implications for military recruitment*. Committee on the youth population and military recruiting. Washington, DC: The National Academies Press.

National Training Systems Association (2008). *Executive summary: National defense industrial association*. Retrieved September 2, 2008, from http://www.trainingsystems.org/publications/T2012_2006_10_ExecutiveSummary.pdf.

Newstrom, J. W. (1978). Catch-22: The problems of incomplete evaluation of training. *Training and Development Journal*, 32(11), 22-24.

Office of the Assistant Secretary of Defense for Personnel and Readiness (2008). *2008 population representation in the military services*. Retrieved September 9, 2010 from <http://prhome.defense.gov/MPP/ACCESSION%20POLICY/PopRep2008/summary/chap2.pdf>.

- Office of the Under Secretary of Defense (2006). *2006 Population representation in the military service*. Retrieved October 13, 2008, from http://www.defenselink.mil/prhome/PopRep_FY06/pdf/ExecSum2006.pdf
- Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal, 44*, 113-160.
- Opdenakker, M., & Van Damme, J. (2006). Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice. *Teaching and Teacher Education, 22*, 1-21.
- Paradise, A., & Patel, L. (2009). *Industry report*. Alexandria, VA: American Society for Training and Development.
- Patrick, J., & Smart, R. M. (1998). An empirical evaluation of teacher effectiveness: The emergence of three critical factors. *Assessment & Evaluation in Higher Education, 23*(2), 165-179.

- Perry, R. P., Hall, N. C., & Ruthig, J. C. (2007). Perceived (academic) control and scholastic attainment in higher education. In R. P. Perry, & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 477-551). Memphis, TN: Springer.
- Plant, R. A., & Ryan, R. J. (1994). Who is evaluating training? *Journal of European Industrial Training, 18*(5), 27-30.
- Plax, T. G., Kearney, P., & Downs, T. M. (1986). Communicating control in the classroom and satisfaction with teaching and students. *Communication Education, 35*, 379-388.
- Polk, J. A. (2006). Traits of effective teachers. *Arts Education Policy Review, 107*(4), 23-29.
- Pounder, J. S. (2007). Is student evaluation of teaching worthwhile? An analytical framework for answering the question. *Quality Assurance in Education, 2*, 178-191.
- Powers, B. (1992). *Instructor excellence: Mastering the delivery of training*. San Francisco, CA: Jossey-Bass.
- Powers, R., & Lawler, J. (2007). *ASVAB for dummies*. Hoboken, NJ: Wiley Publishing, Inc.
- Pregent, R. (2000). *Charting your course: How to prepare to teach more effectively*. Madison, WI: Atwood.

- Rae, L. (1993). *Evaluating trainer effectiveness*. Homewood, IL: Business One Irwin.
- Rand, A. (1957). *Atlas shrugged*. New York: Dutton.
- Raskin, B. L., & Plante, P. R. (1979). The student devaluation of teachers. *Academe: Bulletin of the American Association of University Professors*, 65(6), 381-383.
- Richardson, J. T. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment & Evaluation in Higher Education*, 30(4), 387-415.
- Richmond, V. P., Gorham, J. S., & McCroskey, J. C. (1987). The relationship between selected immediacy behaviors and cognitive learning. In M. A. McLaughlin (Ed.), *Communication yearbook 70* (pp. 574-590). Newbury Park, CA: Sage.
- Rodin, M., & Rodin, B. (1972). Student evaluations of teachers. *Science*, 177, 1164-1166.
- Rosenberg, M. J. (1982). The ABCs of ISD. *Training and Development*, 36(9), 44-50.
- Ryan, J. J., Anderson, J. A., & Birchler, A. B. (1980). Student evaluation: The faculty responds. *Research in Higher Education*, 12(4), 317-333.

- Sanders, R. E. (1998). Questioning techniques. In M. W. Galbraith (Ed.), *Adult learning methods* (pp. 187-196). Malabar, FL: Krieger.
- Schmoker, M. J. (2006). *Results now: How we can achieve unprecedented improvements in teaching and learning*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Scribner, B. D., Smith, A., Baldwin, R., & Phillips, R. L. (1986). Are smarter tankers better? AFQT and military productivity. *U.S. Armed Forces and Society*, 12, 193-206.
- Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: Love me, love my lectures? *Assessment & Evaluation in Higher Education*, 25(4), 397-405.
- Simpson, M. D. (2008, October 2). Students grading teachers. *National Education Association*, p. 20.
- Simpson, P. M., & Siguaw, J. A. (2000). Student evaluations of teaching: An exploratory study of the faculty response. *Journal of Marketing Education*, 22(3), 199-213.
- Simpson, R. D. (1995). Uses and misuses of student evaluation of teaching effectiveness. *Innovative Higher Learning*, 20(1), 3-5.

- Small, A. C., Hollenbeck, A. R., & Haley, R. L. (1982). The effect of emotional state on student ratings of instructors. *Teaching of Psychology, 9*(4), 205-211.
- Spector, P. E. (1992). *Summated rating scale construction*. Newbury, CA: Sage.
- Spooren, P., & Mortelmans, D. (2006). Teacher professionalism and student evaluation of teaching: Will better teachers receive higher ratings and will better students give higher ratings? *Educational Studies, 32*(2), 201-214.
- Sproull, N. L. (2002). *Handbook of research methods* (2 ed.). Lanham, MD: Scarecrow Press, Inc.
- Stein, L. (2000). Combat readiness through education. *Marine Corps Gazette, 94*(10), 53-56.
- Swanson, R. A., & Holton, E. F. (2001). *Foundations of human resource development*. San Francisco, CA: Berrett-Koehler.
- Theall, M., & Franklin, J. (2001). New directions for institutional research. *New Directions for Institutional Research, 109*, 45-56.
- Thompson, K. D. (2001). *Adult educator effectiveness within the training context: A study of trainee perceptions of effective trainer characteristics*. (Doctoral dissertation). University of Wyoming, Laramie, WY.

- Thorndike, E. L. (1910). The contribution of psychology to education. *Journal of Educational Psychology*, 1(1), 5-12.
- Thorndike, E. L. (1933). The influence of irrelevant rewards. *Journal of Educational Psychology*, 24(1), 1-15.
- Trout, P. (2000). Teacher evaluations. *Commonweal*, 127(8), 10-11.
- Turman, P. D., & Schrod, P. (2006). Student perceptions of teacher power as a function of perceived teacher confirmation. *Communication Education*, 55(3), 265-279.
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. Chicago, IL: University of Chicago Press.
- U.S. Department of Education. (1998). *Educational and labor market performance of GED recipients*. Retrieved September 30, 2008, from <http://www.ed.gov/pubs/GED/gedusaf.html>.
- U.S. Government Accounting Office (2005). *Military personnel: Financial costs and loss of critical skills due to DOD's homosexual conduct policy cannot be completely estimated*, GAO-05-299. Retrieved September 9, 2010, from <http://www.gao.gov/cqi-bin/getrpt?GAO-05-299>.

- U.S. Office of Management and Budget. (2005a). *Detailed information on the Department of Defense training and education programs - accession training*. Retrieved October 13, 2008, from ExpectMore.gov:
<http://www.whitehouse.gov/omb/expectmore/summary/10003>.
- U.S. Office of Management and Budget. (2005b). *Detailed information on the Department of Defense training and education programs - basic skills and advanced training assessment*. Retrieved October 13, 2008, from ExpectMore.gov: <http://www.whitehouse.gov/o>.
- vanDriel, J. H., Beijaard, D., & Verloop, N. (2001). Professional development and reform in science education: The role of teachers' practical knowledge. *Journal of Research in Science Teaching, 38*(2), 137-158.
- Wanzer, M. B., Frymier, A. B., Wojtaszczyk, A. M., & Smith, T. (2006). Appropriate and inappropriate uses of humor by teachers. *Communication Education, 55*(2), 178-196.
- Watkins, S. J., & Sherk, J. (2008). *Who serves in the U.S. military? Demographic characteristics of enlisted troops and officers*. Washington, DC: Heritage Foundation.
- White, E. E. (1886). *Elements of pedagogy*. Cincinnati, OH: American Book Company.

- Wilson, J. H. (2006). Predicting student attitudes and grades from perceptions of instructors' attitudes. *Teaching of Psychology, 33*(2), 91-95.
- Wolfowitz, P. (2004). *DOD Directive 1322.18, Military personnel assignments*. Washington, DC: Department of Defense.
- Wolfowitz, P. (2007). *DOD Directive 1315.07, Military training*. Washington, DC: Department of Defense.
- Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment, 28*(3), p. 186-191.
- Wortruba, T. R., & Wright, P. L. (1975). How to develop a teacher-rating instrument: A research approach. *The Journal of Higher Education, 46*(6), 653-663.
- Wulff, D., Staton-Spicer, A., Hess, C., & Nyquist, J. (1985). *The student perspective on evaluating teaching effectiveness*. Retrieved September 2008, 9, from Association for Communication Administration Bulletin.
- Yang, B. (2005). Factor analysis methods. In R. A. Swanson & E. F. Holton (Eds.), *Research in organizations: Foundations and methods of inquiry*, (pp. 181-199). San Francisco, CA: Berrett-Koehler.

Young, S., & Shaw, D. G. (1999). Profiles of effective college and university teachers. *The Journal of Higher Education*, 70(6), 670-686.

APPENDICES

Appendix A

Item Deletion List

Motivation for discarding item	Qty
• These items were discarded during original review as duplicative author and/or study	310
• These items were discarded in later review as duplicate by the same author	9
• These items were discarded because in most cases developmental factors course design, structure, objectives and level of difficulty are outside instructor's control	162
• These items were discarded because they failed to identify a single measurable item (double barreled).	145
• These items were discarded because in most cases instructional materials, texts, examination and assignments are outside instructor's control	86
• These items were discarded due to relevance and/or application to the study's military training environment	100
• These items were discarded as incomplete	53
• These items were discarded for polarity; i.e., another item is opposite and retained	41
• These items were discarded for lacking definition and/or clarity	39
• These items were discarded as redundant to other items; although, not exact enough to combine	31
• These items were discarded because all lesson materials are entry-level, performance based standardized instruction to perform a specific task that disallows free expression and deviation from published training plans	24

- These items were discarded because instructor may not be afforded additional time to provide individualized training outside of published schedule; therefore, student may not be able to determine if instructor or course is responsible 11
- These items were discarded for lack of specificity (related to specific lesson material vice entire course or within or outside of classroom) 9
- These items were discarded as the student may not be qualified to answer for entire student body. 9
- These items were discarded as stereotypical item; not gender neutral 7
- These items were discarded as facilities and training areas are outside instructor's control 6
- These items were discarded at face value (2nd Round) 49
- These items were discarded at face value (3rd Round) 52

Appendix B

Initial Factors, Sub-factors and Items

- Factor 1 - Effective Delivery: The instructor effectively delivered the lesson material.
- Sub-factor 1.1 The instructor's teaching style captured and held the student's interest
- Item 1 The instructor presented the lesson in an interesting way
 - Item 2 The instructor maintained student attention as the lesson progressed
 - Item 3 The instructor used threats to make the students listen
- Sub-factor 1.2 The instructor was able to stimulate interest in subject material
- Item 4 The instructor introduced stimulating ideas about the subject material
 - Item 5 The instructor failed to stimulate interest in the subject matter
- Sub-factor 1.3 The instructor employed training techniques designed to enhance student learning
- Item 6 The instructor gave preliminary overview of the lesson
 - Item 7 The instructor signaled the transition from one topic to the next during the lesson
 - Item 8 The instructor summarized periodically during the lesson
 - Item 9 The instructor repeated difficult ideas several times
 - Item 10 The instructor stressed the important points of the lesson
 - Item 11 The instructor provided sample exam questions
- Sub-factor 1.4 The instructor exhibited distracting lecture techniques

- Item 12 The instructor taught by reading directly from their notes
 - Item 13 The instructor taught from behind the podium or desk
 - Item 14 The instructor had a tense body position when talking to the class
 - Item 15 The instructor played with chalk or pointer while teaching
 - Item 16 The instructor talked to the board (dry-erase, chalk, etc.) with their back to the class while teaching
- Sub-factor 1.5 The instructor was well prepared to present lesson material
- Item 17 The instructor was well prepared for the lesson
 - Item 18 The instructor showed a lack of planning for classroom work
 - Item 19 The instructor ensured that training equipment and/or training aids were set-up in advance
- Sub-factor 1.6 The instructor used technology and training aids to enhance the lesson
- Item 20 The instructor used technology comfortably while teaching
 - Item 21 The instructor used computerized electronic media that enhanced the learning experience
 - Item 22 The instructor used non-computerized training aids that enhanced the lesson
- Sub-factor 1.7 The instructor used board (chalk, dry erase, etc.) to enhance the lesson
- Item 23 The instructor used the board (dry-erase, chalk, etc.) in a legible way that enhanced the lesson
- Sub-factor 1.8 The instructor used appropriate examples to enhance the lesson
- Item 24 The instructor used examples to get points across during the lesson
 - Item 25 The instructor used concrete, everyday examples to explain the lesson

Factor 2 - Learning Environment: The instructor created learning environment that was conducive to learning.

Sub-factor 2.1 The instructor maintained a safe learning environment that was conducive to learning

- Item 26 The instructor maintained a classroom atmosphere conducive to learning
- Item 27 The instructor ensured the learning environment was safe

Sub-factor 2.2 The instructor was punctual and used time efficiently

- Item 28 The instructor was punctual in starting the lesson
- Item 29 The instructor made effective use of lesson time
- Item 30 The instructor effectively paced the lesson delivery
- Item 31 The instructor allowed lesson discussions to wander from subject

Factor 3 - Enjoys Teaching: The instructor appeared to enjoy teaching and was well prepared.

Sub-factor 3.1 The instructor appeared interested in training students

- Item 32 The instructor was enthusiastic about training students
- Item 33 The instructor was not interested in training students

Sub-factor 3.2 The instructor appeared to enjoy training students

- Item 34 The instructor seemed to genuinely enjoy training students
- Item 35 The instructor enhanced the lesson with the use of humor

Factor 4 - Clear Communication: The instructor clearly and effectively communicated.

- Sub-factor 4.1 The instructor is an effective communicator
 - Item 36 The instructor communicated effectively throughout the lesson
 - Item 37 The instructor spoke clearly during the lesson
 - Item 38 The instructor used distracting words like "um" or "ah" during the lesson
 - Item 39 The instructor used inappropriate language (cursed, swore, etc.) that detracted from the lesson
- Sub-factor 4.2 The instructor clearly communicated student goals
 - Item 40 The instructor clearly defined lesson objectives
 - Item 41 The instructor gave a clear idea of the student requirements
 - Item 42 The instructor gave clear instructions concerning lesson assignments
 - Item 43 The instructor clearly demonstrated how the work should be approached
 - Item 44 The instructor made the aims of the assessment clear
- Sub-factor 4.3 The instructor's presentation clarified lesson material
 - Item 45 The instructor explained lesson material clearly
 - Item 46 The instructor had a method of teaching that helped me understand the lesson material
 - Item 47 The instructor presented the background of concepts that helped me understand the lesson material
 - Item 48 The instructor used explanations that confused me
- Sub-factor 4.4 The instructor used language that the students could understand
 - Item 49 The instructor used words that students could understand
 - Item 50 The instructor failed to define new or unfamiliar terms

Factor 5 - Subject Expertise: The instructor demonstrated subject matter expertise.

Sub-factor 5.1 The instructor introduced practical, current, real world application of subject material

- Item 51 The instructor demonstrated the importance of the subject matter to future work
- Item 52 The instructor pointed out practical application of lesson material
- Item 53 The instructor related lesson material to real life situations
- Item 54 The instructor described relevant personal experience to enhance the lesson

Sub-factor 5.2 The instructor integrated student knowledge into lesson

- Item 55 The instructor related content of this lesson to other subject matter that students already understand
- Item 56 The instructor integrated student experiences to illustrate specific lesson points

Sub-factor 5.3 The instructor is a subject matter expert on material

- Item 57 The instructor had a background of subject matter experience that enhanced the lesson
- Item 58 - The instructor demonstrated an exceptional knowledge of the subject material

Factor 6 - Provides Feedback: The instructor provided feedback.

Sub-factor 6.1 The instructor provided feedback to student

- Item 59 The instructor provided feedback on my individual work through comments on papers, oral discussions, etc.
- Item 60 The instructor returned class work in time to benefit my progress
- Item 61 The instructor did not review in a way that students could understand their weaknesses

Sub-factor 6.2 The instructor used questions to solicit student feedback

- Item 62 The instructor used questions to test for knowledge, skills and attitudes
 - Item 63 The instructor called on students to answer questions even if they have not indicated that they want to talk
- Sub-factor 6.3 The instructor redirects questions back to student
- Item 64 The instructor encouraged students to think out answers
 - Item 65 The instructor referred questions back to students
 - Item 66 The instructor publicly ridiculed and criticized students who responded incorrectly
- Sub-factor 6.4 The instructor's response to questions
- Item 67 The instructor answered students' questions in a way that helped students understand
 - Item 68 The instructor sticks to the point in answering questions
 - Item 69 The instructor continued to answer questions until points of confusion were cleared up
 - Item 70 The instructor answered student questions in an incomplete or inaccurate way
 - Item 71 The instructor became angry when questions were asked
 - Item 72 The instructor when unable to answer student's questions researched answers and reported results back to students
- Sub-factor 6.5 The instructor listened to students
- Item 73 The instructor listened to students
 - Item 74 The instructor interrupted speaking student without justification
- Sub-factor 6.6 The instructor maintained eye contact with students
- Item 75 The instructor maintained eye contact with students while listening
 - Item 76 The instructor looked at the class when talking

Factor 7 - Concern for Learning: The instructor appeared to be concerned about the student's progress and learning.

Sub-factor 7.1 The instructor appeared to be concerned about the student's progress and learning

- Item 77 The instructor referred to students by name
- Item 78 The instructor gave personal attention to student work
- Item 79 The instructor was concerned with whether the students learned the lesson material
- Item 80 The instructor lacked interest in student's learning

Sub-factor 7.2 The instructor was able to motivate the students to want to learn

- Item 81 The instructor had the ability to motivate you to learn the lesson material
- Item 82 The instructor motivated me to think rather than just memorize material

Sub-factor 7.3 The instructor was able to encourage students

- Item 83 The instructor encouraged students to actively participate in the lesson activities
- Item 84 The instructor did not inspire confidence in individual student performance

Sub-factor 7.4 The instructor seemed to know when students needed help with lesson material

- Item 85 The instructor knew when students didn't understand the lesson material
- Item 86 The instructor suggested specific study skills to improve my performance in this lesson

Factor 8 - Fair and Respectful: The instructor demonstrated a positive, respectful and fair attitude toward students.

Sub-factor 8.1 The instructor demonstrated a positive attitude toward students

- Item 87 The instructor had a positive attitude about training students

- Item 88 The instructor had a cynical attitude that repelled students
- Sub-factor 8.2 The instructor seemed to respect the student as an individual
 - Item 89 The instructor seemed to genuinely respect students as individuals
 - Item 90 The instructor made me feel that I could not learn the lesson material
 - Item 91 The instructor was always arguing with students
- Sub-factor 8.3 The instructor demonstrated fairness and impartiality
 - Item 92 The instructor treated all students equally
 - Item 93 The instructor did not evaluate student work fairly
- Factor 9 - Available and Relevant: The instructor appeared available outside of class and the relevance of lesson material
 - Sub-factor 9.1 The student perceived that the instructor was available out of formal classroom
 - Item 94 The instructor was available outside of the formal class period to explain or clarify lesson material.
 - Sub-factor 9.2 The student perceived the lesson material as relevant
 - Item 95 The instructor has been instrumental in increasing my knowledge of the lesson's subject material
 - Item 96 How would you rate the overall value of this lesson as taught by this instructor?
 - Sub-factor 9.3 Student's opinion of overall instructor performance
 - Item 97 Compared with all instructors I have had, both in high school and since, this instructor was: One of the Best-One of the Worst
 - Item 98 Compared with all instructors I have had, both in high school and since, this instructor was not qualified to be an instructor

Appendix C

Script for Data Collection

BEFORE SESSION

LOCATION: Formal Military Occupational Specialty Training Center Classroom aboard U.S. Marine Corps Base, Camp Johnson, North Carolina.

ACTION-> The investigator has made prior arrangements with school administrators and lesson instructors to solicit volunteer participants. Classes were selected based on three conditions: 1) only one instructor has taught the lesson, (2) the same instructor has taught all of the testable lesson material, and 3) a forthcoming test will measure student mastery of only the lesson material taught by this instructor.

ACTION-> The investigator will arrive with the investigator's script (to be used by the investigator), survey instrument packages (to be completed by volunteer participants), and additional participant consent instruction forms (to be provided to participants upon request). Distribution and accountability of all survey instrument packages will be closely monitored prior to- and after the session.

Two pages make up a survey package: 1) an Informed Consent Information & Signature Sheet (single-sided); and 2) Survey Instrument (double-sided with demographic information, instructions and participant response statements).

The instructor will introduce the investigator and provide participant instructions for actions after the session. The instructor will then depart the session. The investigator will locate and notify the instructor upon completion of the session.

IN CLASSROOM

READ ALOUD-> Good Morning/Afternoon, my name is Dan Lathers and I am a graduate student at North Carolina State University.

As you all may know, the Commanding Officer is always seeking to improve the training provided to Marines. One of the best ways to do that is to ask you - the students.

We're interested in what you are experiencing during your training aboard Camp Johnson and your responses may affect future formal training.

One of the most productive methods of gaining information is by using a questionnaire – which is what we will be using today.

First and foremost, participation is completely voluntary and will not affect your grade or class standing. Because responses from all participants will be collected and tallied together with other classes; responses will be not be linked in reports to any one particular individual or name.

Is everyone willing to participate?

Those of you who would like to participate please stay seated; everyone else - please exit the room at this time.

ACTION-> WAIT FOR NON-PARTICIPANTS TO LEAVE, IF ANY

READ ALOUD-> Thank you all for agreeing to participate. Let's get started. Please take a package and a pencil and pass the rest to the student behind you. Please wait before reading until told to do so.

ACTION-> DISTRIBUTE PACKAGES AND PENCILS

ACTION-> START PENCIL BOX WITH FIRST PARTICIPANT WITH INSTRUCTIONS TO TAKE ONE AND PASS ON

ACTION-> DISTRIBUTE PACKAGES (INCLUDES AN INFORMED CONSENT COVER SHEET TO BE SIGNED AND THE DOUBLE-SIDED SURVEY INSTRUMENT) TO FIRST STUDENT IN EACH ROW

ACTION-> COMPLETE ALL ROWS WITH SINGLE PACKAGE TO EACH PARTICIPANT

ACTION-> STOW ANY UNUSED PACKAGES AND PENCILS IN TRANSPORT CASE

READ ALOUD-> Does everyone have a survey and a pencil?

ACTION-> IF NOT - PROVIDE SURVEY or PENCIL; IF SO – PROVIDE INSTRUCTIONS

READ ALOUD-> If you would look at the top sheet you'll see an Informed Consent Release Form. This form describes your rights and also allows me to collect and use your responses. Please read along as I read aloud.

READ ALOUD-> INFORMED CONSENT FORM

The title of the study is: The Evaluation of U.S. Marine Corps Vocational Training: Creating an Instrument to Assess Effective Instructional Techniques and Attributes

I am the principal investigator and Dr. James Bartlett from NC State University is my faculty sponsor

Research studies are used to gain a better understanding of a certain topic or issue. You are being asked to take part in a research study.

Research studies may sometimes pose risks to those that participate; therefore, your signed consent is required if you wish to participate.

In this consent form you will find specific details about the research in which you are being asked to participate.

If you do not understand something in this form it is your right to ask the researcher for clarification or more information.

A copy of this consent form will be provided to you to keep.

Your participation in this study is voluntary.

You have the right to be a part of this study, to choose not to participate or to stop participating at any time without penalty.

If at any time you have questions about your participation, do not hesitate to contact me (the researcher(s) named above).

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> The primary purpose of this study is to develop an instrument to assess instructional techniques and attributes based on research findings from multiple training disciplines and educational settings for application to U.S. Marine Corps vocational specialty instructors.

If you agree to participate, there will be two parts of this study.

The first part will ask you to respond to statements concerning the vocational training you are receiving in a two-page questionnaire.

For the second part of the study, I will collect your end-of-lesson performance test scores. Once data collection is complete individual questionnaire responses and

associated test scores from all participants will be combined to determine statistically significant relationships.

READ ALOUD – ELABORATE -> This study is only concerned with the lesson you are currently learning.

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> There are some potential benefits in this study.

By identifying redundancies across a wide variety of studies, this study is designed to identify consistent themes that can be used to create a practical “trainee-reaction” instrument to capture and reflect specific educational processes.

Once identified these themes can then be used to focus and enhance trainer developmental programs while increasing student satisfaction and overall performance.

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> You will not be compensated for your participation. There are no direct benefits for being in this study and you will not receive anything for participating.

There are no potential risks identified in connection with this study and I will follow every practical and ethical consideration to maintain participant’s confidentiality.

READ ALOUD – ELABORATE -> Confidentiality is defined as: ensuring the information collected is only accessible to those authorized to have access.

Only two people are authorized to have access to this study’s raw data.

In this case only my faculty advisor (Dr. Bartlett) and I will have access to your individual responses – not your instructors, officers or other participants.

Also, only the two of us will be able to link your individual responses to your individual test scores by your student ID#.

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> Your student ID# will only be used to link your test scores to your questionnaire responses for transfer to a master working data file.

You will be asked to write your student ID# on the consent form.

Once individual data is transferred, all completed questionnaires and test scores will be secured in a locked file cabinet separate from the master working data file.

Keeping participant data secured and separate will ensure that the information collected in this study will be kept confidential.

Only the master working data file will be used for data analysis; therefore, there is no way responses or scores could be linked to any one individual.

Also, no reference will be made in oral or written reports which could link your individual responses or test scores to this study and o.

READ ALOUD – ELABORATE -> Again – only two people – Dr. Bartlett and I – not your instructors, officers or other participants - will not be able to link your responses to your test scores.

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> If you have questions at any time about the study or the procedures, you may contact me at jdlathe@ncsu.edu or 910/934-4332.

If you have questions about your rights as a research participant or feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Deb Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919/515-4514).

I will provide you with a copy of this informed consent form after this session so you will have this contact information with you.

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> If there are no questions, please read, sign and date in the box at the bottom of the page.

You'll also see a place to write your student ID#.

Please fill that in at this time.

I will keep the signed forms in a separate place than the collected responses but I will have a stack of unsigned consent forms for any participant to take after this session.

Has everyone signed the informed consent block?

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> Alright, please turn to the next page where you'll see a box at the top with a few questions asking you to describe yourself.

Please do not begin responding to the statements outside the top box until told to do so.

Is everyone there?

If so, please do not read ahead but follow along and fill in the blanks or check in the boxes when asked to do so.

In the first space provided – write your student ID#.

Next, fill-in the year you were born, check your gender and the state you enlisted from.

Next, check your ethnicity and finally check the highest educational level you have completed.

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> Has everyone completed describing yourselves? If so, then please follow along as I read the instructions aloud. Please do not begin the survey until told to do so.

The following statements are designed to collect information on your reaction to this lesson and this instructor only.

There are no “right” or “wrong” answers to these statements. Instead, please indicate on the scale provided how much you agree or disagree with the statement based on your experience with this lesson and this instructor.

You will need to read each numbered statement and choose the response number that best indicates how much you agree or disagree with the statement.

When you mark your response you will need to find the number of the response that best describes how much you agree or disagree with the statement. Mark your response by completely filling in the circle like the small box in the instructions.

ACTION-> REFER TO BUBBLE SCALE DIAGRAM DESCRIBE CORRECT AND INCORRECT METHODS OF RESPONDING - Correct: ①●③④⑤⑥ Incorrect ①②③④⑤✘

READ ALOUD-> Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> The following categories are available for you select from.

1 - STRONGLY DISAGREE. You would disagree most of the time.

2 - MODERATELY DISAGREE. You would frequently disagree.

3 - SLIGHTLY DISAGREE. You would occasionally disagree.

4 - SLIGHTLY AGREE. You would occasionally agree.

5 - MODERATELY AGREE. You would frequently agree.

6 - STRONGLY AGREE. You would agree most of the time.”

If you have difficulty responding, select the single response that seems to fit best with your experience and move on.

Please do not think too long about the statement as your first reactions will usually be your most accurate response.

Please select only ONE response to every statement and please respond to EVERY statement.

Please reply only to statements as they relate to your current instructor during this phase of training.

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> After this session, I will be conducting other sessions with other classes throughout the coming weeks, so please do not discuss your participation in this session.

It is very important to not discuss any of the actual statements, the type of statements or your actual responses with other participants, fellow students or instructors because it could contaminate the overall study.

Are there any questions?

ACTION-> IF YES, ANSWER QUESTIONS; IF NOT, CONTINUE

READ ALOUD-> You will have up to 45 minutes to respond to all of the statements.

Once you have completed the survey please bring up to me.

I will a copy of the informed consent instruction sheet for you at the time.

Again please only reply to questions as they relate to your current lesson and instructor.

Please begin at this time.

ACTION-> MONITOR PROGRESS

ACTION-> COLLECT SURVEYS COMPLETED

ACTION-> ACCOUNT FOR ALL SURVEY PACKAGES

ACTION-> REMOVE SIGNED CONSENT FORMS AND STACK IN ONE PILE AND SURVEY INSTRUMENTS IN SECOND PILE ON OPPOSITES CORNERS OF THE DESK

ACTION-> VERIFY INFORMED CONSENT SIGNATURE

ACTION-> PROVIDE BOX TO COLLECT PENCILS AS COMPLETE

ACTION-> PROVIDE CONSENT FORM TO STUDENTS WHO REQUEST

ACTION-> THANK VOLUNTEERS FOR PARTICIPATING AND ANSWER ANY QUESTIONS WHERE APPROPRIATE

ACTION-> LOCK SIGNED CONSENT FORMS AND SURVEY INSTRUMENTS IN LOCKABLE FILE CASE FOR TRANSPORT

ACTION-> CONTACT AND THANK INSTRUCTOR WHEN SESSION IS COMPLETE

ACTION-> COORDINATE RECEIPT OF CLASS ROSTER IDENTIFYING PARTICIPANT TEST SCORES BY STUDENT ID#

Appendix D

Institutional Review Board Approval

North Carolina State University is a land-grant university and a constituent institution of The University of North Carolina

**Office of Research
and Graduate Studies
Division of Research Administration**

NC STATE UNIVERSITY

Sponsored Programs and
Regulatory Compliance
Campus Box 7514
2701 Sullivan Drive
Raleigh, NC 27695-7514

919.515.2444
919.515.7721 (fax)

From: Debra Paxton, IRB Administrator
North Carolina State University
Institutional Review Board

Date: April 21, 2010

Project Title: The evaluation of U.S. Marine Corps vocational training: Creating an instrument to assess effective instructional techniques and attributes

IRB#: 1397-10-4

Dear Mr. Lathers,

The project listed above has been reviewed by the NC State Institutional Review Board for the Use of Human Subjects in Research, and is approved for one year. **This protocol will expire on April 5, 2011 and will need continuing review before that date.**

NOTE:

1. You must use the attached consent forms which have the approval and expiration dates of your study.
2. This board complies with requirements found in Title 45 part 46 of The Code of Federal Regulations. For NCSU the Assurance Number is: FWA00003429.
3. Any changes to the protocol and supporting documents must be submitted and approved by the IRB prior to implementation.
4. If any unanticipated problems occur, they must be reported to the IRB office within 5 business days by completing and submitting the unanticipated problem form on the IRB website.
5. Your approval for this study lasts for one year from the review date. If your study extends beyond that time, including data analysis, you must obtain continuing review from the IRB.

Sincerely,



Debra Paxton
NC State IRB

Appendix E

Informed Consent to Release Form for Research

INFORMED CONSENT FORM for RESEARCH - North Carolina State University

This consent form is valid April 10, 2010 through April 11, 2011

Title of Study: The Evaluation of U.S. Marine Corps Vocational Training: Creating an Instrument to Assess Effective Instructional Techniques and Attributes

Principal Investigator Mr. Dan Lathers ~ Faculty Sponsor Dr. James Bartlett

Research studies are used to gain a better understanding of a certain topic or issue. You are being asked to take part in a research study. Research studies may sometimes pose risks to those that participate; therefore, **your signed consent is required** if you wish to participate. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something in this form it is your right to ask the researcher for clarification or more information. A copy of this consent form will be provided to you to keep. Your participation in this study is **voluntary**. You have the right to be a part of this study, to choose not to participate or to stop participating at any time without penalty. If at any time you have questions about your participation, do not hesitate to contact the researcher(s) named above.

The primary purpose of this study is to develop an instrument to assess instructional techniques and attributes based on research findings from multiple training disciplines and educational settings for application to U.S. Marine Corps vocational specialty instructors.

If you agree to participate, there will be two parts of this study. The first part will ask you to respond to statements concerning the vocational training you are receiving in a two-page questionnaire. For the second part of the study, I will collect your end-of-lesson performance test scores. Once data collection is complete individual questionnaire responses and associated test scores from all participants will be combined to determine statistically significant relationships. **This study is only concerned with the lesson you are currently learning.**

Benefits. While there is no direct benefit to you from participation, by identifying redundancies across a wide variety of studies, this study is designed to identify consistent themes that can be used to create a practical "trainee-reaction" instrument to capture and reflect specific educational processes. Once identified these themes can then be used to focus and enhance trainer developmental programs while increasing student satisfaction and overall performance.

Compensation. You will not receive anything for participating.

Risks. There are no potential risks identified in connection with this study.

Confidentiality. Your student ID# will only be used to link your test scores to your questionnaire responses for transfer to a master working data file. Once individual data is transferred, all completed questionnaires and test scores will be secured in a locked file cabinet separate from the master working data file. Keeping participant data secured and separate will ensure that the information collected in this study will be kept confidential. Only the master working data file will be used for data analysis; therefore, there is no way responses or scores could be linked to any one individual. Also, no reference will be made in oral or written reports which could link your individual responses or test scores to this study.

If you have questions at any time about the study or the procedures, you may contact the researcher, Mr. Dan Lathers, at jdlathe@ncsu.edu or 910/934-4332. If you have questions about your rights as a research participant or feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Deb Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919/515-4514).

Consent to Participate: "I have read and understand the Informed Consent information. I understand I will receive a copy of the information provided. I agree to participate in this study with the understanding that I may choose not to participate or to stop participating at any time without penalty or loss of benefits to which I am otherwise entitled. I understand that there are no direct benefits for participating and that I will not receive anything for being in this study."

Participant's Signature _____ Date _____ Student ID# _____

Investigator's Signature _____ Date _____ Survey# _____

Appendix F

Initial Instrument

QUESTIONNAIRE – STUDENT REACTION TO FORMAL INSTRUCTION			
Student ID#: _____	Birth Year: 19____	Male <input type="checkbox"/> Female <input type="checkbox"/>	State Enlisted: _____
Ethnicity: American Indian/Alaska Native <input type="checkbox"/> Asian <input type="checkbox"/> Black <input type="checkbox"/> Hispanic <input type="checkbox"/> Pacific/Hawaiian Islander <input type="checkbox"/> White <input type="checkbox"/> Other <input type="checkbox"/> (_____)			
Highest Education Level? GED <input type="checkbox"/> HS Diploma <input type="checkbox"/> 2-yr Degree <input type="checkbox"/> 4-yr Degree <input type="checkbox"/> 6-yr or More <input type="checkbox"/>			
<p>How to Mark Your Responses: Find the number of the response that best describes how much you agree or disagree with the statement. Mark your response by filling in the circle. Please only reply to statements as they relate to your current lesson and instructor.</p> <p>1 - STRONGLY DISAGREE. You would <u>disagree most of the time</u>.</p> <p>2 - MODERATELY DISAGREE. You would <u>frequently disagree</u>.</p> <p>3 - SLIGHTLY DISAGREE. You would <u>occasionally disagree</u>.</p> <p>4 - SLIGHTLY AGREE. You would <u>occasionally agree</u>.</p> <p>5 - MODERATELY AGREE. You would <u>frequently agree</u>.</p> <p>6 - STRONGLY AGREE. You would <u>agree most of the time</u>.</p>			
			Correct: ① ● ③ ④ ⑤ ⑥ Incorrect: ② ④ ⑤ ⑥
The Instructor:			
1. ① ② ③ ④ ⑤ ⑥	Presented the lesson in an interesting way		
2. ① ② ③ ④ ⑤ ⑥	Maintained student attention as the lesson progressed		
3. ① ② ③ ④ ⑤ ⑥	Used threats to make the students listen		
4. ① ② ③ ④ ⑤ ⑥	Introduced stimulating ideas about the subject material		
5. ① ② ③ ④ ⑤ ⑥	Failed to stimulate interest in the subject matter		
6. ① ② ③ ④ ⑤ ⑥	Gave preliminary overview of the lesson		
7. ① ② ③ ④ ⑤ ⑥	Signaled the transition from one topic to the next during the lesson		
8. ① ② ③ ④ ⑤ ⑥	Summarized periodically during the lesson		
9. ① ② ③ ④ ⑤ ⑥	Repeated difficult ideas several times		
10. ① ② ③ ④ ⑤ ⑥	Stressed the important points of the lesson		
11. ① ② ③ ④ ⑤ ⑥	Provided sample exam questions		
12. ① ② ③ ④ ⑤ ⑥	Taught by reading directly from their notes		
13. ① ② ③ ④ ⑤ ⑥	Taught from behind the podium or desk		
14. ① ② ③ ④ ⑤ ⑥	Had a tense body position when talking to the class		
15. ① ② ③ ④ ⑤ ⑥	Played with chalk or pointer while teaching		
16. ① ② ③ ④ ⑤ ⑥	Talked to the board (dry-erase, chalk, etc.) with their back to the class while teaching		
17. ① ② ③ ④ ⑤ ⑥	Was well prepared for the lesson		
18. ① ② ③ ④ ⑤ ⑥	Showed a lack of planning for classroom work		
19. ① ② ③ ④ ⑤ ⑥	Ensured that training equipment and/or training aids were set-up in advance		
20. ① ② ③ ④ ⑤ ⑥	Used technology comfortably while teaching		
21. ① ② ③ ④ ⑤ ⑥	Used computerized electronic media that enhanced the learning experience		
22. ① ② ③ ④ ⑤ ⑥	Used non-computerized training aids that enhanced the lesson		
23. ① ② ③ ④ ⑤ ⑥	Used the board (dry-erase, chalk, etc.) in a legible way that enhanced the lesson		
24. ① ② ③ ④ ⑤ ⑥	Used examples to get points across during the lesson		
25. ① ② ③ ④ ⑤ ⑥	Used concrete, everyday examples to explain the lesson		
26. ① ② ③ ④ ⑤ ⑥	Maintained a classroom atmosphere conducive to learning		
27. ① ② ③ ④ ⑤ ⑥	Ensured the learning environment was safe		
28. ① ② ③ ④ ⑤ ⑥	Was punctual in starting the lesson		
29. ① ② ③ ④ ⑤ ⑥	Made effective use of lesson time		
30. ① ② ③ ④ ⑤ ⑥	Effectively paced the lesson delivery		
31. ① ② ③ ④ ⑤ ⑥	Allowed lesson discussions to wander from subject		
32. ① ② ③ ④ ⑤ ⑥	Was enthusiastic about training students		
33. ① ② ③ ④ ⑤ ⑥	Was <u>not</u> interested in training students		
34. ① ② ③ ④ ⑤ ⑥	Seemed to genuinely enjoy training students		
35. ① ② ③ ④ ⑤ ⑥	Enhanced the lesson with the use of humor		
36. ① ② ③ ④ ⑤ ⑥	Communicated effectively throughout the lesson		
37. ① ② ③ ④ ⑤ ⑥	Spoke clearly during the lesson		
38. ① ② ③ ④ ⑤ ⑥	Used distracting words like "um" or "ah" during the lesson		
39. ① ② ③ ④ ⑤ ⑥	Used inappropriate language (cursed, swore, etc.) that detracted from the lesson		
40. ① ② ③ ④ ⑤ ⑥	Clearly defined lesson objectives		
41. ① ② ③ ④ ⑤ ⑥	Gave a clear idea of the student requirements		
42. ① ② ③ ④ ⑤ ⑥	Gave clear instructions concerning lesson assignments		

The Instructor:

43.	①②③④⑤⑥	Clearly demonstrated how the work should be approached
44.	①②③④⑤⑥	Made the aims of the assessment clear
45.	①②③④⑤⑥	Explained lesson material clearly
46.	①②③④⑤⑥	Had a method of teaching that helped me understand the lesson material
47.	①②③④⑤⑥	Presented the background of concepts that helped me understand the lesson material
48.	①②③④⑤⑥	Used explanations that confused me
49.	①②③④⑤⑥	Used words that students could understand
50.	①②③④⑤⑥	Failed to define new or unfamiliar terms
51.	①②③④⑤⑥	Demonstrated the importance of the subject matter to future work
52.	①②③④⑤⑥	Pointed out practical application of lesson material
53.	①②③④⑤⑥	Related lesson material to real life situations
54.	①②③④⑤⑥	Described relevant personal experience to enhance the lesson
55.	①②③④⑤⑥	Related content of this lesson to other subject matter that students already understand
56.	①②③④⑤⑥	Integrated student experiences to illustrate specific lesson points
57.	①②③④⑤⑥	Had a background of subject matter experience that enhanced the lesson
58.	①②③④⑤⑥	Demonstrated an exceptional knowledge of the subject material
59.	①②③④⑤⑥	Provided feedback on my individual work through comments on papers, oral discussions, etc.
60.	①②③④⑤⑥	Returned class work in time to benefit my progress
61.	①②③④⑤⑥	Did <u>not</u> review in a way that students could understand their weaknesses
62.	①②③④⑤⑥	Used questions to test for knowledge, skills and attitudes
63.	①②③④⑤⑥	Called on students to answer questions even if they have <u>not</u> indicated that they want to talk
64.	①②③④⑤⑥	Encouraged students to think out answers
65.	①②③④⑤⑥	Referred questions back to students
66.	①②③④⑤⑥	Publicly ridiculed and criticized students who responded incorrectly
67.	①②③④⑤⑥	Answered students' questions in a way that helped students understand
68.	①②③④⑤⑥	Sticks to the point in answering questions
69.	①②③④⑤⑥	Continued to answer questions until points of confusion were cleared up
70.	①②③④⑤⑥	Answered student questions in an incomplete or inaccurate way
71.	①②③④⑤⑥	Became angry when questions were asked
72.	①②③④⑤⑥	When unable to answer student's questions researched answers and reported results back to students
73.	①②③④⑤⑥	Listened to students
74.	①②③④⑤⑥	Interrupted speaking student without justification
75.	①②③④⑤⑥	Maintained eye contact with students while listening
76.	①②③④⑤⑥	Looked at the class when talking
77.	①②③④⑤⑥	Referred to students by name
78.	①②③④⑤⑥	Gave personal attention to student work
79.	①②③④⑤⑥	Was concerned with whether the students learned the lesson material
80.	①②③④⑤⑥	Lacked interest in student's learning
81.	①②③④⑤⑥	Had the ability to motivate you to learn the lesson material
82.	①②③④⑤⑥	Motivated me to think rather than just memorize material
83.	①②③④⑤⑥	Encouraged students to actively participate in the lesson activities
84.	①②③④⑤⑥	Did <u>not</u> inspire confidence in individual student performance
85.	①②③④⑤⑥	Knew when students didn't understand the lesson material
86.	①②③④⑤⑥	Suggested specific study skills to improve my performance in this lesson
87.	①②③④⑤⑥	Had a positive attitude about training students
88.	①②③④⑤⑥	Had a cynical attitude that repelled students
89.	①②③④⑤⑥	Seemed to genuinely respect students as individuals
90.	①②③④⑤⑥	Made me feel that I was could <u>not</u> learn the lesson material
91.	①②③④⑤⑥	Was always arguing with students
92.	①②③④⑤⑥	Treated all students equally
93.	①②③④⑤⑥	Did <u>not</u> evaluate student work fairly
94.	①②③④⑤⑥	Was available outside of the formal class period to explain or clarify lesson material
95.	①②③④⑤⑥	Has been instrumental in increasing my knowledge of the lesson's subject material
96.	①②③④⑤⑥	How would you rate the overall value of this lesson as taught by this instructor? (1) One of the Best, (2) Very Good, (3) Above Average, (4) Below Average (5) Very Bad, (6) One of the Worst
97.	①②③④⑤⑥	Compared with all instructors I have had, both in high school and since, this instructor was: (1) One of the Best, (2) Very Good, (3) Above Average, (4) Below Average (5) Very Bad, (6) One of the Worst
98.	①②③④⑤⑥	Compared with all instructors I have had, both in high school and since, this instructor was <u>not</u> qualified to be an instructor. (1) Strongly Agree, (2) Moderately Agree, (3) Slightly Agree, (4) Slightly Disagree, (5) Moderately Disagree, (6) Strongly Disagree

Appendix G

Frequencies and Percents of Round One Participants' Age, Gender, Education Level, and Ethnicity

	N	f	%
		306	100
Gender			
Male		285	93.1
Female		21	6.9
Age			
18 yrs		1	.3
19 yrs		89	29.1
20 yrs		84	27.5
21 yrs		42	13.7
22 yrs		25	8.2
23 yrs		24	7.8
24 yrs		17	5.6
25 yrs		9	2.9
26 yrs		5	1.6
27 yrs		6	2.0
28 yrs		1	.3
29 yrs		1	.3
30 yrs		1	.3
Non-Response		1	.3
Highest Education Level			
GED		21	6.9
High School Diploma		264	86.3
2yr College		17	5.6
4yr College		4	1.3
Ethnicity			
American Indian/Alaskan Native		4	1.3
Asian		7	2.3
Black		29	9.5
Hispanic		55	18.0
Pacific/Hawaiian Islander		19	6.2
White		168	54.9
Other		24	7.8

Note. One participant did not provide year of birth.

Appendix H

Final Instrument

- 1 - STRONGLY DISAGREE. You would disagree most of the time.
 2 - MODERATELY DISAGREE. You would frequently disagree.
 3 - SLIGHTLY DISAGREE. You would occasionally disagree.
 4 - SLIGHTLY AGREE. You would occasionally agree.
 5 - MODERATELY AGREE. You would frequently agree.
 6 - STRONGLY AGREE. You would agree most of the time.

The Instructor:

1.	①②③④⑤⑥	<i>Repeated difficult ideas several times</i>
2.	①②③④⑤⑥	<i>Signaled the transition from one topic to the next during the lesson</i>
3.	①②③④⑤⑥	<i>Stressed the important points of the lesson</i>
4.	①②③④⑤⑥	<i>Summarized periodically during the lesson</i>
5.	①②③④⑤⑥	<i>Ensured that training equipment and/or training aids were set-up in advance</i>
6.	①②③④⑤⑥	<i>Used technology comfortably while teaching</i>
7.	①②③④⑤⑥	<i>Used concrete, everyday examples to explain the lesson</i>
8.	①②③④⑤⑥	<i>Used examples to get points across during the lesson</i>
9.	①②③④⑤⑥	<i>Ensured the learning environment was safe</i>
10.	①②③④⑤⑥	<i>Effectively paced the lesson delivery</i>
11.	①②③④⑤⑥	<i>Made effective use of lesson time</i>
12.	①②③④⑤⑥	<i>Was enthusiastic about training students</i>
13.	①②③④⑤⑥	<i>Was <u>not</u> interested in training students</i>
14.	①②③④⑤⑥	<i>Seemed to genuinely enjoy training students</i>
15.	①②③④⑤⑥	<i>Clearly defined lesson objectives</i>
16.	①②③④⑤⑥	<i>Clearly demonstrated how the work should be approached</i>
17.	①②③④⑤⑥	<i>Gave a clear idea of the student requirements</i>
18.	①②③④⑤⑥	<i>Gave clear instructions concerning lesson assignments</i>
19.	①②③④⑤⑥	<i>Demonstrated the importance of the subject matter to future work</i>
20.	①②③④⑤⑥	<i>Described relevant personal experience to enhance the lesson</i>
21.	①②③④⑤⑥	<i>Related lesson material to real life situations</i>
22.	①②③④⑤⑥	<i>Related content of this lesson to other subject matter that students already understand</i>
23.	①②③④⑤⑥	<i>Continued to answer questions until points of confusion were cleared up</i>
24.	①②③④⑤⑥	<i>Listened to students</i>
25.	①②③④⑤⑥	<i>Looked at the class when talking</i>
26.	①②③④⑤⑥	<i>Did not review in a way that students could understand their weaknesses</i>
27.	①②③④⑤⑥	<i>Publicly ridiculed and criticized students who responded incorrectly</i>
28.	①②③④⑤⑥	<i>Answered student questions in an incomplete or inaccurate way</i>
29.	①②③④⑤⑥	<i>Became angry when questions were asked</i>
30.	①②③④⑤⑥	<i>Interrupted speaking student without justification</i>
31.	①②③④⑤⑥	<i>Lacked interest in student's learning</i>
32.	①②③④⑤⑥	<i>Did not inspire confidence in individual student performance</i>
33.	①②③④⑤⑥	<i>Had a cynical attitude that repelled students</i>
34.	①②③④⑤⑥	<i>Made me feel that I could not learn the lesson material</i>
35.	①②③④⑤⑥	<i>Was always arguing with students</i>
36.	①②③④⑤⑥	<i>Did not evaluate student work fairly</i>
37.	①②③④⑤⑥	<i>Presented the lesson in an interesting way</i>
38.	①②③④⑤⑥	<i>Referred to students by name</i>
39.	①②③④⑤⑥	<i>Was concerned with whether the students learned the lesson material</i>
40.	①②③④⑤⑥	<i>Had the ability to motivate you to learn the lesson material</i>
41.	①②③④⑤⑥	<i>Motivated me to think rather than just memorize material</i>
42.	①②③④⑤⑥	<i>Encouraged students to actively participate in the lesson activities</i>
43.	①②③④⑤⑥	<i>Knew when students didn't understand the lesson material</i>
44.	①②③④⑤⑥	<i>Suggested specific study skills to improve my performance in this lesson</i>
45.	①②③④⑤⑥	Compared with all instructors I have had, both in high school and since, this instructor was: (1) One of the Best, (2) Very Good, (3) Above Average, (4) Below Average (5) Very Bad, (6) One of the Worst
46.	①②③④⑤⑥	Compared with all instructors I have had, both in high school and since, this instructor was <u>not</u> qualified to be an instructor. (1) Strongly Agree, (2) Moderately Agree, (3) Slightly Agree, (4) Slightly Disagree, (5) Moderately Disagree, (6) Strongly Disagree

Appendix I

Revised Factors, Sub-factors, and Items

Factor 1 - Effective Delivery: The student perceived the instructor clearly and effectively delivered the lesson material.

Sub-factor 1.1 The instructor employed training techniques designed to enhance student learning

- Item 2 Signaled the transition from one topic to the next during the lesson
- Item 4 Summarized periodically during the lesson
- Item 10 Effectively paced the lesson delivery
- Item 11 Made effective use of lesson time
- Item 25 Looked at the class when talking

Sub-factor 1.2 The instructor managed a safe, prepared learning environment.

- Item 5 Ensured that training equipment and/or training aids were set-up in advance
- Item 6 Used technology comfortably while teaching
- Item 9 Ensured the learning environment was safe

Sub-factor 1.3 The instructor appeared enthusiastic and interested in training students.

- Item 12 Was enthusiastic about training students
- Item 13 Was not interested in training students
- Item 14 Seemed to genuinely enjoy training students
- Item 24 Listened to students

- Sub-factor 1.4 The instructor clearly communicated student's goals and objectives.
- Item 15 Clearly defined lesson objectives
 - Item 17 Gave a clear idea of the student requirements
 - Item 18 Gave clear instructions concerning lesson assignments
 - Item 16 Clearly demonstrated how the work should be approached
 - Item 23 Continued to answer questions until points of confusion were cleared up
- Sub-factor 1.5 The instructor used examples to stress the importance lesson.
- Item 1 Repeated difficult ideas several times
 - Item 3 Stressed the important points of the lesson
 - Item 8 Used examples to get points across during the lesson
 - Item 7 Used concrete, everyday examples to explain the lesson
- Sub-factor 1.6 The instructor demonstrated the importance of the lesson material.
- Item 19 Demonstrated the importance of the subject matter to future work
 - Item 20 Described relevant personal experience to enhance the lesson
 - Item 21 Related lesson material to real life situations
 - Item 22 Related content of this lesson to other subject matter that students already understand
- Factor 2 - Negativity: The student perceived the instructor as disinterested in, and uncooperative toward, student's learning.

Sub-factor 2.1 The instructor did not review or evaluate student work accurately.

- Item 26 Did not review in a way that students could understand their weaknesses
- Item 28 Answered student questions in an incomplete or inaccurate way
- Item 36 Did not evaluate student work fairly

Sub-factor 2.2 The instructor did not demonstrate a positive, respectful or professional attitude toward students.

- Item 27 Publicly ridiculed and criticized students who responded incorrectly
- Item 29 Became angry when questions were asked
- Item 30 Interrupted speaking student without justification
- Item 35 Was always arguing with students
- Item 33 Had a cynical attitude that repelled students

Sub-factor 2.3 The instructor did not inspire confidence and lacked interest in student learning.

- Item 31 Lacked interest in student's learning
- Item 32 Did not inspire confidence in individual student performance
- Item 34 Made me feel that I could not learn the lesson material

Factor 3 - Concern for Learning: The student perceived the instructor appeared to be concerned about the student's progress and motivation to learn.

Sub-factor 3.1 The instructor was able to motivate the students to learn.

- Item 37 Presented the lesson in an interesting way
 - Item 40 Had the ability to motivate you to learn the lesson material
 - Item 41 Motivated me to think rather than just memorize material
 - Item 42 Encouraged students to actively participate in the lesson activities
- Sub-factor 3.2 The instructor knew the students and was concerned with their progress.
- Item 38 Referred to students by name
 - Item 39 Was concerned with whether the students learned the lesson material
 - Item 43 Knew when students didn't understand the lesson material
 - Item 44 Suggested specific study skills to improve my performance in this lesson

Appendix J

Final Factors, Sub-factors, and Items

Factor 1 - Effective Delivery

- Item 7 Used concrete, everyday examples to explain the lesson
- Item 8 Used examples to get points across during the lesson
- Item 9 Ensured the learning environment was safe
- Item 12 Was enthusiastic about training students
- Item 14 Seemed to genuinely enjoy training students
- Item 20 Described relevant personal experience to enhance the lesson
- Item 21 Related lesson material to real life situations
- Item 22 Related content of this lesson to other subject matter that students already understand
- Item 37 Presented the lesson in an interesting way
- Item 39 Was concerned with whether the students learned the lesson material
- Item 40 Had the ability to motivate you to learn the lesson material
- Item 41 Motivated me to think rather than just memorize material
- Item 42 Encouraged students to actively participate in the lesson activities

- Item 43 Knew when students didn't understand the lesson material
- Item 44 Suggested specific study skills to improve my performance in this lesson

Factor 2 - Negativity

- Item 13 Was not interested in training students
- Item 25 Looked at the class when talking
- Item 27 Publicly ridiculed and criticized students who responded incorrectly
- Item 28 Answered student questions in an incomplete or inaccurate way
- Item 29 Became angry when questions were asked
- Item 30 Interrupted speaking student without justification
- Item 31 Lacked interest in student's learning
- Item 32 Did not inspire confidence in individual student performance
- Item 33 Had a cynical attitude that repelled students
- Item 34 Made me feel that I could not learn the lesson material
- Item 35 Was always arguing with students
- Item 36 Did not evaluate student work fairly

Factor 3 - Concern for Students

- Item 3 Stressed the important points of the lesson

- Item 5 Ensured that training equipment and/or training aids were set-up in advance
- Item 6 Used technology comfortably while teaching
- Item 10 Effectively paced the lesson delivery
- Item 15 Clearly defined lesson objectives
- Item 16 Clearly demonstrated how the work should be approached
- Item 17 Gave a clear idea of the student requirements
- Item 18 Gave clear instructions concerning lesson assignments
- Item 23 Continued to answer questions until points of confusion were cleared up
- Item 24 Listened to students