The Performance Impact of Leadership Training: an Empirical Study on the United States Coast Guard

Chad A. Long
Embry Riddle Aeronautical University

ABSTRACT

Training programs that focus on leadership and management are becoming more prevalent in society with little regard to the training's impact. This study's purpose was to determine if there was a relationship between leadership training and performance. The sequential mixed-method study examined the impact of a 33-day resident training course on the graduate's performance. The measurement of performance was obtained quantitatively through annual performance evaluations and qualitatively through interviews. The study spanned 4 years, collecting one pre-graduation evaluation and three post-graduation evaluations. The results of this study indicated that leadership training significantly increased overall performance.

Training has assumed an ever-increasing status of importance in government and industry. Technological advances and organizational complexity have created an environment that forces a corporation to prepare its employees for current and future developments. This corporate preparation has become extremely costly. In 2011, employers in the United States spent $59.7 billion on education and formal training (Freifeld, 2011).

While expensive, the need for training is apparent to many managers. However, the impact is often difficult to see. Most employers intuitively feel that training is valuable yet never measure its benefit to the organization. Phillips (2003) explains, "[Employers] logically conclude that training can pay-off in important bottom-line measures such as productivity improvements, qualities enhancements, cost reductions, and time savings. . . . While the payoffs are assumed to exist and training appears to be needed, more evidence is needed, or training funds may not be allocated in the future" (p. 2).

There has been a large increase in the quantity of training programs that focus on leadership and management (Sessa, Matos, & Hopkins, 2009). Many organizations are concerned about the leadership inadequacies of their employees and, as a result, are committing to education and training that deepens the skills, perspectives, and competencies of their leaders (Conger & Benjamin, 1999). Based on historical training budgets, the annual spending for leadership training and development will continue to grow throughout the next decade. This growth in leadership training will be short lived if there is little or no correlation between the money spent on training and performance improvement for the corporation.

The solution to this impending crisis is to start evaluating leadership training through appropriate research (Dvir, Eden, Avolio & Shamir, 2002). Sogunro (1997) explains "[Leadership training providers] have rarely assessed impacts in terms of effectiveness and efficiency regarding cost and benefits to the funders; many lack assessment of impacts on participants of the program, especially through a combination of pre-training, during-training, post-training and follow-up evaluation procedures; and most lack in-depth data-gathering strategies involving mixed research methods" (p. 714). While there are examples of post-training evaluations, most are conducted by end-of-the-program questionnaires. These evaluations primarily provide input on the participant's reaction to the course and provide minimal information on the real impact of the training on the participant's performance (Sogurno, 1997).

In an effort to identify the real impact of leadership, the following three areas need to be explored in more detail: (a) understanding leadership, (b) training leaders, and (c) evaluating leadership training. The first section will provide a background to leadership theory. The next section reviews literature on the fundamental of leadership education and training. The last section examines the methods used to evaluate leadership training.

UNDERSTANDING LEADERSHIP

Leadership is a complex phenomenon. This becomes evident when one examines the terminology used to conceptualize leadership and describe its many dimensions. Through the ages, scholars have had a difficult time getting a theoretical grasp on the word *leadership*. Stogdill (1974), followed by Bass (1990), analyzed thousands of studies on the topic of leadership. Stogdill noted that "the endless accumulation of empirical data has not produced an integrated understanding of leadership" (p. vii). An investigation of leadership research

published in *Leadership Quarterly* from 1994 to 2003 found a wide variety of leadership

definitions and no explicit patterns or evolutions of a single definition over time (Martin and

Ernst, 2005). While most definitions discuss empowerment (Lohmann, 1992), real change (Rost,

1993), and shared goals (Northouse, 1997), the United States Coast Guard has a more

authoritarian definition for leadership. They define it as "the ability to influence others to obtain

their obedience, respect, confidence, and loyal cooperation" (USCG, 2006, p. 1). Although the

field of leadership has developed since Stogdill's research, it is still a field that is very

ambiguous in nature (Roberts, 2007).

While leadership definitions are important, the dimensions, or competencies, of leadership have

a more direct linkage to performance (Young & Dulewicz, 2005; Obiwuru, Okwu, Akpa,

Nwankwere, 2011). There are many dimensions which combined make an individual more

likely to be a successful leader (Van Fleet & Yukl, 1986). These dimensions are a mixture of

traits and behaviors. Stogdill's 1974 study is the largest and most referenced study on leadership

traits and behaviors (Bass, 1990). When the United States Coast Guard Chief Petty Officer

evaluation categories are compared against Stogdill's leadership dimensions (Table 1), it is

apparent that there is a strong relationship between the two. Twenty of the 24 enlisted evaluation

categories match up with one of Stogdill's dimensions. The only exceptions are safety and

occupational health, evaluations, work-life sensitivity/expertise, and customs and courtesies.

While these characteristics are important to the Coast Guard on an organizational level, they

have little to no connection to leadership literature. Additionally, there are some leadership

dimensions that are not encompassed by the Coast Guard performance evaluation form.

Table 1

Leadership Dimensions: Stogdill's Survey of Findings Compared to Coast Guard Performance Evaluation Categories

Coast Guard Evaluation Categories	Stogdill's Leadership Dimensions
Performance Section	
Professional Specialty Knowledge	Knowledge
Professional Development	Education
Administrative Ability	Administrative Ability
Organization	Task Organization
Using Resources	Resourcefulness
Monitoring Work	Responsibility in the pursuit of objectives
Safety and Occupational Health	----
Stamina	Persistence against obstacles
Communication	Fluency of speech
Leadership Section	
Directing Others	Dominance
Working with others	Cooperativeness
Developing subordinates	Nurturance
Responsibility	Drive for Responsibility
Evaluations	----
Work-life Sensitivity/ Expertise	----
Setting the Example	Maintaining a standard of performance
Military Section	
Military Bearing	Appearance/ Grooming
Customs and Courtesies	----
Professional Qualities Section	
Health and Well Being	Activity/ Energy
Integrity	Personal Integrity
Loyalty	Belonging, Identification, and Loyalty
Respecting Others	Tolerance of follower's freedom of action
Human Relations	Diplomacy
Adaptability	Adaptability

Note. **Bold** indicates that an evaluation category matches up with one of Stogdill's leadership dimensions. Coast Guard evaluation section headings are given in *italics*. The data in the second column are from *Bass and Stogdill's Handbook of Leadership*, by B. M. Bass, 1990, New York: Free Press.

Characteristics such as creativity, self-confidence, enthusiasm, and sociability are some of the important leadership dimensions in the literature that are not part of the Coast Guard performance evaluation.

TRAINING LEADERS

Given the need of so many organizations, including the Coast Guard, to develop leaders exemplifying strength in all leadership dimensions, the question arises whether they can be developed through training. On a conceptual level, leadership training is built on two fundamental premises. The first is that human beings are capable of learning, and the other is that leadership can be taught. While few people question that humans can learn, some still doubt whether leadership can be taught. Those who do not believe it can be learned view leadership as a set of innate abilities and advocate the view that great leaders are "born, not made," similar to the Great Man theory of leadership. Fortunately for those not well endowed with the natural qualities of a successful leader, social science has repeatedly demonstrated that leadership can be taught (Horner, 1995).

The purpose of leadership training differs from organization to organization, but the primary purpose of most leadership development interventions is to improve individual managerial skills and on-the-job performance (Burke & Day, 1986). Burke and Day completed a meta-analysis of 70 published and unpublished studies spanning over 30 years on the topic of managerial leadership development. Collins (2002) completed a follow-up meta-analysis on the Burke and Day research, evaluating 83 studies over a 19-year period. Collins found that more and more organizations were looking at leadership training programs to have a positive effect on both the individual and the organization. McCauley and VanVelsor (2003) explain the purpose of

leadership training by describing the three phases of leadership development education. The first phase is the improvement of an individual's capacities, even when the training is directed at teams or organizations. From there, the training focuses on making an individual effective in a variety of leadership roles and processes. This phase does not create a leader, but it focuses on giving the trainee the skills they need to act in both formal and informal leadership positions. Lastly, the goal of training is to expand an individual's leadership capacities. While identifying the phases of leadership training is relatively simple, assessing the impact of the training often becomes challenging.

EVALUATING LEADERSHIP TRAINING

The effectiveness of a training program is not something that is always apparent to the coordinators or stakeholders. The primary method to determine the effect of a program is through evaluation. The evaluation findings are then used to determine if a program is producing its intended results (Torres & Preskill, 2001). Many companies are using the results of program evaluations also to improve organizational performance and determine the return on investment of the program (Phillips, 2003).

There are many models that are being used to evaluate training programs such as the CIPP (Context, Input, Process, and Product) model, Constructionist Evaluation model, Results Assessment Model and Kirkpatrick's four-level model (Guba & Lincoln, 1989; Stufflebeam, Foley, Gephart, Hammond, Merriman & Provus, 1971; Swanson & Holton, 1999). Kirkpatrick's four-level model (1959a, 1959b, 1960a, 1960b) is the most prevalent framework for evaluating training (Behrens & Benham, 2007). It measures the effectiveness of a training program on four levels. Each level builds upon the next. Level 1 measures participant reaction to the training

material. A positive reaction does not guarantee learning, but a negative reaction almost certainly reduces the possibility. Level 2 measures the student level of learning. This information is normally attained through pre- and post-intervention testing. The third level in Kirkpatrick's model is the measurement in learning transfer or generalization (Baldwin & Ford, 1988). A training program whose participants use many of the techniques learned during training in their daily routine would score well in this level. Level 4, the final level, measures the results of the training. This level is focused on performance improvement such as, increased production, improved quality, reduced frequency of accidents, and even higher profits. Performance improvement is the apex in Kirkpatrick's training effectiveness model.

Three studies have completed a large-scale analysis of the performance effects (Kirkpatrick's level 4) of leadership training. These meta-analysis studies are Burke and Day (1986), Zhang (1999), and Collins (2002). These studies, when combined together, evaluate most all of the leadership development literature on performance, both published and unpublished, from 1956 to 2001. They encompass leadership development programs in the majority of major industries in the United States and throughout the world including: automotive, financial, manufacturing, technology, utilities, education, government, medical, military, and many others. All three studies used effect sizes to evaluate the impact of performance. Several standards exist in the literature to assess the meaningfulness of effect size. Cohen (1977) suggests 0.2 as a minimal effect, 0.5 as a moderate effect, and 0.8 as a meaningful effect. The three studies showed a consistent effect range from minimal to moderate levels. The studies are encouraging, but the majority of results are based on post-training participant questionnaires. The shortage of performance-based studies on the issue is distressing.

METHOD

This study has a mixed-method design. The format of the study was sequential where the quantitative section was performed prior to the qualitative segment. The quantitative portion evaluated enlisted performance evaluations and then qualitative interviews were used to describe the results in more detail.

QUANTITATIVE DESIGN

The quantitative design of the study focused on the change in performance evaluations after completing the Chief Petty Officer Academy, a 33-day, resident, leadership development program. The Coast Guard has a mandatory attendance policy for the Academy, making it unfeasible to have a control group for the study. Collins and Holton (2004) state that there should be more single group pretest-posttest studies in the area of leadership development which is the format of this study. The pretest was composed of the performance information prior to graduation, while the posttests were focused on performance information after graduation. The performance evaluations spanned from approximately 1 year before a Chief graduates from the Academy to 3 years after their graduation, for a total of four performance evaluations.

Three years is the standard amount of time required to determine whether a change in performance has occurred as a result of an intervention (Collins, 2002). A one-way repeated measure analysis of variance (ANOVA) was used to establish if there were any significantly different performance appraisal totals over the 4-year time periods. Additionally, a pairwise analysis was conducted to establish which time periods were statistically unique.

QUALITATIVE DESIGN

For the qualitative portion of the study, four Chiefs were probed in-depth with interviews. The four Chiefs were a purposive sample consisting of two graduates who show a significant increase in overall performance and two graduates with limited overall performance change (Merriam, 1988). The interviews were coded, looking for similarities and differences in experience among the four Chiefs. The qualitative data were compared against the quantitative results in an effort to better describe the performance effects of the Academy.

QUANTITATIVE INSTRUMENTATION

The instrumentation for the quantitative section of this study was the enlisted employee review system, or enlisted evaluation performance form. The Coast Guard enlisted appraisal system uses both quantitative and qualitative analysis. The system is mixed between trait-based and effectiveness-base, with the majority of the evaluation focused on trait-based performance dimensions. The performance appraisal uses a 7 point, graphic rating scale with narrative comments. The scale goes from a low score of 1 to a high of 7. Narrative comments are required only for ratings of 1, 2, or 7. This study evaluated only the numerical scores used on the performance evaluation and did not address narrative comments.

The Chief Petty Officer performance appraisal form is computer generated with 24 categories divided into four sections of performance. The Coast Guard enlisted evaluation sections of performance are leadership, professional qualities, performance, and military. For ease of understanding and to avoid confusion in this study, the titles of the Coast Guard enlisted evaluation sections have been altered. The sections are defined by the enlisted evaluation portion of the *Coast Guard Personnel Manual* (USCG, 2011) as,

Leadership Abilities. Measures a member's ability to direct, guide, develop, influence, and support others' performing work.

Professionalism. Measures those qualities the Coast Guard values in its people.

Organizational Responsibilities. Measures a member's willingness to acquire knowledge and the ability to use knowledge, skill, and direction to accomplish work.

Military Protocol. Measures a member's ability to bring credit to the Coast Guard through personal demeanor and professional actions. (p. 5-3)

The sections of performance evaluation are further divided into seven leadership abilities categories, six professionalism categories, nine organizational responsibilities categories, and two military protocol categories. Table 1 displays the categories in each section. The Coast Guard enlisted evaluation instrument and process of implementation has been deemed reliable and valid based on internal examination (USCG, 2011).

QUALITATIVE INTERVIEW PROTOCOL

The interviews consisted of five questions. The first four of the five questions were derived from McCauley and Hughes-James's (1994) study on the effects of a leadership development program on 38 school superintendents. The questions were chosen because they were specifically directed towards the impact of leadership training. The first two questions were general questions aimed at examining how the Academy impacted the participants (Mason, 1996). The next two questions gave the participants the opportunity to discuss the good and bad aspects of the program, regardless of its personal impact. The last question was created to examine the relationship between the Academy and enlisted performance evaluations. The interviews were formulated around the following five questions.

1. What are the two or three most important ways the Academy has had an impact on you?

2. If I asked those who work with you what you are doing differently compared to 6 to 7 years ago, what would they say? How much would you attribute these differences to the Academy?

3. What were the highlights or most positive aspects of the Academy for you?

4. What were the lowlights or least positive aspects of the Academy for you?

5. What effect do you think that the Academy had on your performance as measured through your enlisted evaluations?

Although the interview was formulated around five questions, follow-up questions were used occasionally to better understand the participant's answers. The interview process was carried out as equally as possible between the four participants to ensure maximum uniformity.

QUANTITATIVE SAMPLE

From January 1, 1999, to December 31, 1999, three hundred and ninety-nine Chief Petty Officers graduated from the Chief Petty Officer Academy. However, this study included only a portion of those individuals in the sample. The Chiefs excluded fell into two categories: irretrievable data and those promoted during the period of the study.

An individual's data were no longer available when they retired or departed from the Coast Guard. The second reason that a Chief Petty Officer was excluded from the sample was promotion during the 4-year study. When a First Class Petty Officer (E-6) is promoted to a Chief Petty Officer (E-7) he or she is evaluated with a different performance evaluation form. Additionally, individuals promoted from Chief Petty Officer (E-7) to Senior Chief Petty Officer

(E-8) were also excluded from the study. While the performance evaluation forms are the same between the two Chiefs, the newly promoted E-8 is now being evaluated against a different set of peers. The interpretation on how this newly promoted E-8 performs against his or her new peer group sometimes results in a decrease in performance evaluation scores.

Of the 399 Chiefs who graduated in 1999, one hundred and thirty-nine Chief Petty Officers agreed to participate in the study. The sample was reduced to 71 Chiefs when it was discovered that 68 Chiefs had data that were irretrievable. Of the 71 with retrievable data only 40 met the requirement of not being promoted during the 4-year study. The effective sample for this study was the 40 Chiefs who met all the required conditions of the study.

QUALITATIVE SAMPLE

The four Chiefs in the qualitative portion of the study were separated into two categories: Chiefs with no significant overall performance variation ($p > 0.05$), as measured through their enlisted evaluation totals, and Chiefs with significant positive overall performance variation ($p \leq 0.05$), as measured through their enlisted evaluation totals. The final two Chiefs for the no-significant-performance variation were selected because they had the least amount of change in performance after Academy graduation. The two significant variation Chiefs were not selected because they had the largest increase in performance after graduation, but because their performance increase mirrored the significant increase found in the quantitative sample.

QUANTITATIVE PROCEDURE

The performance evaluation process assessed an enlisted member's performance and value to the Coast Guard through a system of multiple evaluators who present independent views and thus

ensured accurate, prompt, and correct reporting (USCG, 2011). The quantitative information for this study was collected from data available on Chief Petty Officers' annual Coast Guard performance evaluation forms. In higher education, there is talk of grade inflation (Babcock, 2010). The same phenomenon exists in performance appraisals of the American workforce (Glover, 1996). The Coast Guard's enlisted evaluation system also displayed inflation over time. The mean evaluation total for Chief Petty Officers between 1999 and 2002 showed a consistent annual increase. Over the 4-year period, evaluation totals increased from a mean of 129.32 to 134.79, a 4.21% increase. The data in this study were adjusted by reducing the Chief's evaluation scores relative to this annual increase. The adjusted scores were analyzed using SPSS for windows, version 11.

QUALITATIVE PROCEDURES

Qualitative interviews were used to expand on the impact of the leadership training. The interviews allowed the respondents to elaborate, in their own words, on their feelings about the Coast Guard Chief Petty Officer Academy. Sources of data were analyzed for content using Lincoln and Guba's (1985) constant comparative method. This method involved the process of unitizing and categorizing information into emergent themes. The data were divided into the simplest feasible units of information to define the categories. This process was repeated until all the data were broken up into units of information. The broad themes that developed were very similar to many of the 24 categories in the Coast Guard's enlisted evaluation form. To better complement the quantitative data, the interviews were coded using the Coast Guard enlisted evaluation categories to sort the data. A file for each enlisted evaluation category was established. Each category was checked to ensure that it was internally homogeneous, externally

heterogeneous, and mutually exclusive. The categorizing was continued until the sources were exhausted and the categories were saturated.

HYPOTHESIS

The yearly Coast Guard enlisted evaluation form was the instrument used to test the hypothesis. The overall performance included all 24 categories of the enlisted evaluation. The total score was presented as the mean of all 24 categories, not the sum of all the categories. All of the means were corrected for the annual creep in scores found in the Chief Petty Officer performance evaluation data.

> *H1.*   A statistically significant positive relationship exists between Academy
>
> graduation and an individual's performance.

RESULTS

QUANTITATIVE RESULTS

A one-way repeated measures ANOVA was conducted with the factor being time interval since the year prior to Academy graduation and the dependent variable being the mean enlisted evaluation total score or overall performance, corrected for evaluation inflation. Table 2 presents the results of the repeated measures ANOVA. The mean and 95% confidence interval are graphically presented in Figure 1.   The results for the ANOVA indicated a significant time effect, Wilks' $\Lambda = 0.96$, $F(3, 2877) = 15.07$, $p < 0.01$. These results suggest that leadership training positively impacts overall performance as measured through enlisted evaluations.

The results for effect size indicated less than minimal effect, $d = 0.18$. These results suggest that while leadership training positively impacts performance as measured through enlisted evaluations, the impact of the training on enlisted evaluations was minor.

Follow-up polynomial contrasts indicate a significant linear and quadratic effect with means increasing over time, $F(1,959) = 30.87, p < 0.01$, and $F(1,959) = 7.85, p < 0.01$, respectfully. The cubic polynomial contrast was nonsignificant.

Table 2

The Results of the Repeated Measures ANOVA for Overall Performance

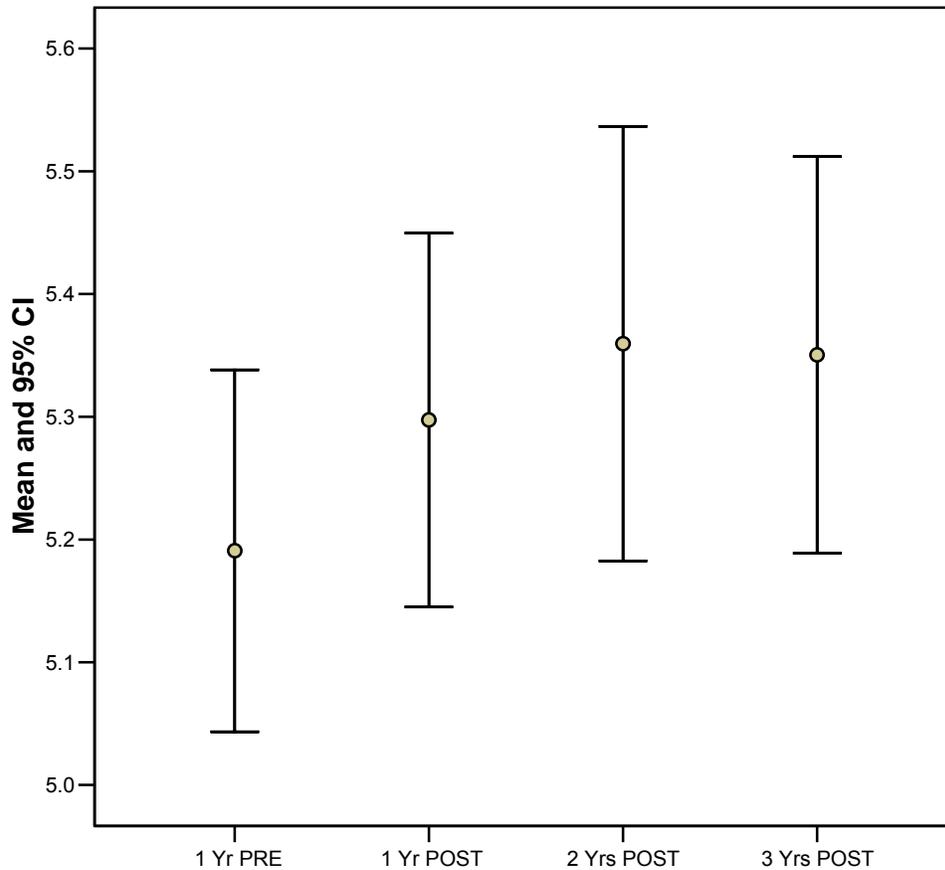| Source	SS	df	MS	F	p
Between Subjects	1397.36	959	1.46		
Within Subjects	17.30	3	5.77	15.07	< 0.01
Error	1100.71	2877	0.38		
Total	2515.37	3839			

Figure 1. Mean and 95% confidence interval for enlisted evaluations.

While the repeated measures ANOVA shows that there was a significant difference between the means, the paired sample *t*-test examines when the significant difference was occurring. The *t*-test results are presented in Table 3. The results indicate the pre-graduation evaluation (M = 5.19, SD = 0.81) was significantly less than all three post-graduation evaluations. The post 2 (M = 5.36, SD = 0.83) evaluation totals were significantly higher than the 2 previous years in the study but statistically equivalent to post 3 (M = 5.35, SD = 0.77). Post 3 was considered statistically equivalent to both post 1 (M = 5.30, SD = 0.82) and post 2. In summary, the effects of the Academy training continually increased evaluation totals the first 2 years after graduation and then returned to somewhere between the post 1 and post 2 levels on the 3rd year.

This research shows that Chief Petty Officers increased in total performance when they were given leadership development training. The overall increase in performance was 2.8%. This measurement was an average of each of the three post Academy graduation years. All measurements were adjusted for evaluation inflation. The increase in overall performance the first post-graduation year was larger than the subsequent 2 years. The performance peaked on the 2nd year after graduation with an overall gain of 3.3%. The effect size was 0.18 for this study, which was slightly less than Collins found ($M = 0.38$) in her research on the effects of leadership development training on performance (Collins, 2002).

Table 3

Paired Sample t-*test Comparing the Four Time Intervals of Enlisted Evaluations (Overall Performance)*

Pair	*M* Diff	*SD*	*SE*	*t*
1PRE-1POST	-0.11	0.88	0.03	-3.76**
1PRE-2POST	-0.17	0.89	0.03	-5.90**
1PRE-3POST	-0.16	0.94	0.03	-5.26**
1POST-2POST	-0.06	0.77	0.02	-2.50*
1POST-3POST	-0.05	0.93	0.03	-1.77
2POST-3POST	0.01	0.83	0.03	0.34

*The *t*-test was significant at the 0.05 level. ** 0.01 level.

QUALITATIVE RESULTS

The qualitative results were based on graduate interviews. The emergent themes that developed were very similar to many of the 24 categories in the Coast Guard's enlisted evaluation form. The Coast Guard enlisted evaluation form for the Chief Petty Officer consists of 24 categories which are grouped into four sections, or meta-themes. The four meta-themes are leadership abilities, organizational responsibilities, military protocol, and professionalism. Fifteen of the 24 enlisted evaluation categories, or 62.5%, were discussed in the qualitative interviews. The missing categories include: responsibility, setting the example, professional/ specialty knowledge, administrative ability, organization, monitoring work, safety and occupational health, loyalty, and adaptability. The structure of the meta-themes and discussed subordinate themes are represented in Figure 2.

To determine the amount of interview discussion, two types of manifest effect sizes were calculated on the qualitative results (Onwuegbuzie & Teddlie, 2003). The frequency of occurrence was obtained by tabulating the number of times a theme or meta-theme was mentioned during the interviews. The intensity effect size was calculated by converting the frequency of occurrence into a percentage. The manifest effect sizes and percentage endorsement for each of the categories or themes are calculated and displayed in Table 4. The personal endorsement statistic represents the percentage of interviewees who discussed a particular theme during the course of the interview.

The 24 themes in the Coast Guard enlisted evaluation form did not completely encompass all of the information from the interviewees. Unitizing and categorizing were continued on these

outlying data until the sources were exhausted, the categories were saturated, regularities emerged, or over-extension occurred as described by Lincoln and Guba (1985). Four new themes immerged as a result of this additional assessment (Table 5). The themes were no low point/ good school, life changing experience, should focus more on traditions, and should be earlier in career.
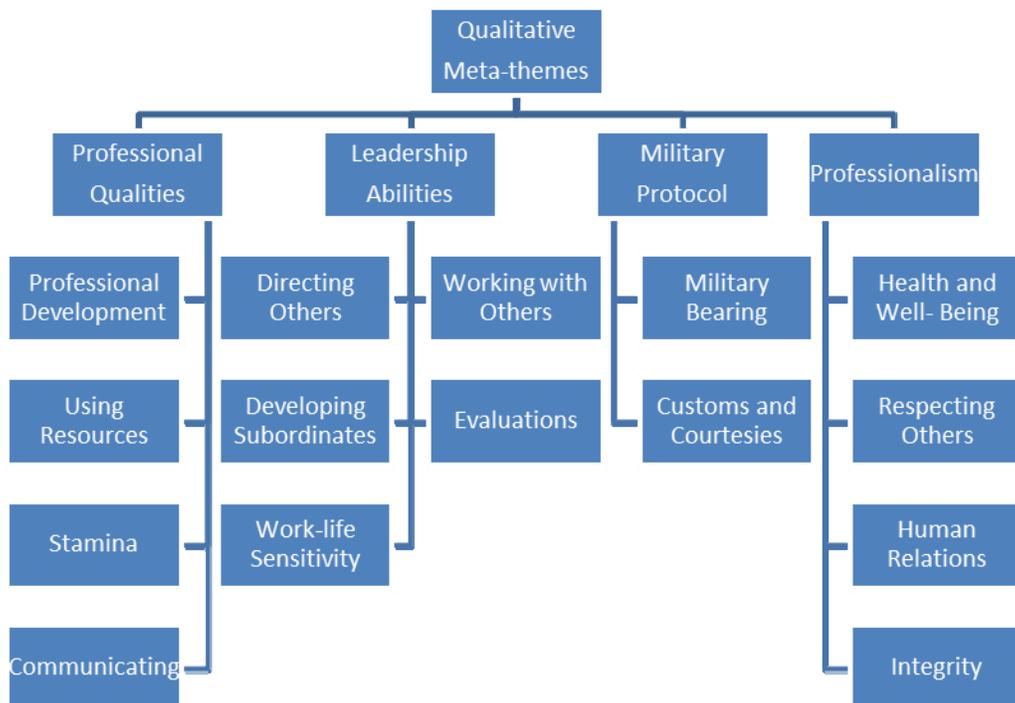

Figure 2. The structure of qualitative meta-themes.

Table 4

Manifest Effect Size and Frequency Distribution of the Meta-themes Associated with the Chief

Petty Officer Academy

Theme	Meta-Theme	Frequency of occurrence (units)	Intensity effect sizes (% of total)	Percentage endorsement ($n = 4$)
Directing Others	Leadership Abilities	10	17.8	100.0
Working with Others	Leadership Abilities	15	26.8	75.0
Developing Subordinates	Leadership Abilities	2	3.7	25.0
Responsibility	Leadership Abilities	0	0.0	--
Evaluations	Leadership Abilities	1	1.8	25.0
Work-life Sensitivity	Leadership Abilities	1	1.8	25.0
Setting an Example	Leadership Abilities	0	0.0	--
Leadership Abilities Subtotal			*29*	*51.8*
Prof./ Specialty Knowledge	Organizational Resp.	0	0.0	--
Professional Development	Organizational Resp.	1	1.8	25.0
Administrative Ability	Organizational Resp.	0	0.0	--
Organization	Organizational Resp.	0	0.0	--
Using Resources	Organizational Resp.	5	8.8	50.0
Monitoring Work	Organizational Resp.	0	0.0	--
Safety and Occ. Health	Organizational Resp.	0	0.0	--
Stamina	Organizational Resp.	4	7.1	50.0
Communicating	Organizational Resp.	9	16.1	100.0
Organizational Responsibilities Subtotal			*19*	*34.0*
Military Bearing	Military Protocol	1	1.8	25.0
Customs and Courtesies	Military Protocol	3	5.3	50.0
Military Protocol Subtotal			*4*	*7.1*
Health and Well-being	Professionalism	1	1.8	25.0
Integrity	Professionalism	0	0.0	--
Loyalty	Professionalism	0	0.0	--
Respecting Others	Professionalism	1	1.8	25.0
Human Relations	Professionalism	1	1.8	25.0
Adaptability	Professionalism	1	1.8	25.0
Professionalism Subtotal			*4*	*7.1*
Total		56	100.0	

Table 5

Manifest Effect Size and Frequency Distribution for the Outlying Themes Associated with the

Chief Petty Officer Academy

Theme	Frequency of occurrence (units)	Intensity effect sizes (% of total)	Percentage Endorsement ($n = 4$)
No Low Point/ Good School	6	30	100
Life Changing	5	25	25
More Traditions	3	15	25
Should Be Earlier	3	15	50
Other	3	15	75
Total	20	100	

DISCUSSION

The ultimate goal of the Coast Guard's Chief Petty Officer Academy is "to provide critical

leadership skill sets need by all [Chief Petty Officers]" (USCG, 2003, p. 4). Collins (2002)

would classify the goal of this program as educational. The program was not specifically

developed to improve individual or organizational performance. Collins's research found that

only 30% of leadership development programs had increased performance as their desired

outcome. While the curriculum of the Chief Petty Officer Academy was not specifically directed

towards improving an individual's performance, it appears that it has that effect based on the

quantitative and qualitative results.

The overall performance for Academy graduates was significantly higher than their pre-Academy totals. Although understanding that the post-graduation totals were significantly higher is important, that information does not provide a complete picture of the results of the Academy. To better explain the results, it was imperative to thoroughly examine the paired sample *t*-tests for overall performance.

1. Post 1 overall performance totals were significantly higher than pre-graduation totals. This information implies that the program had an immediate effect on the graduating Chiefs.

2. Post 2 overall performance totals were significantly higher than Post 1 graduation totals. This information implies that the program's effect lasted 2 years, continually affecting the graduating Chiefs in a positive manner.

3. Post 3 overall performance totals were not significantly different from either the Post 1 or Post 2 totals. This information implies that the program's lasting effects peaked on the 2nd year after graduation.

Summarizing the three details, the Academy had an increasing positive effect the first 2 years after graduation but this effect started to decline on the 3rd year. Baldwin and Ford (1988) explain that knowledge and skills used in many training programs commonly erode over time. There are many possible explanations for the increase in performance and then slow decline, but the qualitative interviews lead to the following assumption. After graduation, the students are very motivated and they are looking forward to using all the new skills they just learned. Some of the skills that they learn, such as leadership skills, are immediately applied, while other skills, such as conflict resolution and writing awards, can be used only when the opportunity presented itself. Additionally, some of the skills acquired, such as building a quality Chief's mess and

establishing a beneficial organizational network, even if immediately applied could take a couple of years before the results from their efforts became visible. When the immediate and delayed effects are combined together, they create 2 years of significant improvement after the Academy. The 1st-year improvement was the result of the immediate applied skills and the 2nd-year was the combined result of delayed impact of some immediately applied skills with those skills that could not be immediately applied.

The 3rd year after graduation showed a decrease in organizational effectiveness. As mentioned earlier, Baldwin and Ford (1988) discuss that it is common for training programs to lose effectiveness over time. They go on to add that the best way to slow this decay is to use follow-up training. This slows the reduction and keeps the material fresh in the trainee's mind. Without the use of follow-up training, the data in this study raise the question of whether over time the overall performance would eventually return to its pre-graduation starting levels. While this was outside the scope of this study, some insight can be drawn from the qualitative interviews. While the quantitative data were collected only from 1998 to 2002, the qualitative interviews were collected in 2006, seven years after graduation. The vivid memory the interviewed Chiefs had about their experience at the Academy combined with the similarity between the qualitative and quantitative results leads to the likelihood that a large portion of the training remained with the participants and probably continues to be reflected in higher performance evaluations.

CONCLUSION

This study explored the question "In what ways does the leadership training affect its graduates?" using the Coast Guard's enlisted performance evaluation as an instrument. Using a mixed-method design, a one group pre-posttest study was conducted to determine the effects of

the Coast Guard leadership development program on its graduates' performance. The 4-year study was evaluated prior to the intervention and the three consecutive annual evaluations following graduation. Using a repeated measures analysis of variance statistical analysis and case study interviews, the study shows that the Coast Guard leadership development program resulted in significant increases in overall performance as recorded by enlisted performance evaluations.

REFERENCES

Babcock, P. (2010). Real Cost of Nominal Grade Inflation? New Evidence from Student Course Evaluations. *Economic Inquiry*, *48*, 896-996.

Baldwin, T. T. & Ford, K. J. (1988). Transfer of training: A review of direction for future research. *Personnel Psychology, 41*, 63-105.

Bass, B.M. (1990). *Bass and Stogdill's handbook of leadership: Theory, research and managerial applications*. New York: Free Press.

Behrens, T. R. & Benham, M. K. P. (2007). Evaluating community leadership programs, In K. M. Hannum, J. W. Martineau, and C. Reinelt (Eds.). *The handbook of leadership development evaluation* (pp. 284-314). San Francisco, CA: Jossey-Bass.

Burke, M. J. & Day, R. R. (1986). A cumulative study of the effectiveness of managerial training. *Journal of Applied Psychology, 71*, 232-245.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Collins, D. B. (2002). *The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982-2001*. Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.

Collins, D. B. & Holton, E. F., III. (2004). The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982-2001. *Human Resource Development Quarterly, 15* (2), 217-225.

Conger, J. A. & Benjamin, B. (1999). *Building leaders*, San Francisco, CA: Jossey-Bass.

Dvir, T. Eden, D., Avolio, B. & Shamir, B. (2002). Impact of Transformational Leadership on Follower Development and Performance: A Field Experiment. *Academy of Management Journal, 45* (4), 735-744.

Freifeld, L. (Ed.) (2011). 2011 industry report. *Training*, *48* (6), 22-34.

Glover, R. B. (1996), Why are we ignoring performance appraisal research? *Parks and Recreation, 31* (11), 24-27.

Guba, E. & Lincoln, Y. (1989). *Fourth generation evaluation*. Sage, Newbury Park, CA.

Horner, D., Jr. (1995). Leadership development and why it remains important. *Military Review, 75* (4), 76-87.

Johnson, V. E. (2003). *Grade inflation: A crisis in college education.* New York: Springer-Verlag.

Kirkpatrick, D. L. (1959a). Techniques for evaluating training programs. *Journal of ASTD,* Vol. 13, 3-9.

Kirkpatrick, D. L. (1959b). Techniques for evaluating training programs, Part 2- Learning. *Journal of ASTD, 13*, 21-26.

Kirkpatrick, D. L. (1960a). Techniques for evaluating training programs, Part 3- Behavior. *Journal of ASTD, 14*, 13-18.

Kirkpatrick, D. L. (1960b). Techniques for evaluating training programs, Part 4-Results. *Journal of ASTD, 14*, 28-32.

Lincoln, Y. & Guba, E. (1985). *Naturalist inquiry*. Beverly Hills, CA: Sage.

Lohmann, D. (1992). The impact of leadership on corporate success: A comparative

    analysis of the American and Japanese experience. In K. E. Clark, M. B., Clark,

    and D. P. Campbell (Eds.), *Impact of leadership* (pp. 59-80). Greensboro, NC: Center for

    Creative Leadership.   .

Martin, A. & Ernst, C. (2005). Leadership, learning and human resource management: Exploring

    leadership in times of paradox and complexity. *Corporate Governance*, *5* (3), 82-94.

Mason, J. (1996). *Qualitative research*. Sage: London.

McCauley, C. D. & Hughes-James, M. W. (1994). *An evaluation of outcomes of a leadership

    development program.* Greensboro, NC: Center for Creative Leadership.

McCauley, C. D. & Van Velsor, E. (Eds.) (2003). *The Center for Creative Leadership handbook

    of leadership development (2nd Edition).* San Francisco, CA: Jossey-Bass

Merriam, S. B. (1988). *Case study research in education*. San Francisco, CA: Jossey-Bass.

Northouse, P. G. (1997). *Leadership: Theory and practice*. Thousand Oaks, CA: Sage.

Obiwuru, T. C. Okwu, A.T., Akpu, V.O. & Nwankwere, I.A. (2011). Effects of leadership style

    on organizational performance: A survey of detected small scale enterprises in Ikosi-Ketu

    council development area of Lagos State, Nigeria. *Australian Journal of Business and

    Management Research,* 1 (7), 100-111.

Onwuegbuzie, A. & Teddlie, C. (2003). A framework for analyzing data in mixed-methods

    research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed-methods in social and

    behavioral research.* Thousand Oaks, CA: Sage.

Phillips, J. J. (2003). *Return on investment in training and performance improvement programs

    (2nd ed.).* San Francisco, CA: Butterworth Heinemann.

Roberts, D. (2007). *Deeper learning in leadership: Helping college students find the potential within*. San Francisco, CA: Jossey-Bass.

Rost, J. C. (1993). *Leadership for the twenty-first century.* Westport, CT: Greenwood.

Sessa, V. Matos, C, & Hopkins, C. (2009). Evaluating a College Leadership Course: What do Students Learn in Leadership Course with a Service Learning Component and How Deeply do They Learn It? *Journal of Leadership Education*, *7* (3), 167-200.

Sogunro, O. A. (1997). Impact of training on leadership development: Lessons from a leadership training program. *Evaluation Review*, 21 (6), 713-737.

Stogdill, R. M. (1974). *Handbook on leadership*. New York: Free Press.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Hammond, L. R., Merriman, H. O. & Provus, M. M. (1971). *Educational evaluation and decision-making in education*. Itasca, IL: Peacock.

Swanson, R. A. & Holton, E. F. (1999). *Results: How to assess performance learning and perceptions in organizations,* San Francisco, CA: Berrett-Koehler.

Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present and future. *American Journal of Evaluation*, *22*, 387-394.

United States Coast Guard. (2003). *Chief Petty Officer Academy* (Commandant Instruction 1500.15F). Washington, DC: U.S. Government Printing Office.

United States Coast Guard. (2011). *Enlisted Accessions, Evaluations, and Advancement. (Commandant Instruction M1000.2).* Washington, DC: U.S. Government Printing Office.

United States Coast Guard. (2006*). Leadership development framework (Commandant Instruction M5351.3).* Washington, DC: U.S. Government Printing Office.

Van Fleet, D. D. & Yukl, G. A. (1986). *Military leadership: An organizational behavior perspective.* Greenwich, CT: JAI Press.

Young, M. & Dulewicz, V. (2005), A model of command, leadership and management competency in the British Royal Navy, *Leadership and Organization Development Journal*, *26* (3). 228-241.

Zhang, J. (1999). *Effects of management training on trainees' learning, job performance, and organizational results: A meta-analysis of evaluation studies from 1983-1997,* Unpublished doctoral dissertation, Oklahoma State University, Stillwater.