Academic Parallels from a Military Merit List

Major Jason M Newell, Embry-Riddle Aeronautical University

Major Newell is a Operations Research System Analyst for the United States Army (Functional Area 49). He previously served for 10 years as a Field Artillery Officer in multiple positions and served as cadre for the Embry-Riddle Aeronautical University (Daytona Beach, FL) ROTC Program. He is currently stationed at Fort Leavenworth, KS as an analyst for the The Research and Analysis Center.

Bryan Watson, Embry-Riddle Aeronautical University

Bryan Watson, PE earned his Ph.D. at the Georgia Institute of Technology and his B.S. in Systems Engineering at the United States Naval Academy in 2009. After graduating, Bryan joined the nuclear Navy, serving as a submarine officer onboard the U.S.S Louisville and at the Naval Prototype Training Unit from 2009-2017. Significant milestones include earning the Master Training Specialist Certification (the military's highest instructor accreditation), Nuclear Professional Engineer Certification, two Naval Achievement Medals, the Military Outstanding Volunteer Service Medal, and a Naval Commendation Medal for his work troubleshooting and repairing the Moored Training Ship 635's reactor and electrical distribution faults. Following his transition from active duty, Bryan earned his PhD as a member of both the Computation and Advancement of Sustainable Systems Lab, where he developed a new method for distributed system demand estimation, and at the Sustainable Design and Manufacturing lab, where his work focused on increasing System of System resilience. Bryan's work has been published in the Journal of Industrial Ecology, Journal of Mechanical Design, and IEEE's Systems Journal.

At Embry-Riddle, Bryan's current work is focused on investigating the use of biologically inspired design to increase the resilience of modern system. The goal of their work is more reliable services to users, increased user safety, and increased sustainability for connected manufacturing, energy, and infrastructure systems.

Academic Parallels from a Military Merit List

Abstract

To explore the extent that military training can provide pedagogical insight, this research seeks to discover relationships between classroom (on-campus formative assessment) and "real world" performance (summer training summative assessment) for Army Cadets. This analysis examines data from three years of one program's Army Reserve Officer Training Corps participants. We establish what factors on-campus correlate to Cadet Summer Training success.

The goal of this research is to spur a discussion about different assessments that could reflect real-world performance. Another goal of this research is to establish reasonable criteria for on-campus evaluation that most accurately predicts future Cadet Summer Training (CST) success. Discussions about resourcing and improvement of the on-campus assessment processes are also included. Of the available on-campus measurements, a Cadet's leadership grade is hypothesized to provide the closest correlate to their level of success at CST, both in their final ranking, but also in their leadership evaluations.

Methodology included regression models, including linear regression and logistical regression, to identify relationships, between on-campus evaluations and CST rankings. Different variables were analyzed against each other in multiple iterations and combinations to establish the results. Logistical regression was used to evaluate the impact of academic majors.

The results show that the most highly correlated variable was the Physical Fitness score of each Cadet. The article discusses different possible reasons for this relationship. Results also show moderate to weak relationships between academic performance and any event at Cadet Summer Training. These findings suggest that evaluations outside of a classroom environment could be more effective at predicting future real-world success.

Keywords

Job Preparation, Army, ROTC, Order of Merit List, Linear Regression, Logistic Regression

1. Introduction

An age-old question is how to prepare students most effectively for post-graduation life. In most circumstances, it is difficult to measure student's readiness for post-graduation success with solely academic metrics. Additionally, there are several difficulties in quantitatively assessing subsequent job performance in the workforce. The military structure has built-in comparative tools and evaluation forms that allow for quick, accurate feedback on an Officer's job performance that many civilian institutions lack. The Army Reserve Officer Training Corps (ROTC) program offers a unique opportunity to bridge this gap. The structure of an ROTC program involves multiple assessments of college students in both on-campus classrooms and field exercises (a proxy for workforce performance).

The mission of the Army Reserve Officer Training Corps (ROTC) is to "partner with universities to recruit, educate, develop, and inspire Senior ROTC Cadets in order to commission officers of character for the Total Army".[1] Each student enters the ROTC program as a Cadet, associated both with their academic institution and the Army. They are both students and Army officers in training. Their responsibilities include routine physical fitness in the mornings, military science classes, and weekly labs to practice military skills. All three of these events allow Cadets to be evaluated by active-duty Army Staff with classroom exams, leadership assessments, and military skill evaluations while on their home academic campus. Along with on-campus evaluation, all cadets are evaluated at Cadet Summer Training (CST). On-campus evaluations mirror traditional academic evaluations, while CST is focused on practical skill assessment (mirroring real-world performance). This research examines data to explore relationships between on-campus evaluations and practical exercises at Cadet Summer Training (CST). An assumption of this study is that of the available on-campus measurements, a Cadet's leadership grade will provide the closest correlate to their level of success at CST, both in their final ranking and in their CST leadership evaluations. The contribution of this study is two-fold:

1. The results can help ROTC programs understand which methods of on-campus evaluations are reliable in predicting success (or lack thereof) at CST. Currently, each ROTC program is free to determine which on-campus evaluation methods to use. We hope these findings will help inform individual program's decisions.
2. This study performs a first analysis that compares on-campus training evaluations to real-world performance for ROTC students. The authors' long-term goal is to examine the extent to which the unique requirements of the ROTC program provide fundamental insights into designing assessment techniques to more accurately reflect post-graduation performance. The results of this study provide insights into future work to achieve this long-term goal.

The remainder of the article proceeds as follows. First, the ROTC ranking approach (Order of Merit List), on-campus assessments, and CST assessments are discussed. Next, methodology presents the models that were used to evaluate our research question. The results section of this paper presents the mathematical findings, and the discussion section ties the findings to the contributions.

---

[1] US Department of the Army, ROTC. *About Army ROTC.* Accessed November 8, 2023. https://www.army.mil/rotc#org-about

2. Background

We begin by discussing the role of on-campus and CST evaluation in creating the National Order of Merit List. Next, the on-campus evaluation procedures are presented, followed by the off-campus evaluation (CST).

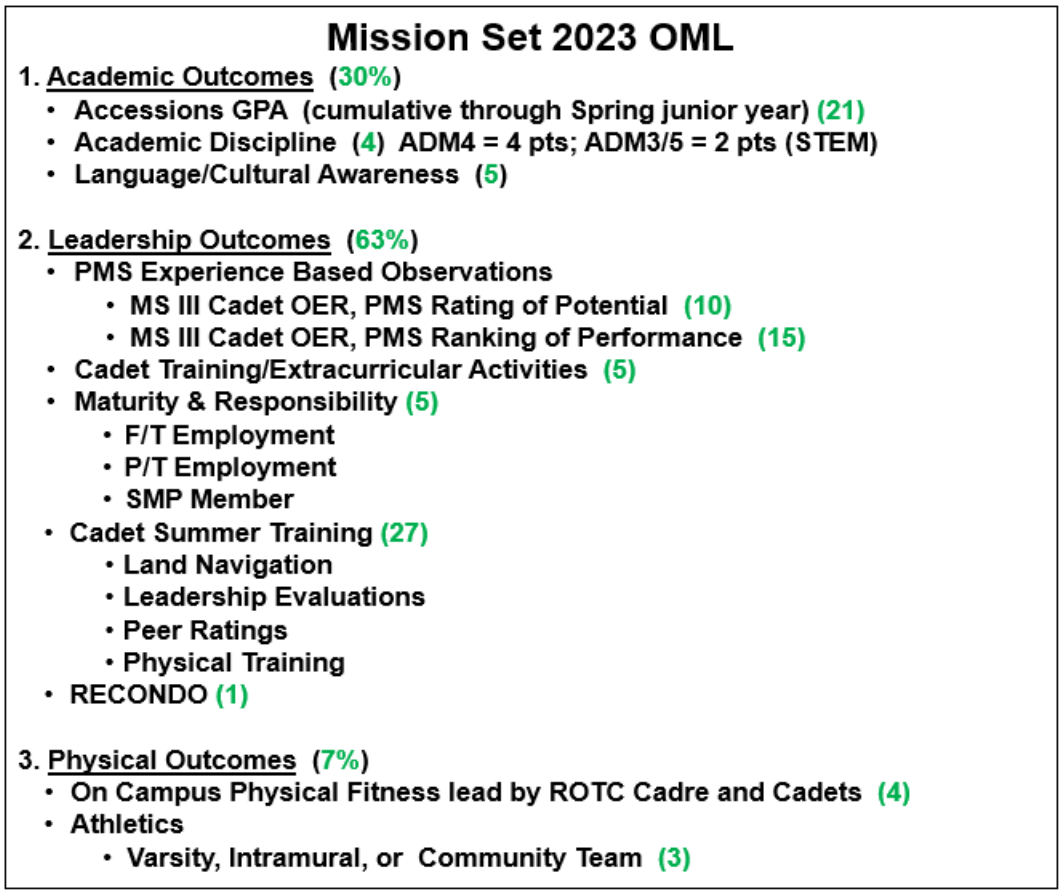
2.1 Order of Merit List

The National Order of Merit List (OML) is the national ranking of all Army ROTC cadets.This list is used in assigning initial job placements as well as ensuring the cadets are eligible to commission into the Army. We begin with a discussion of the National Order of Merit List, as both the on-campus evaluation and CST performance are inputs for its calculation.

The creation of an OML for over 6,000 Cadets nationwide each year is a dynamic process. Each year, small modifications are made to the ranking process. However, most of those modifications simply change the weighting used to combine the criteria into an aggregated metric. The criteria themselves typically remain the same and are focused around five categories, that are described in Figure 1: Academic GPA, Physical Fitness Score (most recently evaluated by the Army Combat Fitness Test (ACFT)), extracurricular activities and community involvement, and their on-campus OML their junior year, and their ranking at Cadet Summer Training. The most heavily weighted are the on-campus OML and CST ranking (Figure 1). Research to refine the OML at the national level is conducted at the United States Army Corps of Cadets (USACC) Headquarters located at Fort Knox, Kentucky. There is an Operations Research unit that has years of data collected and is constantly analyzing separate variables at the national level following the guidance of assigned Major General (2 star). While the weights themselves for the OML are available the data used to determine the national rankings are unavailable for this study.

An example of how the variables used to calculate the National OML may change, such as in 2021, when the Physical Fitness Score was not included (due to COVID-19 restrictions). The Physical Fitness Score was removed, and an extra 5% and 10% were added to the on-campus OML and CST OML, respectively. Now that the ACFT is fully in effect, the percentages have shifted back to 25% each. Figure 1 shows an example of OML weighting. All events except “Cadet Summer Training” and “RECONDO” take place at the host campus.

Figure 1: OML criteria for Cadets Commission in FY2023 [2]



2.2 On-Campus Evaluation

The rankings given for on-campus Assessments are very subjective. There is no national guidance on how to create the on-campus OML. Cadets are still ranked against their peers, but only on their own campus. Typically, the criteria includes the following (but vary by campus):

- Academic criteria (either overall GPA, or only military science grades).
- Physical criteria (sometimes consists of only ACFT score, but other events like ruck marches, varsity sports participation, or other factors may be included).
- Leadership evaluation based on garrison and field training leadership opportunities.

Though there is this general understanding of what should be included at a minimum, the decision on how to rank Cadets ultimately rests in the Professor of Military Science (PMS) and the MSIII (Military Science III) instructor at the host campus. The PMS is the highest-ranking military person on university staff and the MSIII instructor is the teacher and coach

[2]Johnny K. Davis, MG, personal communication, March 10, 2022.

for all junior Cadets. For both on-campus evaluations and CST ranking, Cadets are ranked on a five-tier scale as described below:

"O" (Outstanding) – Top 15% of cohort
"E" (Excellent) – 15%-50% of cohort
"P" (Proficient) – 50%-85% of cohort
"C" (Capable) – 85%-100% of cohort
"U" (Unsatisfactory) – Cadet did not meet the requirements

2.3 Off-Campus Evaluation

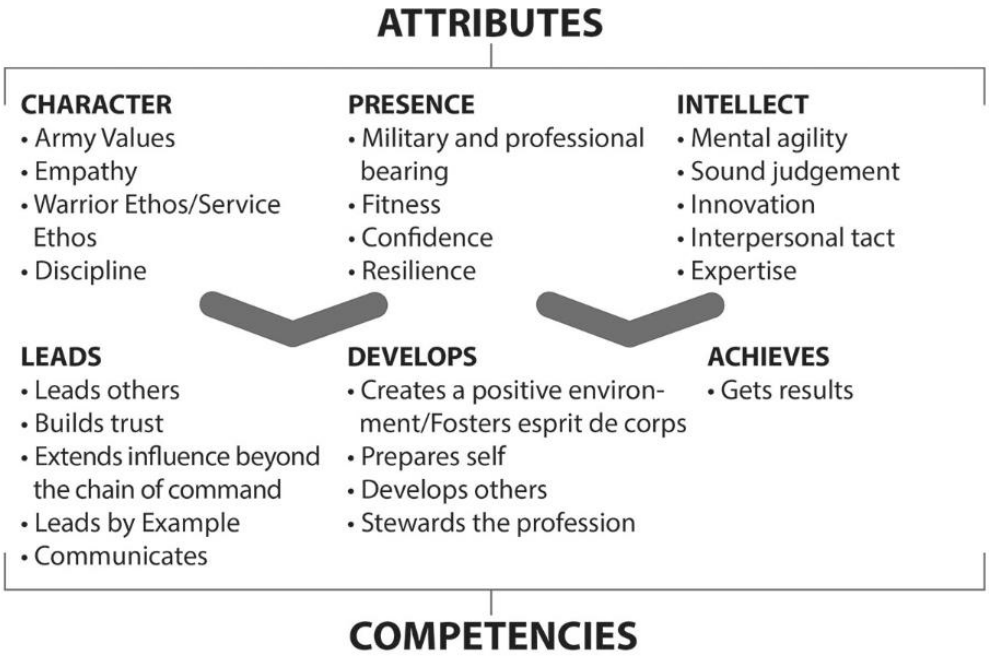
CST is the culminating exercise of pre-commissioning education. Each summer approximately 6,000 Cadets converge on Fort Knox, Kentucky from universities around the country. Cadets are divided into regiments (separated by time), Companies, and finally Platoons (those last two only separated by Active Duty evaluator supervision). Each Platoon consists of about 40 cadets divided into four squads. Over a period of 40 days, each Cadet is tested on individual Warrior Tasks and leadership abilities. Their CST ranking compares them to the other 40 personnel in their Platoon. Like the national OML, the rankings at CST are dynamic and based off several metrics. The yearly changes to calculating the CST rankings are typically small percentage shifts.

Table 1: Garrison and Field Events during CST

Garrison Requirements	Field Requirements
Health Physical (to include meeting height and weight requirements)	One Platoon Leader or Platoon Sergeant Leadership rotation
Tactical Combat Casualty Care (First aid)	One Squad Leader Leadership rotation
Army Combat Fitness Test	Offensive Operations (attack, movement to contact, ambush, or raid)
Call For [Indirect Fire]	Defensive Operations (area or mobile defense)
Tear Gas Chamber	8 mile formation ruck march
6 mile timed ruck march	12 mile formation ruck march

The assessments begin with a "Garrison" schedule (Table 1). This section contains individual skills such as a Physical Fitness Test, Basic Rifle Marksmanship, and Land Navigation Testing (both written and practical). Other tests such as Artillery Call for Fire, Chemical, Biological, Radiological and Nuclear Defense, Confidence Courses, and Grenade Courses are evaluated on Go/No-Go criteria and not included in this research. Throughout these tasks, Cadets are given leadership roles at the Platoon and Squad levels and evaluated by Active Duty Officers and Non-Commissioned Officers (NCOs) on their abilities in accordance with Army Doctrine Publication (ADP) 6-22 and the Army Leadership Requirement Model (Figure 2). After the conclusion of these events, Cadets begin a tactical portion running scenario-based exercises where they are tested on their tasks and leadership abilities. Each leadership evaluation is conducted using a SOAR Card (Situation, Observations, Assessment, Recommendation) on a four-tier grading system. While the ultimate grade is subjective, the tiers are: Excellent, Proficient, Capable, and Untrained.

Figure 2: Army Leadership Requirements Model (reprinted from ADP 6-22)[3]
This table lists the 26 Attributes and Competencies Army leaders can be evaluated on. The top row is Attributes which describes what the leader is as a person. The bottom is Competencies which are the skills the leader displays. ADP 6-22 provides definitions of each.



3. Data Collection and Methodology

The data for this research were collected on 162 junior level Cadets enrolled in the Embry-Riddle Aeronautical University (ERAU) Army ROTC program, from the period August 2020 – May 2023 (three different groups). The data are from two sources – the ERAU ROTC Program (on-campus performance) and United States Army Cadet Command (USACC) (real-world performance). The ERAU ROTC program compiles the on-campus rankings and posts it on their share portal. It is unclassified and available upon request. USACC compiles their data and puts it on their share portal. The USACC data is also unclassified. Therefore, there are six different files of data: three years of on-campus evaluations, and three years of CST performance.

It is important to note that ERAU ROTC is the largest of 39 programs in the Southeastern United States (including Puerto Rico), has been nationally recognized for their CST performance over the last five years, and is consistently a top performer in the Southeastern United States. Therefore, although this study uses a large dataset for analysis, the results are not representative of all ROTC programs. While most of the data is focused on overall campus performance and overall performance results at CST, there are also individual tasks and specific events considered

[3] Department of the Army, ADP 6-22, *Army Leadership and the Profession* (Washington, DC: Government Printing Office, July 2019), 29.

that may influence their overall ranking. Ancillary aspects like academic major are also considered during analysis. Table 2 lists the variables that will be used for this research.

Table 2: List of Variables Evaluated

On-campus	Major	The undergraduate academic major the Cadet is pursuing.
	Major_Tier	Tier that academic major is included in using the Navy ROTC Tiered system. (See section 3.5)
	Class_Rank	The overall ranking in regards to on-campus performance (lower number is better)
	Leadership	The grade given to the Cadet in regards to their leadership performance; assessed on a 1-100 scale (higher number is better)
	Academic	The grade received during Military Science III classes (higher number is better)
	A_ACFT	ACFT score on-campus (out of 600)
CST	Campus_BRM	Dummy variable establishing whether or not the Cadet conducted Basic Rifle Marksmanship on-campus or not.
	C_ACFT	ACFT score at CST (only graded on "GO" "NOGO" criteria)
	LNWT	Score received on the written Land Navigation test
	LNEX	Score received on the practical Land Navigation test (out of 6 possible points)
	Camp_BRM	Score received on Basic Rifle Marksmanship at CST (10 means Expert qualification on first attempt; score decreases with status or additional attempts)
	G1_Pos	Garrison Leadership Position (Squad Leader, Platoon Sergeant or Platoon Leader)
	G2_Pos	
	G1_Eval	Garrison Leadership Evaluation (Excellent, Proficient, Capable, or Untrained)
	G2_Eval	
	F1_Pos	Tactical Leadership Position (Squad Leader, Platoon Sergeant or Platoon Leader)
	F2_Pos	
	F3_Pos	
	F1_Eval	Tactical Leadership Evaluation (Excellent, Proficient, Capable, or Untrained)
	F2_Eval	
	F3_Eval	
	Peer	Peer ranking within their squad, based off an average of others rankings in the squad. (lower number is better).
	Overall	Overall evaluation of performance at CST (as outlined at the end of the Background portion) (OEPCU)
	PLT_Rank	The numerical rank within the Platoon (lower is better)
	PLT_Size	The number of Cadets in the Platoon
	Percent	The percentage the Cadet placed within their Platoon (Lower is better)

To establish relationships and examine our research question, five statistical tests were performed.

3.1. *Individual Task Relationships*

The first evaluation of models is a series of linear regression models to compare the relationship between academic or leadership grades on-campus, to individual tasks at CST. This allows comparison between traditional classroom teaching (e.g. slideshows on ballistics, map reading) and on-campus practical exercises to CST evaluation.

3.2 *Overall Rank Relationships*

The second regression model establishes the linear relationships between all four of the on-campus evaluations (except Major) and compare them to overall CST success ("Percent"). This can indicate which method of on-campus evaluation most correlated with success at CST.

3.3 *Peer Evaluation Relationships*

The final linear regression model examines if any of the on-campus evaluations correlate with CST peer evaluations. Peer evaluations do not have a heavy weight on the overall CST ranking, but it is important to gain the respect of your peers, especially in the environment CST creates. Thus, peer evaluations provide a different approach to measure "real world" success.

3.4 *SOAR Card Relationships*

Since SOAR card evaluations are an assessment of practical leadership abilities, it is important to evaluate if certain on-campus activities impact Cadet leadership. This model is constructed as a logistic regression model using the same on-campus variables, but establishes dependent variables based off the "E, P, C, U" standard. Logistic regression must be used to individually analyze each SOAR card possibility since it is a discrete variable with only four options.

3.5 *Academic Majors*

The final model is a multi-variable regression using dummy variables to see if cadet major correlates with overall CST outcome. It is important to distinguish if certain academic majors perform differently and if the findings should consider different academic concentrations. It might be important to group Cadets by academic concentrations because different programs often explore their material differently. For example, humanities majors are more likely to include readings and group discussions; while engineers are more likely to solve problems or visually examine testbeds. By separating academic majors, we can see if one group has different results, suggesting a potential impact of different teaching techniques.

The levels of academic major was defined using the Navy ROTC Tier System.[4] This system groups majors into three tiers. Tier 1 includes engineering courses that directly relate to military operations (Systems Engineering and Civil Engineering were the only engineering disciplines excluded). Tier 2 includes all sciences and engineering majors not listed in Tier 1. Tier 3 is all the humanities, and most importantly for ERAU specifically, Homeland Security. ERAU ROTC has a large number of Cadets enrolled in the Homeland Security program and many of those Cadets are high performers focused on a long career in the military. Using a different sorting method that put Homeland Security in a separate bin may produce different results for the ERAU program. Army ROTC does not have a tiered system other than "STEM" and "Non-STEM". The US Navy ROTC System is the closest, DOD endorsed tiering system.

4. Results

Summary

For our results, we use the following descriptive statistics (R^2) baseline [5]:

.64 to 1.00 = Very Strong

.35 to .63 = Strong

---

[4] US Department of Defense, Diversity, *Navy ROTC Resources.* Accessed November 8, 2023, https://diversity.defense.gov/Portals/51/Documents/Resources/Docs/ROTC/Navy%20ROTC.pdf

[5] James D. Evans, *Straightforward Statistics for the Behavioral Sciences*, (Brooks/Cole Publishing Company, Pacific Grove, CA, 1996), 146.

.16 to .35 = Moderate
.04 to .16 = Weak
.00 to .04 = Negligible to very weak

The findings were that the majority of the criteria had moderate or weak relationships, especially between any on-campus evaluations and individual skills tests. The highest R^2 value was 0.33, but the majority were lower than 0.2. The differences in the CST overall rank relationships were large enough to identify that academic performance had a much weaker relationship than leadership and ACFT scores, but overall, all of the relationships were still weak. All but two *p*-values were below the significance threshold of 0.05.

Table 2 presents the overall results. The first column annotates which section of the paper these models address. The second and third column identify the dependent and independent variables for the regression models. The final columns are the model parameters. R^2 is the R^2 for that regression model. After that is the *p*-value and the final column easily depicting whether or not the *p*-value is below the 0.05 threshold or not. Sections 3.4 and 3.5 did not use linear regression, so only values for the first three models run are included in the below table.

Table 2: Results of Linear Regression Models

Section	Dependant	Independent	R^2	Pvalue	Pval Val
3.1 4.1 5.1	LNWT	Academic Grade	0.04	1.50E-02	Y
	LNEX	Academic Grade	0.014	1.39E-01	N
	BRM	Academic Grade	0.026	4.20E-02	Y
	LNWT	Leadership Rank	0.165	8.38E-08	Y
	LNEX	Leadership Rank	0.002	5.28E-01	N
	BRM	Leadership Rank	0.14	7.43E-07	Y
3.2 4.2 5.2	CST %	Campus Rank	0.32	5.02E-15	Y
	CST %	ACFT	0.31	1.34E-14	Y
	CST %	Academic Grade	0.07	9.94E-04	Y
	CST %	Leadership Rank	0.22	2.21E-10	Y
3.3	Peer Eval	ACFT	0.16	1.05E-07	Y
4.3	Peer Eval	Leadership Rank	0.17	2.73E-08	Y
5.3	Peer Eval	Academic Grade	0.08	1.40E-04	Y

4.1. *Individual Task Relationships*

In these regression models, higher leadership rankings tended have the strongest relationship to success in individual tasks. All regression models run with Academic grade as the independent variable resulted in an R^2 less than 0.05, with the lowest being 0.03. The p-values also are relatively close to the 0.05 threshold, and one regression model (Land Nav practical exercise- LNEX) does not even meet that threshold.

However, the models between two of three individual tasks and Leadership Ranks resulted in R^2 values of 0.16 and 0.14. While only a weak relationship, these are still considerably higher than the models regarding academic grades. Overall, Land Nav practical exercise (LNEX) did not seem to display any level of relationship to any on-campus variables.

4.2 *Overall Rank Relationships*

Analyzing CST percentage resulted in a moderate correlation for two of the models (Campus Rank and ACFT). Overall campus rank had the highest R^2, Test statistic, and lowest *p*-value. This should not necessarily come as a surprise, but as we will discuss a little later, it does increase confidence in ERAU's evaluation standards. The ACFT score also has a relatively high R^2, and at 0.32, is only one tenth lower than the overall campus rank. Leadership ranks were also moderately correlated (R^2=0.22, p<.01). Once again, Academic grades showed a weak correlation with the overall ranks (R^2=0.07, p<.01).

4.3 *Peer Evaluation Relationships*

Peer Evaluations seemed to have little effect, based off changes in any of the variables. Leadership rank and ACFT score had nearly identical moderate correlations while Academic grades were weakly correlated.

4.4 *SOAR Card Relationships*

Logistic regression for the SOAR cards yielded very low levels of reliability. The confusion matrix had a very low rate predicting many "Cs" and "Es". The number of "Ps" was much more closely predicted, but still fell short of any discernable relationship. In most cases, the precision for "Cs" was zero (when rounded to two decimals) and "Es"were always below 0.5. "Ps" ranged from 0.5 to 0.65, which is a higher prediction, but also with a bell curve for grading, "Ps" were simply more common, and likely resulted in more precision. but The model that did stand out, was that the combined class rank (the professors weighted overall rank) had a much higher level or predictability than any of the individual metrics.

4.5 *Academic Majors*

Breaking down the cohorts by Academic Majors resulted in very poor correlations. The combined R^2 for the multi-variable regression model was 0.029 (p>.05).

5. Discussion

5.1. *Individual Task Relationships*

Overall, it is very difficult to predict performance in CST individual tasks because scores for events like rifle marksmanship and Land Navigation often have too many external factors to simply link them to on-campus activities. On-campus leadership rank has higher correlation values than any other on-campus metric, but is only moderately correlated to high scores on individual tasks. It is shocking that there is such a low level of relationship between academic grades and the land navigation written test, as typically, the score for the written test is incorporated into their military science grade. However, for the rest of the models, the likely biggest contributor lies in how well the Cadet meets the Attributes and Competencies (as discussed in ADP 6-22 and the Army Leadership Requirements Model). Expertise is only one of those 26 Attributes and Competencies. For example, if a Cadet is excellent at communication, developing others, and military bearing, but shoots poorly on their marksmanship attempts, then it does not impact their leadership results on-campus. There still usually needs to be a baseline of proficiency in individual tasks to gain the trust of others, and that may explain why the relationship is stronger than for academic grades.

5.2 *Overall Rank Relationships*

Two things stand out in the findings from these models. As already discussed, it is encouraging to see that the combined class rank has the strongest relationship of all on-campus variables. This result likely indicates that the weights that the program used for the variables were appropriate to be able to predict those who would succeed at camp. Additionally, the physical

fitness test having such a high relationship with CST performance could indicate that a level of dedication and desire is necessary for succeeding. The physical fitness test is an event where someone can achieve the minimum standard and still earn a commission, so the only reason to achieve a higher score is to go beyond what is required. Therefore, those who have shown the ability to put in the work and achieve more than the minimum requirements, could reflect a higher level of dedication to the craft. This genuine dedication could produce better results in other areas, but cannot be directly proven by this study.

5.3 *Peer Evaluation Relationships*

The weak correlations between peer evaluations and any other metrics means that it is nearly impossible to predict how personalities may interact off-campus, and training plans should not be modified due to the results of this study. In fact, Cadet Command has already indicated that future CST peer evaluations will only be used for self-development purposes, and not included in the final weight of CST OML. Because peer evaluations can be influenced so heavily by a few interactions, rather than actually achieving an objective grade, Cadet Command has taken out the numerical effect. It has been replaced with comments from their platoon mates, to hopefully help improve social and professional interactions in the future. Higher echelons of the Army making this adjustment further supports our recommendation.

5.4 *SOAR Card Relationships*

As previously discussed, the grading criteria for the on-campus leadership rank is based off the ALRM in ADP 6-22[6]. The same grading criteria is used at CST. Our finding that leadership rank was unable to predict the outcome of SOAR cards was quite surprising. The likely reason for this disparity is the different graders at CST. Each cadre has separate attributes or competencies that they deem more important than others. So, where home campus cadre may think of a Cadet highly for one aspect while overlooking deficiencies in others, the cadre at CST have different biases or expectations. Additionally, the stress levels on-campus and at CST are different. The stressors on-campus are driven by upcoming assignments and peer relationships. At CST, the stressors involve less sleep, lower caloric intake, and the realization that CST performance has serious implications on their future. The duration of the events is also drastic. On-campus evaluations span an entire year and CST is 40 days. This causes additional stress and urgency to for each evaluation. Therefore, some Cadets may adjust to the stress differently and perform may change.

The application to other academic areas we believe lays mainly in those environmental conditions. Under ideal circumstances in both settings, we suspect that the relationship between leadership ranks and evaluations would likely improve. Conversely, academic and ACFT scores have so little to do with handling that kind of stress, that they would likely remain the same. It is important for instructors to keep their students in the "developmental zone" [7] to ensure preparation for future success. Moving to either extreme, like the difference between campus life and CST, will likely lead to increased frustration and less reliability in performance.

5.5 *Academic Majors*

The poor predictive power of academic majors was not necessarily surprising. This is for several reasons, including the host university used in this study. A number of high performing

---

[6] ADP 6-22, 29

[7] Bill McCollum and Matthew Broaddus, "Leader-Imposed Stress and Organizational Resilience"; reprinted in US Army Command and General Staff College, *L100 Book of Readings* (Fort Leavenworth, KS: USACGSC, July 1992), 273.

students are specializing in areas relating to aviation and aerospace, and the criteria used for this study assigned all non-engineering sciences to the second tier. The third tier also contained a faction of high performing Cadets majoring in Homeland Security. While this can be incredibly applicable to their future, it was not considered Tier 1 by the ranking system used.

Overall, it is likely not important to consider majors when trying to predict success in an environment that relies so heavily on personal interaction and practical exercises. All academic majors seemed to perform more in line with how their military science performance than when divided by major. Additionally, the Army does not set quotas for majors within ROTC. In the past, Navy had set quotas for the ROTC programs using the tiered system, but since the Army does not, this provides further indication that the Army does not believe that academic major impacts real-world job performance.

6. Conclusion

The primary goal of this study was to identify relationships between on-campus evaluations and real-world performance. Our long-term goal is to find ways to relate these findings to other academic endeavors. Our hypothesis was that of the available on-campus measurements, a Cadet's leadership grade would provide the closest correlate to their level of success at CST, both in their final ranking, but also in their CST leadership evaluations (SOAR). This hypothesis was not supported by our findings. Rather, physical fitness scores provided a stronger correlation to CST OML and no meaningful correlations were discovered for SOAR card results. Surprisingly, we found that the academic grade in the military science classes was the least reliable in terms of predicting future success. Leadership ranks and physical fitness were much better predictors and should remain more heavily weighted in future on-campus evaluations. In order to establish a link with engineering education and apply these findings to other areas, instructors should attempt to develop methods for students to find the desire and motivation put in extra work as students with those traits correlate with a higher likelihood of success.

As far as the secondary goal is concerned, it is impossible to assign exact weights based off this study because of the low correlations found. Overall, our results indicate that the grading criteria already established is more successful at predicting success than any individual metric. This is encouraging and may point to the past success of the program in general. This finding also aligns with prior work that holistic success is multidimensional and requires considering the whole human experience[8] (cite. However, through the period that data was collected, academic grades carried the same weight as ACFT score. Our recommendation would be slightly flex ACFT weighting to make it more valuable than academic score. The sacrifice to compensate for that increase could come from decreasing the weights of either the academic score or leadership rank. Either case would improve the predictive power of on-campus OML.

Future research should be dedicated to expanding the pool of Universities. Using only one program – especially the largest in the region – does not allow the ability to look at different curriculums, impact of program size, and prestige of the University itself. If more Universities were able to be analyzed, it would open up possibilities to explore more evaluation criteria (based off what each program decides is valuable criteria) and have an increased diversity of Cadets. Future research could also build off the simple models used in this study and more data points

[8] Gesun, J. S., Major, J. C., Berger, E., Godwin, A., Jensen, K. J., Chen, J., & Froiland, J. M. (2021). A scoping literature review of engineering thriving to redefine student success. *Studies in Engineering Education*, *2*(2).

could open the door for more sophisticated machine learning models. Our study, however, provides an important first step in leveraging the unique curriculum of ROTC students to explore the question: How does on-campus (formative assessment) performance correlate with real-world (summative assessment) performance.