MOTIVATING AIRMEN TO ENGAGE WITH TECHNICAL EDUCATION: EXPERIMENTATION AND ANALYSIS USING MODERN GAMIFICATION TECHNIQUES

THESIS

Landon G. M. Tomcho, 2d Lt, USAF

AFIT-ENG-MS-19-M-061

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENG-MS-19-M-061

MOTIVATING AIRMEN TO ENGAGE WITH TECHNICAL EDUCATION: EXPERIMENTATION AND ANALYSIS USING MODERN GAMIFICATION TECHNIQUES

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Software Engineering

Landon G. M. Tomcho, BS

2d Lt, USAF

March 2019

AFIT-ENG-MS-19-M-061

MOTIVATING AIRMEN TO ENGAGE WITH TECHNICAL EDUCATION: EXPERIMENTATION AND ANALYSIS USING MODERN GAMIFICATION TECHNIQUES

Landon G. M. Tomcho, B.S.

2d Lt, USAF

Committee Membership:

Lt Col Mark A. Reith, Ph.D.
Chair

David Long, Ph.D.
Member

Lt Col Alan C. Lin, Ph.D.
Member

Lt Col Eric Trias, Ph.D.
Member

AFIT-ENG-MS-19-M-061

Abstract

The development and integration of computer systems into today's society and the subsequent growth of cyber as a warfighting domain has led to changes in military and civilian conflict. Several traits unique to cyber, including disruption and fast pace of change, has led to issues never before seen in the military environment, especially with educating and training. A new approach that leverages crowd-sourced content has been proposed. This approach relies on motivating military members to voluntarily engage with technical (cyber) education.

The application of gamification, a design practice aimed at increasing user engagement by targeting core motivators in humans, in the military context is presented in this paper. The adaptation and evaluation of unique game elements onto the platform is also discussed. A human-subject study involving a survey and engagement-tracking experiment is implemented. Results are analyzed using visualization software and a novel framework we created.

We then present results explaining what core drives motivate military members on average and within subgroups. We also show that engagement data can be attributed to motivation levels. Finally, we present recommendations to military leadership and education platform designers based on our findings before discussing ideas for future work.

Acknowledgments

I would like to thank everyone who has been a part of my journey at AFIT. I appreciate the many hours sacrificed by my research sponsor Col Mark Coggins, my entire committee, and my advisor Lt Col Mark Reith, to provide guidance and direction in ensuring that my efforts on this thesis were fruitful and worthwhile. I would also like to thank my family, friends, and peers for their support during this eventful and demanding eighteen months. I truly learned a lot; not only about my research topic and about being a better officer in our Air Force, I also learned a great deal about myself.

Landon G. M. Tomcho

Executive Summary

In this thesis document we discuss foundational work that pointed out specific shortfalls of current Air Force cyber education and training. These include: currency, by the time educational and training content is pushed down from admins to users it is out of date; complexity, most content is either too difficult or too basic to appeal to all users and it is often not from a perspective that relates to the military member; and scalability & breadth, specific content that is useful and free of extra information is not easily delivered to a large number of users in the appropriate timeframe. The Cyber Education Hub (CEH), a platform inspired by crowd-sourced sites like YouTube and curated platforms like Netflix, was developed as part of an attempt to address some of these issues. The main efforts of this thesis deal with motivating military members to voluntarily use this platform such that the critical mass of contributors and consumers is reached.

A human-focused design technique known as gamification borrows from successful practices mastered by the gaming and social media industry to motivate target users to engage with products and platforms. The Octalysis Framework (Chou 2015) is heavily utilized throughout this thesis. This framework breaks down human motivation into eight core drives. Level I of the framework is used for initial analysis and design while higher levels deal with developing a strong experience for different types of users at various stages of the 'game.'

Unique game elements applied to the CEH include the Cyber Topic Map and Knowledge, Skill, and Ability (KSA) Trees. The Cyber Topic Map is a way to organize educational and training content in a visual manner that allows users to orient themselves, navigate to new topics, explore the vastness of the cyber domain, and discover new things along the way. KSA Trees are used to present users with tasks and challenges that allow users to develop their personal KSAs. KSA Trees add a visualized progression dynamic to the CEH and are more restrictive than the Topic Map, but more empowering than a strictly prerequisite-style learning experience.

We hypothesized and concluded that implementing gamification on the CEH via basic and specialized game elements help raise user engagement with the CEH website above that of similar online military education platforms. A human-subject study including surveys and an engagement tracking experiment were utilized for data collection.

During analysis we reached several findings to answer our research objectives. The Topic Map was primarily used when participants were contributing content to the CEH website; users hoped that they would gain more views because others would more easily find and view their content, but the Topic Map was hardly utilized in that manner. KSA Trees helped motivate users to learn more if the user was already interested in the related topic; additional specialized KSA Trees should be developed. Military members are most commonly and most strongly motivated by the following Core Drives: Development & Accomplishment, which deals with reaching goals and visualizing your progress; Social Influence & Relatedness, which involves interacting and competing with friends and peers and engaging with things that relate to your perspective; and Curiosity & Unpredictability, the Core Drive associated with not knowing what will happen next or what is available in an experience.

We also found differences in motivation levels between different subgroups, primarily between gamers and non-gamers. Also, while these users did not report a difference in their overall reaction to the CEH, non-gamers were most positively affected by the Topic Map an KSA Trees in terms of enjoyment and motivation to learn; unfortunately, non-gamers only made up 29% of our population. Finally, we showed that differences in engagement with the CEH platform can be attributed to differences in motivation. In summary, we can reasonably believe that targeting human motivators and user desire in design can lead to more engagement with your product, platform, or experience, even in the military environment.

Table of Contents

List of Figures

List of Tables

Page

MOTIVATING AIRMEN TO ENGAGE WITH TECHNICAL EDUCATION: EXPERIMENTATION AND ANALYSIS USING MODERN GAMIFICATION TECHNIQUES

I. Introduction

1. Motivation

Throughout history, warfare has evolved alongside technology and human innovation. This holds true today as cyber warfare becomes increasingly present in civilian and military conflict. Several characteristics about this young domain and the associated warfare are unique, including the low barrier-to-entry into the cyber 'battlefield.' While more conventional warfare takes place on a physical landscape such as land, sea, air, or space, cyber warfare takes place in a complicated man-made realm where nearly *everyone* may be involved either as an attacker, defender, or target. In the United States (U.S.) military, every warfighter's daily actions or inactions can induce cyberspace events.

The emergence of warfare in this domain comes with unique challenges including educating and training Airmen not only to be compliant to cybersecurity best-practices, but to be resilient to ensure a fully integrated warfighting force (Reith 2016). Traditional methods of education and training in the military environment are not keeping pace with the disruption and rapid pace of change in the cyber domain. Finely-curated content delivered via a top-down approach often in a classroom setting has its place, but new solutions must be explored to meet the force's demands in cyber warfare. In order to reach more Airmen at a quicker pace and deliver the content they need, we must look into a new approach that taps into individual Airmen's talent and personal experience to spread knowledge to keep Airmen up to speed with cyber (Reith et al. 2018).

The Cyber Education Hub (CEH) website is part of the Air Force Institute of Technology's (AFIT) recognition of the issues we are facing in this domain. This platform is built off of the framework detailed in (Reith et al. 2018) where crowdsourcing is proposed to leverage the benefits of a diversity of content within a multi-modal educational experience. In order to achieve long-term success the CEH must pass the tipping-point (Gladwell 2002) and reach critical mass for content producers and consumers on the platform. This CEH is also designed for voluntary-use and motivating Airmen to engage with the platform is a challenge we seek to unravel in this work. The efforts of this paper also directly align with the Continuum of Learning concept (Roberson and Stafford 2017) where Airmen are encouraged to voluntarily seek education and become life-long learners.

Motivating users to engage with products, platforms, and experiences is a practice that has nearly been mastered by the social-media and gaming industry. Applying these fruitful design principles and techniques in other contexts is known as gamification (Chou 2015). The CEH is the vehicle for this thesis research, but the more general problem we seek to explore is: *"How can modern gamification principles and techniques be effectively utilized in a military context?"* Findings about motivating military members should not only be of interest to CEH developers, but all military leaders; military leadership is the art motivating others to complete the mission. Specific research questions are presented in *VII.* on page 100 and answered on page 179.

2. Research Approach

This thesis explores current gamification theory and walks through some of the initial design applied to the CEH platform before experimentation and analysis is performed to answer specific research questions. We discuss taking the website from its base features of allowing users to upload and search for content to a more complete user-experience. Driving software requirements and applying some foundational game elements that appear in the minimum viable product are covered. Adaptations of larger game elements such as maps and skill trees and their implementation on the CEH are also detailed. The 'Cyber Topic Map' element is an alternative way to present content that allows a user to orient themselves, visualize the cyber universe, find topics that they are interested about, and navigate there while potentially discovering new information during their exploration. The KSA (knowledge, skill, and ability) Tree (KSAT) game element presents uses with tasks and challenges focused on increasing their KSAs. The KSAT allows users to visualize their unique development that occurs as part of a hybrid directed and self-guided learning experience.

With our initial design implementations, we aimed at increasing user engagement with the CEH beyond what it otherwise may have been. From the software-engineering perspective, exploring techniques that increase user happiness with your product is important, and potentially crucial when it comes to this type of education. When software is unusable or inconvenient users will explore other options, that are often less secure or unapproved, to solve their problems. In our scenario, by providing a platform that users enjoy, they will be more likely to share potentially-sensitive content and seek information about critical systems on our secure system instead of on a public website such as

YouTube. When boiling it down, the first articles of this thesis discuss our aim at increasing user engagement by targeting human core-motivators in our design. This may have some readers going one step further and begging the question “Can we predict or quantify how much more users will engage as a result of these gamification implementations?”

In order to begin unraveling the question presented above, we look at another question which may springboard us forward in our investigation: “How can differences in engagement with a platform be attributed to differences in motivation?” We present a human-subject study where we seek to answer this question and gain other insights. The study includes a survey to gauge users’ motivation models based on a gamification framework and an experiment to track user engagement with different facets of the CEH website.

3. Assumptions and Limitations

A primary goal of the CEH is to increase the cyber KSAs of Airmen. While examining the best teaching practices and modalities warrants its own focused research, we assume that learning occurs when the users are present on the platform such that increasing user presence increases total learning. Additionally, learning may not always lead to changes in behavior, potentially limiting our efficacy.

The research in this thesis hinges heavily on the Octalysis Framework (Chou 2015). We assume that this framework is effective for analysis and design and that insights gained can enhance user motivation and increase engagement with the platform. When gauging the motivation and enjoyment levels of participants with respect to the framework, we assume that user responses are reflective of their actual behavior/feelings and that the survey questions are adequately designed. We also assume that the limited data collected from our study is representative of a greater population of target users.

Other assumptions are that the applied CEH design decisions positively affect learning and do not negatively affect military structure and discipline. Empowering Airmen to continually learn and contribute (and hopefully have fun while doing so) may have unforeseen impacts in a war-fighting environment based heavily on command and rank-structure. For example, could the exposure to these new methods lead to a lack of structure, discipline, and operational focus or less respect for regulations and authority?

4. Contributions

Major contributions of this work include:

- We show that statistically significant differences in engagement between groups could be attributed to differences in Motivation Levels with a platform using a novel framework.
- Research-based software requirements/recommendations are given to Cyber Education Hub developers.
    - CEH development using an Agile software engineering approach.
    - Design using gamification design principles utilizing frameworks such as Octalysis (Chou 2015).
    - Creation and application of Topic Map and KSAT game elements within an operational military context.
- Gauged motivation of military members based off of the Octalysis Framework and translated findings into specific recommendations for AF leadership.
- Drove requirements for engagement tracking database and developed engagement visualization software that presents resulting data in a meaningful way. Program to be handed off to CEH development team.
- Developed multiple complex algorithms for use on CEH Topic Maps and KSATs. Users select what they 'want' to learn, algorithms tell them what they also 'need' to learn to bridge gaps between 'wants.'

5. Document Structure

This thesis document primarily follows the scholarly article format. Each article contains their own abstract, introduction, analysis, etc. as appropriate. Some transition pages are included to guide the reader between articles. Article supplements and extracts are also included throughout the document and some additional material is also available in the Appendix. There is also overall Introduction and Final Conclusions sections to summarize the thesis document. A preview of each article is provided below.

IV. Scholarly Article: Engaging Airmen with Cyber Education and Training: Designing a Platform Using Gamification (Tomcho and Reith 2019) is published in the *Journal for the Colloquium for Information Systems Security Education.* This article builds on the Framework presented in (Reith et al. 2018). Some ideas from this article are further refined in *V. Scholarly Article: Applying Game Elements to Cyber eLearning: An Experimental Design* (Tomcho et al. 2019). That article is published in the conference proceedings of *the 14th International Conference on Cyber Warfare and Security* and discusses an experimental design and some unique game elements added to the Cyber Education Hub: Topic Maps and KSA Trees. The KSA Tree information was omitted from the published version of the article due to word count restrictions but is discussed in detail in a supplemental section following that article.

The last article is also unpublished and further discuss the human-subject study originally proposed in *Scholarly Article V*. The article discusses the design of the surveys delivered as part of the overall study and the results of the surveys including the motivation models of the participants and the feedback received about the Cyber

Education Hub website. It also discusses the results of the engagement data collection and also analysis of different subgroups based on insights from survey data.

Scholarly Article: Complex Optimization Algorithm Design Project: Minimal Steiner Tree in Graphs Variant is an unpublished article detailing the design and analysis of several algorithms developed to be used on the Cyber Education Hub's Topic Maps and KSA Trees. This article can be found in the Appendix.

II. Literature Review

1. Introduction

This chapter is not fully in the scholarly article format an primarily serves as a background section. This chapter digs into the importance of cyber and cyber education based off of official military documents and recent works by others within the community. Gamification, the main design principle utilized throughout this thesis is also detailed. After this chapter we discuss a contribution to a paper that details some current issues with Air Force cyber education and training and also presents a framework to help solve these issues; gamification is a part of the framework presented in that paper. Afterwards, two published articles are presented which also briefly discuss our motivation and the principle of gamification that are more thoroughly presented in this chapter.

2. Importance of Cyber

Like many things that provide benefits to society, a vast cyber domain has not come about without its consequences. As global networks have grown and became more connected, the ever-extant struggle between international actors manifested in the new cyber domain. The overflow from these conflicts affects more people, and consequently states, as the domain continues to grow. Thus, over a short period of time, the cyber realm has evolved into a highly contested war-fighting domain (Reith 2016).

The idea of cyber as a war-fighting domain is readily apparent in the mission statement of the United States Air Force, which is “to fly, fight and win in air, space and cyberspace.” In addition to the United States, many other nations have operational military cyber forces, further demonstrating the reality of fighting wars in cyberspace. To expand on the US Air Force mission and the importance of cyber, one can explore the *Department of Defense Cyber Strategy*.

The very first strategic goal from this document is based on the need to educate and train cyber operators in the military. The Air Force has made advancements in recent years, but more progress can certainly be made to better prepare Airmen for the fight. Several challenges of and recommendations for achieving this goal are presented in (Reith 2016). The second goal depends on educating and training anyone using DoD information networks in order to secure sensitive data. Based off of the findings of the United Nations and the focus placed by the United States Department of Defense, one could conclude not only that the cyber domain is important, but education and training about securing and fighting in cyberspace is critical. The importance of having well-gamified motivating educational experiences for United States Airmen cannot be overstated.

Due to the fact that cyber is a war-fighting domain that involves active engagement with adversaries, 'cyber resilience' is a more appropriate term than cyber security. Cyber operators fighting within this domain must realize vulnerabilities, assess risks, and make difficult decisions on what assets should be protected and what territory can be given lower priority. The cyber community must shift away from its culture of compliance and move into a state of readiness and resiliency (Reith 2016). Although cyberspace is a war-fighting domain it should not be assumed that it is the same as, or even closely related to, other more traditional war-fighting domains such as land, sea, air, and space. The terrain of cyber is much different from that of other domains in many ways including the concepts of not having a rigid address or proximity, being able to easily replicate a tool in cyber (file), the ability to change rapidly, and being limited by electronic capabilities rather than physics (Reith et al. 2017). In the same way that cyber is different from other war-fighting domains, the education related to cyber must be different; what has always worked won't necessarily work in this new realm.

Therefore, due to the large involvement of human actors in cyberspace, gamified education and training for cyber operators as well as daily users about sound decision-making and safe use may be one way to begin to solve some of our current issues. In addition, the education itself must be implemented in such a fashion as to inspire people to learn and also practice what they learn. Before digging into the development of a new education platform, it may prove useful to assess the education that is currently offered. The effectiveness of some of the current and past Air Force cyber education tools are evaluated through the lens of gamification in Appendix "Literature Review Supplement: Air Force Cyber Education and Octalysis Level I" §1 and § 2 beginning on page IX-1.

3. Gamification

Merriam-Webster defines gamification as "the process of adding games or gamelike elements to something (such as a task) so as to encourage participation" ("Definition of Gamification" 2018). Gamification is a modern buzzword that essentially means 'human-focused design,' (as opposed to 'function-focused design') according to Yukai-Chou, author of (Chou 2015) and creator of the Octalysis Framework. Through this design practice, products and processes are tailored to maximize human motivation (like one would see in a game) as opposed to maximizing efficiency (like one would see in an assembly line). Chou explains that “gamification is the craft of deriving all the fun and engaging elements found in games and applying them to real-world or productive activities.”

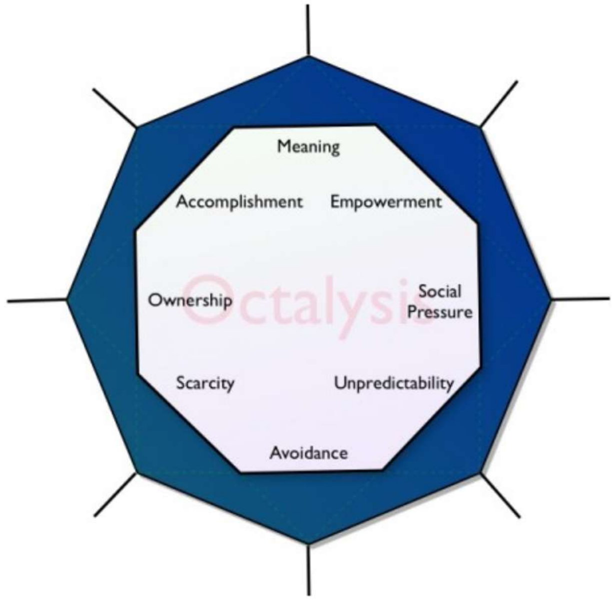
Yukai-Chou also explains that gamification can apply to two different sub-categories: explicit gamification and implicit gamification. Explicit gamification encapsulates what many refer to as ‘serious games’ as well as other common games. A game that is designed to meet some purpose such as education or training, or some other real-life goal is commonly referred to as a ‘serious game.’ Explicit gamification is essentially developing a game where the user knows that they are playing a game. In contrast, implicit gamification is using ‘Human-Focused Design’ methods and applying game elements to create something that users want to do without explicitly calling it a game. A few examples of common implicit games are Wikipedia, Facebook, and Legos (Chou 2015). The authors of (Werbach and Hunter 2012) present the definition that gamification is “the use of game elements and game-design techniques in non-game contexts.” This directly excludes what is described as explicit gamification above. Nonetheless, Yukai-Chou’s Octalysis Framework can be used to analyze either type of gamification.

Since cyber is a complex realm involving human actors and the success of the United States Military depends on effective cyber education, we suggest that gamification of education may be used to help solve our current problems.

Octalysis Framework

This subsection details the Octalysis Framework and digs into each Core Drive. If you are already familiar with this framework and/or prefer to skip this section, note that it is used often throughout this thesis document and context may or may not be presented when this framework is discussed.

Octalysis gets its name because the framework facilitates 'analysis' based off of eight motivators represented by an 'octagon' shape. This Octalysis Framework, which can be seen in Figure 1, encapsulates the eight core drives which motivate people to take action. The author, Youkai-Chou, states that one of the eight core drives are present in everything we do.



Core Drive	Name
CD1	Epic Meaning & Calling
CD2	Development & Accomplishment
CD3	Empowerment of Creativity & Feedback
CD4	Ownership & Possession
CD5	Social Influence & Relatedness
CD6	Scarcity & Impatience
CD7	Unpredictability & Curiosity
CD8	Loss & Avoidance

Figure 1. Octalysis Framework (Chou 2015).

Epic Meaning & Calling: The first core drive Epic Meaning & Calling "is the drive where people are motivated because they believe they are engaged in something bigger than themselves." The author uses several real-life examples where this core drive is successfully employed. The author states that people edit and watch over Wikipedia content not because they get paid or earn a reward, but rather they believe that they are protecting humanity's bank of knowledge, something bigger than themselves. He also uses the example of school rivalries and how they are used to promote a university. More specifically, the author uses examples where schools that have sports teams with significant rivalries often sell more tickets and merchandise while also soliciting more college applications and donations from graduates because they feel like they are part of something larger, that university's community. This core drive is often very powerful when users are just discovering or beginning to experience a product (education is the product in this scenario) and this core drive can also strengthen the other seven core drives when correctly implemented (Chou 2015).

Development & Accomplishment: Core Drive 2: Development & Accomplishment "is the core drive where people are driven by a sense of growth and a need to accomplish a targeted goal." This core drive motivates people to see how far they have come and what they have can achieve. This can be the drive behind people learning new skills or focusing on a career path. The most common implementation of this core drive can be seen in the game attributes of PBLs: points, badges, and leaderboards. The author stresses that this core drive (and all core drives, really) must be carefully designed for when applying Human-Focused Design. If these game elements are simply slapped on a product or experience, people may feel insulted and the desired behavior is never achieved. For example, if one earns points or wins a game by doing something non-challenging or simple, they will likely not repeat the action. One specific example presented by the author is the game of golf. If you could simply carry the ball and drop it in the hole to get a hole-in-one every time, these achievements would be meaningless and no-one would have the desire to play. Another example of a successful implementation of this core drive is LinkedIn's progress bar. By showing users their profile completeness through this simple bar, LinkedIn was able to increase profile completeness by 20 percent (Chou 2015).

Empowerment of Creativity & Feedback: The third core drive Empowerment of Creativity & Feedback is emphasized where people 'play around,' use their imagination, enjoy making their own decisions, and experiment with new designs, strategies and ideas. The author explains that core drive is behind why people play with Legos, demonstrate creativity in Pictionary, and test different strategies in chess. When discussing this core drive in more detail, the author stresses in bold text that "When you design a great gamified system, you want to make sure that there isn't one standard way to win. Instead, provide users with enough meaningful choices that they can utilize drastically different ways to better express their creativity, while still achieving the Win-State" (Chou 2015).

Ownership & Possession: Ownership & Possession is the fourth core drive in the Octalysis Framework. This core drive represents the motivation people have to obtain something and consequently their desires to improve and protect it. This is the core drive behind developing collections and also accumulating wealth. The author explains that even if you did not desire to own or possess something, you are still motivated to care for it once it is yours. This core drive is so powerful that it could even cause someone to care for a Pet Rock, or a virtual Tamagotchi pet. Overall, this core drive can motivate people to do things that may be objectively viewed as irrational, yet those actions may give people a sense of well-being and comfort (Chou 2015). Social Influence & Relatedness: This core drive involves activities motivated by what others think, do, or say. Mentorship, competition, companionship, and group quests are all inspired by this core drive of Social Influence & Relatedness. People's desire to compare and connect with one another is the heart of this core drive. When people are drawn to use a product or participate in an activity that reminds them of something that they know well, or gives them a sense of nostalgia, this is also a part of the Relatedness piece of this core drive. Intrinsic motivation surrounds this drive, meaning that it can give the users deeply rooted satisfaction and motivation to continue to seek products and activities that are strong in this core drive (Chou 2015).

Scarcity & Impatience: Wanting something solely because it is currently unavailable or simply because it is hard to obtain are the basic examples of Core Drive 6: Scarcity & Impatience. The phrase "the grass is always greener on the other side" is a classic demonstration of this. People simply want what they don't have. Also, seeing very few people who are able to accomplish an objective or possess an item will inspire more people to want to follow suit. If something is readily available and easy to access, the value of that object or achievement is low. This core drive is good at causing impulsive actions from users that stem from the desire to gain a scarce good (Chou 2015).

Unpredictability & Curiosity: The seventh core drive in the Octalysis Gamification Framework is the motivating force behind people's obsession with experiences involving uncertainty and chance. This core drive also encompasses people's desire to explore the unknown in the search of surprises. For more specific examples that demonstrate how strong this core drive, one can look at gambling addictions, the skinner box experiments, the lottery, and even Google's "I'm feeling lucky" button. If a little bit of surprise and change can be incorporated into a product or experience, it can be much more productive in terms of attracting people (Chou 2015). Loss & Avoidance: The final core drive in this framework is Loss & Avoidance. This drive motivates people through the fear of losing something or having negative events occur. One ‘real-life’ example of this core drive is the feeling of not wanting to give up on a project and pick a new topic because your hard work up to this point would be lost. Other examples are ‘limited-time-only’ offers, not wanting to fold with a ‘good hand’ in poker even if someone else goes ‘all in,’ and practicing healthy daily habits because one does not want to lose their level fitness. Many studies have shown that people “are much more likely to change behavior to avoid a loss than to make a gain.” This aspect can make this core drive very powerful, but abusing it can cause undesired results (Chou 2015).

White Hat & Black Hat Core Drives: The author gives several examples and explains these core drives in much more detail; the above is not a book report, but a summary of each core drive which will be useful for context during the upcoming analysis section. The author also details some other properties of the Octalysis Framework and the motivation for placing them in certain areas of the framework. The author explains that Core Drives 1, 2, and 3 are White Hat core drives which are positive motivators that influence people by encouraging them to express creativity, give them a sense of empowerment and make them feel like they are part of a larger cause. In contrast, Core Drives 6, 7, and 8 are Black Hat. These are negative techniques that motivate people because they do not know what to expect, cause people to struggle to attain things that seem out of their reach, or act out of fear of losing something. White Hat core drives make people feel good and want to continue the activity, but they are less motivated to act quickly. The Black Hat core drives impulse actions, but users typically do not feel good about these actions and try to eventually wean themselves from anything that overuses these core drives (Chou 2015).

Left-Brain & Right-Brain Core Drives: The Octalysis Framework is also designed to distinguish between 'Left-Brain, and 'Right-Brain' core drives. The author notes that these descriptors are more illustrative than scientific. Core Drives 2, 4, and 6 (on the left side of the model) are in the 'Left-Brain' category. These core drives utilize extrinsic motivation causing people to want to obtain something such as a goal, a skill, a good, or an item that is out of reach. The Right-Brain category includes Core Drives 3, 5, and 7 (on the right side of the model). The force behind these core drives is intrinsic motivation. The interesting thing about intrinsic motivators is that people do not participate in hanging out with friends, being creative, and unpredictable events for some goal or objective, the activity itself is the reward. Right-Brain core drives are often better motivators in the long term, but Left-Brain core drives are usually much easier to implement in a product or experience (Chou 2015).

4. Conclusion

In closing, there are many challenges involved when trying to ensure that the United States Air Force brings the premier fighting force into the cyber domain. Motivational education and training will likely prove critical in the process of developing these Airmen to think critically and act effectively in such a complex and young environment. Gamification can be a critical design consideration when developing new educational platforms that are successful because they are developed with the human user in mind.

III. Scholarly Article Excerpt: Rethinking USAF Cyber Education & Training

Contribution Overview

The author of this thesis was not the primary author of (Reith et al. 2018), the article titled above, and therefore only part of the paper is presented. (Reith et al. 2018) details current problems in Air Force cyber education and training and proposes a framework that can be used to begin addressing these issues. A summary of the problems stated and the framework developed in that paper is presented in the two scholarly articles (IV. and V.) that follow this excerpt. The two extracted sections below include an analysis of the DoD Cyber Awareness Challenge (DISA 2018) and a section on future work proposing applying gamification to the framework presented in (Reith et al. 2018).

1. DoD Cyber Awareness Challenge Analysis

The DoD Cyber Awareness Challenge is a computer-based training module used to provide foundational user training on cyber and information assurance concepts. DoD policy requires all employees to accomplish this training annually. Topics include social engineering, removable media hygiene, protection of sensitive information and information systems, and anti-malware familiarization to name a few (DISA 2018). Clearly the use of points, badges, mini-games and role play suggest that the DoD intended to "gamify" this cyber training, however according to a well-known gamification evaluation system known as the Octalysis Framework (Chou 2015), it fails on several Core Drive dimensions. Consider a Core Drive as a motivating factor that influences users to repeatedly engage, and thus reinforce and expand, learning.

We observe at least three problems with this training module. First, the lack of cyber training options eliminates user choice since the same module, and every topic within, is required to be reviewed annually. Second, the point/badge system is both overly generous and isolated. Described as a challenge, the training gives many points for correct behavior, yet extracts few for mistakes. Even upon attaining a perfect score, the results are lost and the generated certificate merely indicates passing the minimum threshold. This violates the Development & Accomplishment core drive by failing to inspire excellence, and violates the Ownership & Possession and Social Influence & Relatedness core drives by failing to hold the user accountable within a social context. Thus, to improve cyber understanding, the remedy might include progressively tougher grading criteria and resulting score on the certificate and training record. Third, the content is the same every instance with a largely linear gameplay, which violates the core drives of Unpredictability & Curiosity and Empowerment of Creativity & Feedback. Equally as important, the training is presented as a solo activity despite the fact that cyber activity tends to be a highly interdependent team sport. For example, the game fails to associate poor cyber hygiene with increased risk to missions, but instead reinforces absolute rules without any clear concept of likelihood or gravity of consequences. We suspect a better type of training module that addresses these deficiencies is possible under our proposed framework.

2. Future Work

How can gamification be applied to address current shortfalls and challenges presented above? The basis of the concept of gamification is 'human-focused design' which is apparent in a game or theme-park, for example, as opposed to 'function-focused design' which is applied in contexts such as an assembly line (Chou, 2016). Due to the complexity and rapid advancement of the cyber domain and its associated technology it is critical to ensure that those involved are motivated to learn and perform. Also, due to these same factors, it may prove beneficial to create specialized education and training specific to each different USAF community related to cyber as it will allow for better context and more specific relevant technology. There are several prominent gamification frameworks (Burke, 2014; Chou, 2016; Werbach et al, 2012) that should be considered when developing future platforms in order to maximize human motivation and ultimately success through USAF cyber education and training.

Transition to IV. and V.

Thus far we have covered the introduction to the thesis document, some literature review that elaborated on our motivation for the thesis research and the concept of gamification, and an early publication contribution of the author. Next, two published articles are presented, with very minute alterations. Because these articles were published as standalone documents, you may notice some repetition especially when discussing gamification or our specific application, cyber education and training.

IV. Scholarly Article: Engaging Airmen with Cyber Education and Training: Designing a Platform Using Gamification (Tomcho and Reith 2019) was the first published article with the thesis author as the primary author. Some of the problems and the framework detailed in (Reith et al. 2018) are outlined. This paper expands on the framework by beginning to apply basic gamification design to the target platform.

V. Scholarly Article: Applying Game Elements to Cyber eLearning: An Experimental Design (Tomcho et al. 2019) continues to work on the framework discussed in (Reith et al. 2018) and (Tomcho and Reith 2019). Specific game elements are discussed including the Cyber Topic Map and Knowledge, Skill, and Ability Trees (this element is actually discussed in VI. Which is a supplement to V.) An experimental study is also proposed. The experiment and research questions presented in that article are modified. The changes are detailed in *VII. Unpublished Scholarly Article: Analyzing the relationship between Motivation and Engagement: Experimental Study Results and Analysis*.

IV. Scholarly Article: Engaging Airmen with Cyber Education and Training: Designing a Platform Using Gamification

2Lt Landon Tomcho
landon.tomcho@afit.edu

Lt Col Mark Reith
mark.reith@afit.edu

Air Force Institute of Technology
2950 Hobson Way, Wright-Patterson AFB, OH 45433

Abstract

Several issues have impeded the effectiveness of United States Air Force cyber education and training in terms of ensuring that enough Airmen at all different levels of cyber education and training are appropriately prepared. The framework proposed in 'Rethinking USAF Cyber Education and Training' (Reith et al. 2018) is a response to this issue. The framework suggests a platform built around the idea of crowd-sourced content, community engagement, and feedback. This paper proposes several ideas of implementing gamification and human-focused design concepts on the platform and includes an analysis of how this can affect Airmen at different tiers of cyber development. Ideas relating to social involvement, introducing non-cyber-experts to the platform, and a navigable cyber topic map are proposed. These ideas are only a subset of the foundational concepts that can be applied to the platform; data from the platform

should be used to continuously tailor the platform to maximize user engagement and consequently users' cyber knowledge.

Keywords

Cyber education and training, gamification, human-focused design, topic map, 21st century learning

1. Introduction

The rapid development of cyber technology as well as its increasing integration into various US Air Force career fields has led to a demand for better, more accessible cyber training and education for all Airmen. The USAF would benefit from a 21st century approach to education and training where individual Airmen contribute to and consume crowd-sourced content that is up to date and presented at different levels from multiple perspectives. This approach has been proposed as a response to the present Air Force cyber education and training problems of currency, scalability and breadth, and complexity (Reith et al. 2018). The framework in Rethinking USAF Cyber Education & Training (Reith et al. 2018) emphasizes the application of gamification and human-focused design to motivate and engage Airmen with cyber education and training. In order to experience the benefits of the platform entirely, the users must first be attracted to the platform and convinced to stay. Pulling in all types of Airmen to voluntarily learn about cyber is one of the non-trivial challenges we seek to overcome by designing an experience that considers human motivation in each stage of the learner's journey.

2. Current Cyber Education and Training Problems

Many of the current problems relating to Air Force cyber education and training are stated in (Reith et al. 2018). Among these are the currency problem, the scalability and breadth problem, and the complexity problem. The low number of sufficiently cyber-educated personnel in the Air Force is likely a result of several impediments. These include the military's approach to training and problems related to education in general. These observations have led to a crowd-based approach to keep content fresh and sourced from multiple perspectives for users of different skill levels.

The military's general approach to training and educating may not lend to the absorption of cyber knowledge by Airmen involved at all stages of cyber education and training. Note that it can work better in other areas that do not apply to all Airmen and are not as dynamic as cyber, however. The military's sink or swim approach to training means that Airmen either meet a required minimum threshold or fail out. This approach can ensure that everyone has completed some preset benchmark but does not inspire further progress. Specialized cyber training is generally only offered to those in the cyber career field or those in certain leadership positions; all other Airmen are only required to click through the hour-long annual Cyber Awareness Challenge, which has its own abundance of challenges. There are several development tiers of cyber education and training in the USAF based on career field and leadership position and are more specifically described in (Reith et al. 2018). The effect of our proposed platform on each of these groups in discussed in the Analysis section of this paper.

Even outside of the military, the global demand for cybersecurity jobs is skyrocketing, resulting in a dramatic deficiency in the supply of cyber professionals.

Cisco estimates there is more than one million unfilled cybersecurity jobs worldwide (Cisco Advisory Services 2015). Even most college graduates with cybersecurity-related degrees come into the workforce unprepared and ineffective for some time (Endicott-Popovsky and Popovsky 2017). How we can begin to resolve these problems from the cyber education perspective? Before proceeding, we must understand why many students elect not to purse cyber education and why those who do are unprepared. Some of the issues stem from traditional education delivery, a perceived lack of relatedness and relevance of cyber, and the idea that learning about cyber is simply too challenging.

The aforementioned issues are noted in two different papers (Kearney 2016)(Shernoff et al. 2014) as they relate to getting students involved and interested in STEM and in the classroom in general, respectively. When applying these observations to cyber, one may note that in the typical classroom setting, learning generally builds off of prior knowledge from prerequisite courses. Students may either be dissuaded by the prerequisite courses or fear of receiving lower marks for taking harder classes (like cyber) in an environment with a large emphasis on extrinsic motivators (grades). Also note that in a typical classroom, content is passed over only once; whereas learning based on Spaced Repetition decreases the slope of the forgetting curve and leads to longer-term knowledge (Kelley and Whatson 2013).

The education and training for USAF's specialized cyber forces has similar challenges. In addition, the benefits of updating course material and lab infrastructure must be weighed against causing setbacks in an already clogged pipeline (currency problem). Also, only half of recent cyber accessions have STEM degrees (Wingo 2017), resulting in a wide range of background knowledge among students. Teaching to the highest level may yield a handful of well-educated Airmen at the cost of leaving the majority of Airmen frustrated and in the dust. Instead, these courses are generally taught to the lowest level, lending to boredom and cynicism in the Airmen with more background knowledge and skill.

The third issue is that just like STEM, learners may get the impression that cyber is only for the ultra-bright students. The idea that it takes too much time to learn about cyber or that it is simply just too hard for the average person is harmful. Csikszentmihalyi's Flow Theory tells us that optimal performance occurs when the challenge meets the user's skill (Csikszentmihalyi 1990). If people believe that the challenge of learning cyber is too far out of their reach then you can't reasonably expect them to invest their time. How can we introduce Airmen, or civilians, to cyber in a manner that the perceived challenge meets their current skill?

3. Getting Airmen Up to Speed with Cyber

As discussed previously, the worldwide cybersecurity force is severely undermanned. This issue even trickles down and affects the USAF. The commercial sector of cyber has several practices that can be adopted and adapted to help alleviate some of the USAF's challenges (Schmidt et al. 2015). Although these techniques may help, they will not be silver bullets. The limited manning of dedicated USAF cyber forces means that it is every Airman's duty to uphold security standards to diminish cyber threats. The USAF cannot only worry about recruiting and selecting Airmen that will be proficient in cyber career-fields but must also ensure that every Airman is more than just compliant with cybersecurity; they must be educated and inspired to increase cyber fortitude and resiliency (Reith 2016).

One strategy set forth by different sources such as the National Integrated Cyber Education Research Center (Newhouse et al. 2017) and Sobiesk, et al. (Sobiesk et al. 2015) is to place more emphasis on cyber as part of elementary, secondary, and undergraduate core curriculum. This strategy can certainly prove useful (even if there are problems with traditional education) for future cyber professionals, but current professionals cannot be forgotten. Airmen of all ages must be familiarized with cyber. It is harmful to assume that the younger generation understands cyber, and especially cyber conflict, just because they grew up with technology. "A perception exists that using a computer equates to knowing how it works" which is simply not the case (Yannakogeorgos and Geis 2016); and cyber conflict involves more than just understanding technology.

Some other ways to help set the foundation for Airmen could be to get everyone familiar with the cyber domain and terminology via reading and understanding Building and Ontology of Cyber Security (Oltramari et al. 2014) or Cybersecurity: What Everyone Needs to Know (Singer and Friedman 2014). Aside from the currency issue, it is not simple to ensure that everyone in the USAF reads and understands this material without creating mandates and tests, which are troublesome techniques in themselves. As an alternative to these ideas, we propose a well gamified system that builds on the platform proposed in (Reith et al. 2018) in hopes to avoid the pitfalls of the other ideas and current education and training methods.

A. Gamification of Cyber Education

Gamification is a relatively young term that encapsulates the idea of using human-focused design and applying game elements to systems, platforms, and experiences to motivate users and increase engagement. This technique is effective (Hamari, Koivisto, and Sarsa 2014) and has been successfully applied to many successful modern-day platforms. Some examples of well-gamified platforms include Facebook, YouTube, and Netflix. Together, these platforms accounted for over half of all internet traffic in North America in 2016 (Sandvine Incorporated ULC 2016). Well-implemented gamification can certainly motivate people and can even cause ethical dilemmas in certain cases. For example, after backlash from parents, Netflix decided to retract a system that rewarded children with stickers for watching episodes (Desta 2018). However, cyber education is arguably a better goal to push people towards than watching television.

There are several examples of gamification applied to various forms of education including Software Engineering, Information Systems, Math and Science, Programming, etc. Some popular examples are Stack Overflow, Khan Academy, and WebWork. These platforms use game elements to produce outcomes of motivation, engagement, increased interest, and a sense of achievement in the learners (Fui-Hoon Nah et al. 2014). There are also several papers discussing self-determination theory, which relates to human motivation and gamification, and its place in education (Alm 2006)(Vallerand, Pelletier, and Ryan 1991)(Kusurkar, Croiset, and Ten Cate 2011). For a juxtaposition, one can look at the DoD's Cyber Awareness Challenge (evaluated in (Reith et al. 2018)). This education/training module which DoD members must complete annually is an example of a platform that is not well-gamified.

The books (Burke 2014) and (Werbach and Hunter 2012) explain many of the foundational concepts of gamification. Actionable Gamification (Chou 2015) presents the Octalysis framework (Figure 2) which breaks down human motivation into eight Core Drives. The author mainly focuses on Level 1 Octalysis but also presents Level 2 and 3 Octalysis which relates to designing for different stages of the game and different player types, respectively. Level 1 Octalysis can be leveraged to apply common game elements when designing our educational platform. Level 2 Octalysis considers motivating people to join, buy-in to, and continue an experience; the cyber-education journey in this case.

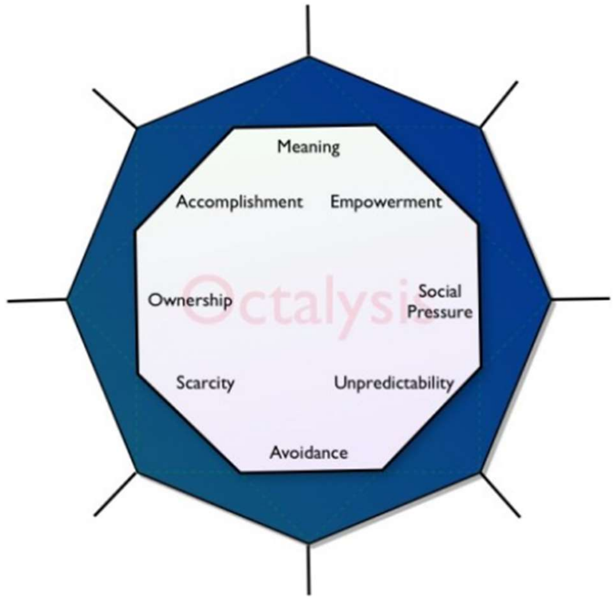


Figure 2. Octalysis Framework (Chou 2015).

B. Applying Basic Game Elements

Crowd-sourced modular educational content not only alleviates the problems of currency, scalability and breadth, and complexity mentioned earlier (Reith et al. 2018), it naturally lends to several Octalysis core drives. The Empowerment of Creativity & Feedback core drive relates directly to users creating their own content and receiving feedback from the community (also Social Influence & Relatedness core drive). This reinforces good content production and allows producers to learn and improve. Most games provide instant feedback to the user. Although a crowd-based feedback system may not be instant, the process of uploading and sharing content should be as smooth as possible with little to no barriers such as a review process to allow for the quickest possible feedback. This may raise a concern over 'false' or unprofessional content. Outside the fact that all content is attributable to someone's actual identity, content that fails to meet community guidelines can be reported as 'inappropriate' or ‘misinformative.’ The reporter can then be required to give a detailed explanation of the problem with the content so that the exact problem is noted.

The community can provide positive feedback on the content via comments or ‘likes.’ Commenting can not only be a place for high praise or compliments on certain parts of the content, but also a place for users to ask questions to clear up confusion or have a discussion among other community members. ‘Likes’ are an easy way for a user to demonstrate that they received the content positively whether it was useful, thoughtful, interesting, engaging, etc. Not having a ‘dislike’ or ‘thumbs down’ option helps prevent early users from becoming dissuaded from posting content that may not be ‘expert’ quality. In the early stages especially, gaining content from as many sources and as many perspectives as possible will be vital to the success of the platform. If the content is ‘bad’ enough to warrant negative feedback, the report options can be used. This design decision will circumvent users disliking content presented in a manner not best for that specific user (i.e. a learner that prefers videos downvotes all blog posts) and also users disliking content that presents new disruptive ideas that are valuable in their own way.

Recommending and sharing videos to specific people takes advantage of the Social Influence & Relatedness core drive. If you share content with a specific person it can remind them to log in to the platform or create a profile and join. Simply being on the platform can spark more content consumption. Allowing users to join groups can establish a sense of community and encourage sharing between units as small as a squadron or as large as an entire career field. Giving a social aspect to the platform ensures that it is more than just a media dump or a distributed learning system.

Personal profile on the platform lends to the Ownership & Possession core drive. Being able to build a profile with a user's background, career field, interests, and past experiences can not only allow other users to gain some perspective behind that user's generated content but also give each user a place that they 'own' to display their achievements and content. Other personalization's such as a custom layout, custom lists, and a tailored content recommendation algorithm also give the user a greater sense of ownership and can encourage them to use the platform more often.

Challenges and levels for earning badges and other rewards directly relates to the Development & Accomplishment core drive. Having clear goals and direction, while still allowing autonomy, can motivate users to achieve while developing their 'skills' (education level) along the way. Weekly challenges that urge users to view/create content could be a simple way to increase engagement with the platform. Displaying progress bars on tasks is another game element that has been shown to increase completion rates (Chou 2015). However, when it comes to points, badges, and other extrinsic motivators, the designer must be careful not to apply so much extrinsic motivation that it overtakes the intrinsic motivation of the user. Using extrinsic elements is great to let users know what they've accomplished and inspire them to achieve more, but too much can have negative effects (Chou 2015).

Some game elements from the Scarcity & Impatience core drive could also be utilized on this platform. Artificial caps are a way of putting limits which most users would not usually surpass on some part of the platform. For example, if most users only view eight content items per day, the platform could advertise a limit of viewing only ten. This internally motivates users to maximize the value of the platform by consuming more. If users wanted to break the limit, they could either unlock unlimited views for the week by posting a content item or permanently bypassing the cap after they become a 'power user' by achieving a certain status. The exact limits and rules need to be carefully considered and tailored based on user data from the platform. Another common game technique is to keep extra features hidden from early users so that they can learn the platform basics without being overwhelmed.

Unpredictability & Curiosity could be targeted with a simple 'random content' button similar to the StumbleUpon website. This can help users find unique and interesting content while sparking interest in new topics and giving users' brains a sensation similar to that of playing a slot machine. The game elements discussed are some features that can fit well into the educational platform proposed by (Reith et al. 2018) and increase user engagement, even in a military context.

C. Designing for Each Phase

The Level 2 Octalysis, discussed in Actionable Gamification, deals with the four phases of a game (or gamified platform). The phases are Discovery, Onboarding, Scaffolding, and Endgame (Figure 3). These phases overlap with Kevin Werbach's theories of Identity, Onboarding, Scaffolding, and Mastery. The Discovery phase deals with the first impression and convincing users to try out a product or platform (Chou 2015). This directly relates to the issue of effectively introducing Airmen and the civilian population to cyber education in a motivating manner. In the military context, it is very important that the educational platform is introduced in such a way that the Airman does not associate it with mandatory training or another clunky military website that doesn't properly motivate users.

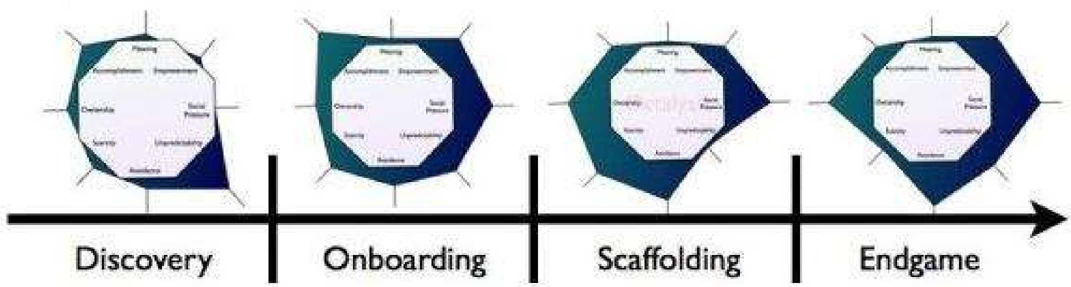


Figure 3. Level 2 Octalysis (Chou 2015).

Discovery and Onboarding Phases

One Core Drive that could be leveraged during the Discovery phase is Epic Meaning & Calling. If Airmen see the platform as a place to contribute to the community and help the USAF, DoD, and the United States, they may be more likely to try the platform. Since Unpredictability & Curiosity is the strongest core drive during this phase (Chou 2015), one may expect humans to constantly be trying new things because of their curiosity, so why do we balk at trying new things sometimes? Nir Eyal, best-selling author of Hooked, claims that “People don’t want something truly new, they want the familiar done differently” (Eyal 2018).

For an example that supports this claim we can look at the California Roll (Eyal 2018). During the 1970s there was hardly any market for sushi in the United States. Nowadays, Americans consume about 2.25 billion dollars of sushi annually (“Sushi Industry Statistics” 2017). So, how did one roll spark the growth of this market? For many Americans in the ‘70s, sushi was too unfamiliar. Much like cyber, the perception of facing too hard of a challenge turned many people away. The California Roll brought the challenge down to a lower level and delivered familiar ingredients like avocado and crab to give consumers a reason to try sushi; the only really strange ingredient for most was seaweed. The later redesigning of the roll to hide the seaweed on the inside was another simple innovation that brought the challenge even lower (Corson 2008).

This innovation of the California Roll fits perfectly with psychologist Mihaly Csikszentmihalyi's flow theory (Csikszentmihalyi 1990) illustrated in Figure 4. It presented Americans that had low exposure to sushi (low skill) to something familiar with a twist (low challenge). After Americans were past the Discovery stage of consuming sushi, they could then try other varieties (harder challenges) and grow their pallet (increase their skill) during the Onboarding phase and eventually become sushi aficionados during the Scaffolding and Endgame phases. There are many other examples of presenting users with the familiar done differently to attract users. For instance, the user interface of personal computers that used common ideas like folders, windows, notepads, trash cans, etc. was more inviting for users than the command line (Eyal 2018). On the other hand, new technologies that do not easily fit into the ecosystem and relate to what users already know and possess often have a hard time taking off (Eyal 2018)(Adner and Kapoor 2016).

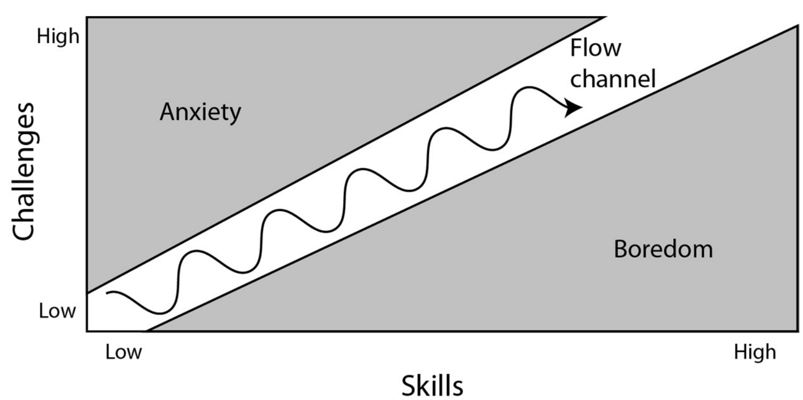


Figure 4. Flow Theory (Csikszentmihalyi 1990).

An example in the domain of education and science is one approach of attracting students to surgical careers. Researchers designed a preclinical surgical experience to introduce medical students to basic surgical skills, familiarize them with the career field, and attract them to the surgical career field by matching the challenge level with their skills (Antiel et al. 2012). As for a cyber example, a four-week summer program was able to increase high-schoolers' interest in pursuing cybersecurity related college majors through hands-on activities relating to cybersecurity (Danforth and Lam 2017). Also, an interactive module detailed in (Peker et al. 2016) that presented the consequences of careless cyber habits to college students was effective in raising cyber awareness, particularly among non-Computer Science students. How can we apply similar ideas to take a diverse population of Airmen that may be unfamiliar with and intimidated by the challenge of cyber and present it in a way that takes them through the Discovery and Onboarding phase to becoming committed to the game of cyber education?

One method would be putting a cyber twist on things that the user is already familiar with. Talk to high school students about the details of how their 'magic' smart phones connect to GPS so that they can navigate to the nearest mall. Get college students, who use Social Networks Sites so much that it negatively effects their GPA (Mcfarlane and Mcfarlane 2017), to realize the impact of cybersecurity in social media and their daily lives. Show an Airman in the aircraft maintenance career field how the aircraft navigation system is equipped to deal with spoofing. Explain how the ID card reader used by Security Forces personnel is connected to the network in order to access the database of valid IDs. Discuss with medical personnel how patient data is encrypted and securely stored on a remote server. After these initial connections are made, it will prove useful for the learner to be able to navigate to other related topics. If the user doesn't know where to go next, they may become frustrated and quit (Chou 2015). A solid tutorial which helps users learn the platform will be crucial, but some sort of map could also prove useful.

Topic Map

A topic map which shows relationships between content can increase the relatedness of content and help ease the educational journey of the learner. But before the learner can navigate, they must have a starting point. By relating topics in cyber to people's everyday lives and careers we can not only clarify that cyber is important to them, but they are more likely be motivated to care and learn about something they can associate with (Chou 2015). Whether the student is in high school or an adult, the student will want to know why they should learn about cyber (Kearney 2016)(NHI 2000). Showing learners that cyber has an effect on their lives can help convince them that cyber education is beneficial.

Accomplishment & Development and Social Influence & Relatedness are the strongest core drives during the Onboarding phase. Having a topic map for cyber similar to Dominic Walliman's maps of biology, chemistry, computer science, etc. (Walliman 2009) would allow people to orient themselves at topics they relate to and navigate to connected topics using the map. As the learner covers more and more topics they can see where they started and how much they have accomplished while they developed themselves along the way. The topic map could also be dynamic and change based on community suggestions. As users realize changes or additions to the cyber domain that are not present in the map, they can suggest edits and be rewarded by this minigame within the platform. The topic map can also show the learner topics which they didn't know that they didn't know. Raising this type of awareness also helps adults want to learn (Leh and Kapp 2018).

Bethesda Softworks games like Fallout 4 and The Elder Scrolls V: Skyrim can give insight into some features that can be adopted for mapping cyber education. In game there is a main storyline to follow, but the user is also free to explore the map on their own or take part in side quests. As the user visits different locations, each with their own challenges and difficulty levels, the user hones their skills, increases their character's abilities, and gains more loot. In the beginning of the game, the map is basic. A single location is highlighted over an immense terrain. The user has direction but can also see the vastness of the world that they can explore at will. As the user explores the map, areas that they come relatively close to will populate on the map. Each location has an associated emblem, which is hollow but becomes solid once that location is explored. Locations can also appear on the map if an in-game character sets you on a quest to that location. If you have visited that location before, you have the option to 'fast travel' or teleport there. Conversely, if you have not been there, your quickest option is to teleport to a location you have visited that is close and then work your way to the quest location, discovering other close locations along the way. Some locations even require you to pay a guide or bring friends along to make sure that you get there safe and complete the objective. In order to minimize clutter on the map (and avoid overwhelming the user), the player can select certain quests from an active list, and flags are placed only at those associated locations.

These maps and quests could be adapted into a cyber education topic map. One main difference is that in these game maps, the terrain defines proximity, whereas nodes and edges may be more appropriate for cyber topics and connections, respectively. Therefore, something like the skill tree in Path of Exile may be more appropriate for representing content on the platform. Unlike navigating terrain, one person can visit multiple cyber topics without visiting all the points in-between. There is also the reality that some topics require more than one pre-requisite to adequately understand.

Nonetheless, applying these game elements can take advantage of human-motivating core drives and increase engagement. The autonomy of exploration fits in with Empowerment of Creativity & Feedback core drive. The user can discover topics that they may not yet have the tools or skills to deal with or find quests that are at the appropriate challenge level and yield helpful rewards. The rewards can include the development of the user's knowledge as well as extrinsic rewards such as job qualifications or progress points that can display on the user's profile. These points could even be used in the Air Force's Talent Marketplace, which is an agile solution to place personnel in appropriate jobs based on experience and skill in different areas (Lamb 2017). These points can demonstrate to unit commanders that the Airman has a certain familiarity level within specific cyber topic areas. This can provide incentive for users to diversify their cyber knowledge and also become experts in certain areas.

Associating quests with locations on the map makes the users feel competent and also gives them achievable goals to work toward (Development & Accomplishment core drive). In the Bethesda games mentioned above, undiscovered locations are not visible on the map, however this should be altered for the cyber map. A three-color system that displays visited topics as solid color, related topics to what has been visited as grey, and topics that have not been 'discovered' as black may be an appropriate hybrid. This way the user can still track where they have been and where they should go next, but they are also not kept in the dark about other topics that exist. These topics may be 'too high of a challenge' for the user's current skill but they could choose a topic that they want to navigate to and can then know where to start and how to get there. Also, in order to manipulate the map as a community (Social core drive), all nodes and edges should be visible.

The topic map and its associated quests and navigation adds significant value and uniqueness to the education platform. This element would be a great distinguisher between this platform and other educational platforms such as Udemy and other crowd-sourced content platforms like YouTube. The topic map gets the community involved to add, change, and remove nodes and edges to alleviate the currency problem, and also empowers users to navigate their own journey while developing themselves and feeling accomplished along the way.

Onboarding (cont.), Scaffolding, and Endgame Phases

To focus on the social part of the Social Influence & Relatedness core drive the user's personal map could be optionally shown on their profile or shared with specific peers. Users can also be motivated to share their progress and take the educational journey with others. Also related to this core drive is the detail that in general, people do what their peers and friends are doing (Chou 2015). A big part of convincing Airmen to voluntarily use a cyber education platform like that presented in (Reith et al. 2018) will deal with what their friends and peers tell them about the platform. The more people on the platform that the user knows, the more likely the user will be to get on board and stay. The phenomena where a product ‘goes viral’ and spreads like a wildfire as more and more people buy in is called getting past the tipping point (Gladwell 2002). Ensuring that users have a good experience with the platform and convincing them to buy-in through the Onboarding phase will be crucial to reaching the tipping point of cyber education for all Airmen.

Another core drive that is strong in the Onboarding and Scaffolding phases is the golden core drive Empowerment of Creativity & Feedback. The idea of including autonomy into education fits perfectly in this core drive. Even with a topic map, some learners may be faced with too many choices and have trouble choosing a path. As mentioned previously, adults learn better when they discover what they don’t know (Leh and Kapp 2018). One idea would be to assess the knowledge of users to determine their baseline in different cyber topics. Quizzes could be procedurally generated from a bank of user generated questions to maintain currency and community ownership. After the learner’s baseline knowledge is determined the system could suggest content that is just challenging enough to motivate the learner to choose to learn about that topic, develop their knowledge, and repeat this cycle without getting bored or frustrated. The repeated engagement with the platform is the key idea of the Scaffolding phase (Chou 2015).

During the Scaffolding phase users will likely find gaps in the content offered or find content that the think could be presented in a better way. The users are encouraged to contribute content as members of the community to add unique perspectives and present ideas in ways that may better reach different types of learners. Through this activity, certain users will likely rise to the top and publish content that is recognized to be valuable to the community. These users will be established as experts within the community and may transition to the Endgame phase of the game (although in the cyber realm learning is never truly complete). Many games struggle to keep users in the Endgame as some users' skill becomes higher than any challenge and they get bored. To alleviate this issue in the new educational platform we can promote these community experts to the role of 'mentor.' Mentors can be assigned to new users to whom they can suggest content to consume, give tips on producing content, and answer questions about the platform and the organization in general.

4. Analysis

The three core problems of current USAF cyber education and training discussed in (Reith et al. 2018) of currency, scalability and breadth, and complexity are addressed by the framework in the same paper and are further developed in this paper. Cyber education and training effects all Airmen, but based on career-field and other factors, the education and training received is different. The broad categories of current USAF education and training discussed in (Reith et al. 2018) demonstrates how the USAF increases investment into smaller and smaller groups of Airmen. Each development tier has challenges which can be alleviated by the gamified platform we have discussed.

All Airmen (and all DoD personnel) are required to complete the Cyber Awareness Challenge annually. While this platform may not replace this module, it is a place where users can learn about how to deal with cyber threats and practice good cyber hygiene daily. If their peers have explained ideas in a manner that relates to them in a style that they can understand (scalability and breadth problem), users may learn more from consuming and creating content than clicking through a rarely-updated (currency problem) 3D quiz once a year. This platform also presents the opportunity to go above and beyond the yearly requirement. Airmen can be motivated to perform deeper research (complexity problem) to develop themselves and also share their unique perspective to contribute to the community. This platform adds value to the USAF by allowing many more Airmen to participate in deeper cyber education and current topics and pushes the boundaries of the current pyramid-shaped cyber education and training model.

Airmen in the cyber career-field experience Initial Skills Training & USAF eLearning as well as Cyber Weapon System Training. These airmen could see several benefits from the platform during these education and training stages and also afterwards. As mentioned previously, there is a wide range of background knowledge for Airmen entering cyber careers. With this platform Airmen can prepare themselves for initial training and education by consuming introductory content (complexity) on the platform as well as skimming the topic map to get an idea of the cyber realm. Course content could not only be hosted on this platform, but as Airmen see that content needs updated or could be presented better, they could upload their own content. Course developers could then pull new content from the platform that is presented in several different styles (scalability and breadth), giving the learners opportunities to learn in their preferred style. Lastly, after initial training, Airmen can share ideas, struggles, and innovations from their units with other cyber squadrons, reducing duplication of effort and increasing force efficiency.

The next group includes Airmen who receive graduate and/or undergraduate cyber education. Again, course content could be hosted on or pulled from this platform, but the research performed at the graduate level along with the projects completed at the undergraduate level could also be posted as content items on the platform. This content may relate to cutting-edge cyber topics (currency) and present views from the educational versus the technical perspective. A great benefit would be that the lessons learned could be shared outside the university bubble to the forces that are actively practicing in the field (scalability and breadth).

The last group discussed is military leadership, who take courses to refine their cyber knowledge as it relates to strategic and operational level decision making. These courses can also take advantage of hosting and pulling content from this platform.

Leadership can stay up-to-date (currency) on new cyber ideas and technology and how the terrain is changing in relation to operations and strategy. With different users presenting content at different levels and depths, commanders can gain insight into how changes in cyber can affect their unit no matter what their cyber background is (complexity).

5. Future Work

This paper leads to several possible future research avenues relating to more efficiently educating people about cyber. The game elements that are applied should be monitored to detail their effectiveness. If something is not effectively motivating users to take desired actions it should be modified or removed. If an element is performing well, analysis should be performed to understand why and how it could be altered to be more successful. More game elements than those presented in this paper could also be introduced to the platform. Empirical research on how different gamification methods work at different states of the game for different player types is of specific interest.

Time should also be invested in deciding how to roll out the platform to different user groups. Should the platform be released as the minimum viable product to all Airmen? Should only the cyber community have access first to find and report bugs with the platform and generate baseline content before rolling out to other career-fields? Would the last option form a stigma that the platform is only for cyber experts? The tipping point as it relates to a military community should also be researched: how many users need to buy in before everyone jumps on the bandwagon?

Another question that should be researched is whether there is a best way to present content to specific people based on their demographics such as age, career-field, education level, personality type, etc. Perhaps machine learning can be used on this data and user's feedback to suggest content that is effective and enjoyable. Along the same lines, work should be put in to developing an effective algorithm for suggesting content to users based on their content consumption. The topic map idea can likely be integrated into this algorithm to suggest content that relates to what the user has seen and avoid content that is too disconnected. The topic map itself also warrants more research. The map could be designed with gamification principles to encourage users to insert/remove topics and edit connections based on the current state of cyber. The design and reward system associated for this map should be further discussed.

6. Conclusion

Although there is work to be done in the future, this paper progresses the framework proposed in (Reith et al. 2018) by introducing gamification and human-focused design techniques to increase human motivation and engagement with the platform. These design ideas may be a good start but the platform should be revised and supplemented based on the response of the users. As the population on the platform increases and as more users enter each stage of the game there will be more opportunities to continue to develop the cyber knowledge of Airmen from different backgrounds at different stages of their careers.

V. Scholarly Article: Applying Game Elements to Cyber eLearning: An Experimental Design

Landon Tomcho[1], Alan Lin[2], David Long[1], Mark Coggins[3], Mark Reith[1]
[1]Graduate School of Engineering and Management, Air Force Institute of Technology, Dayton, USA
[2]Air Force Research Laboratory, Air Force Office of Scientific Research, Tokyo, Japan
[3]Air Education and Training Command Headquarters, Joint Base San Antonio - Randolph, USA
Landon.Tomcho@afit.edu
Alan.Lin@us.af.mil
David.Long@afit.edu
Mark.Coggins@us.af.mil
Mark.Reith@afit.edu

Abstract

As cyber warfare evolves and integrates into military operations, educating and training United States Air Force (USAF) members to be resilient despite a contested digital environment becomes increasingly important. The modern warfighter needs to understand technically-capable adversaries in order to preserve a competitive advantage. Previous analysis identified three core challenges in the USAF involving education opportunity, technical complexity, and content currency. This paper continues the analysis by investigating how gamification applied to cyber eLearning can enhance psychological motivation, increase engagement, and attract non-technical users. One of the objectives of this work is to provide practical options supporting the USAF's continuum of learning education and training strategy. Furthermore, it reports on a cloud-based research platform called the Cyber Education Hub (CEH) that attempts to address the aforementioned challenges by delivering current and relevant crowd-sourced modular educational content to USAF Airmen. The Octalysis gamification framework was selected for analysis and development of the platform. This tool led to the integration of game elements such as maps and skill trees. The map is thoroughly discussed in this paper from theory to design and evaluation. Lastly, this paper outlines a human subjects experiment designed to evaluate motivation and engagement with the platform. The proposed study utilizes a survey to gauge motivation and measures tracking data to evaluate engagement between a control group and experimental group. The authors suggest these findings may be generalizable to other fields of study.

Keywords

cybersecurity education, web-based learning, gamification, topic map, skill trees, human subject experimentation

1. Introduction

Many authors have recognized and elaborated on the current problems in cyber warfare and security. One paramount issue is cyber education and training. Many works have explored the challenges and/or proposed solutions (Singer and Friedman 2014)(Sobiesk et al. 2015)(Endicott-Popovsky and Popovsky 2017)(Cisco Advisory Services 2015). This paper further refines solutions proposed in (Tomcho and Reith 2019) based on the framework created to address problems stated in (Reith et al. 2018). We begin this paper with an overview of currently available platforms which inadequately address cyber education/training problems in the USAF. Discussion on a work-in-progress platform, specific game elements applied to that platform, and an experiment to test these elements follow.

Many educational and training experiences are available for members of the DoD including on the job training, tech schools, formal mentoring, computer-based training, etc. (Reith et al. 2018) discusses the issues with many of these educational/training experiences and sets forth a cloud-based solution to help USAF Airmen who are involved/interested in cyber. There are many other online military platforms available to USAF Airmen, but each have shortfalls in solving current issues in the USAF. Many civilian systems like edX, Udemy, and Khan Academy have had success at engaging their target users in order to foster learning. The overarching problem with these military and civilian platforms is that they have not solved the current issues with USAF education/training (Reith et al. 2018).

2. Cyber Education Hub Platform

The CEH is a crowd-sourced multi-modal cloud-based platform developed to address the current issues in USAF cyber education and training detailed in (Reith et al. 2018). In order to deliver content to a vast array of users at the appropriate level, the platform must foster engagement and content contribution from the community. The platform borrows design principles from successful commercial platforms and aligns them to the military community's goals. These design principles focus on engaging the most vulnerable aspect of cyber warfare, the human. Keynote speakers at the International Conference on Cyber Warfare and Security in 2018 emphasized the human element of cyber. Although technical solutions have their place, human users sometimes circumvent them. By first delivering an engaging platform that humans are motivated to use, effective education and training can be delivered to reduce the risk inherent in the human element of cyber.

Many effective commercial platforms use gamification techniques to increase success. Gamification is about applying game elements to platforms, products, and experiences to increase user motivation and engagement (Burke 2014)(Chou 2015)(Werbach and Hunter 2012). Gamification in education has been surveyed in (Fui-Hoon Nah et al. 2014), and its effectiveness has been shown in (Hamari, Koivisto, and Sarsa 2014) and in many studies on business ROI (Octalysis Group 2018). Not all studies of gamification have shown the same results, however. The study (Kyewski and Krämer 2018) found that presenting badges based on the quality of school work and amount of participation for university students was ineffective. The study was "unable to show that badges help to motivate, foster activity and increase learning results". We believe that their measures of motivation and engagement are effective and fitting, and so these measures are also used in our study. However, we believe that the gamification in (Kyewski and Krämer 2018) was over-simplified and we have acted to avoid this pitfall while designing the platform. Experts emphasize careful design with consideration for the desired outcomes (Burke 2014)(Chou 2015)(Werbach and Hunter 2012); trying a one-size-fits-all solution and carelessly applying points, badges, and leaderboards will not guarantee success (Chou 2015). The human-subject study will provide insight on the benefit of applying gamification to a military platform.

An innovation of the CEH platform is the employment of a Topic Map and Knowledge, Skill, and Ability (KSA) Trees. The Topic Map organizes site content and provides orientation and navigation for the user. KSA Trees present challenges and track user accomplishments. These facets are applications of common game elements of maps and skill trees, respectively. The Topic Map is discussed in detail below and the KSA Tree elaboration is saved for future work.

3. Topic Map

Developers hypothesized that those with cyber backgrounds would be more drawn to the platform while other USAF members may be confused or intimidated by cyber concepts that they do not firmly grasp. At the student level, many are turned away from STEM-type fields because of the notion that it is only for the 'ultra-bright' community (Kearney 2016). We were looking for a way to avoid this notion and present cyber content in an accessible, unintimidating way to increase engagement and motivation among all users, especially those that lack a technical background. Dominic Walliman's maps of Computer Science, Biology (Figure 5), Physics, etc. (Walliman 2009) provided inspiration.

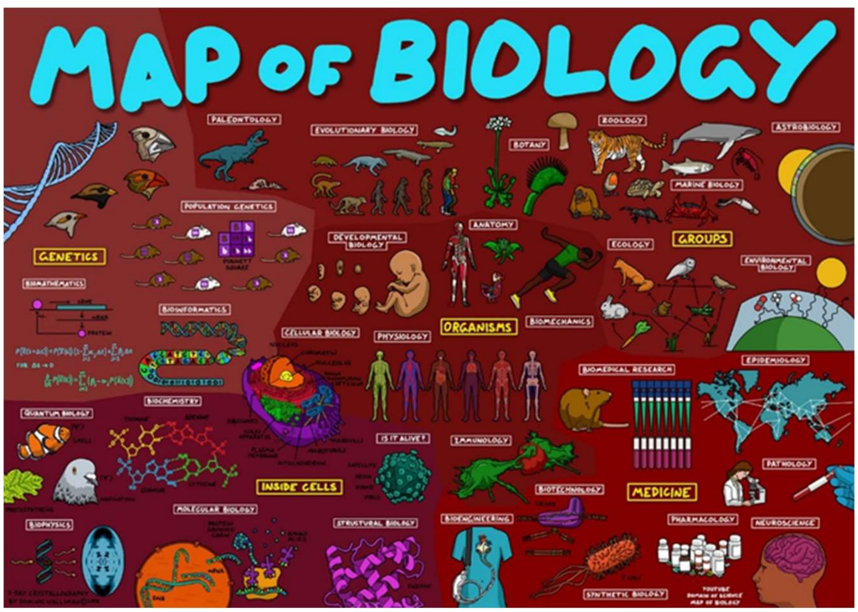


Figure 5. 'Map of Biology' shows the variety of subtopics within the domain of Biology (Walliman 2009).

Platforms like YouTube have minimal visual organization of content, which may be overwhelmingly unorganized for a learner unfamiliar with cyber. Many online educational platforms such as Udemy and Khan Academy have a linear prerequisite content flow where the user is encouraged to consume content in a specific order. The Topic Map lies between YouTube and Udemy and gives the user a sense of orientation and the ability to navigate content independently.

In many games, players access a map for several utilities. Foremost, the user can see the vastness and potential of the world and all of the possible places they can visit. In the early game this could be overwhelming, but the user can zoom in and out to see the world at the appropriate depth. The map makes the current location of the character clear, shows what is close and easily reachable, and what is far away. Another use of the maps in games is to allow users to plan routes for missions. Knowing the current location and seeing all possible routes to the destination gives the player a sense of available options so that they are not overwhelmed, but also gives them autonomy by providing several paths.

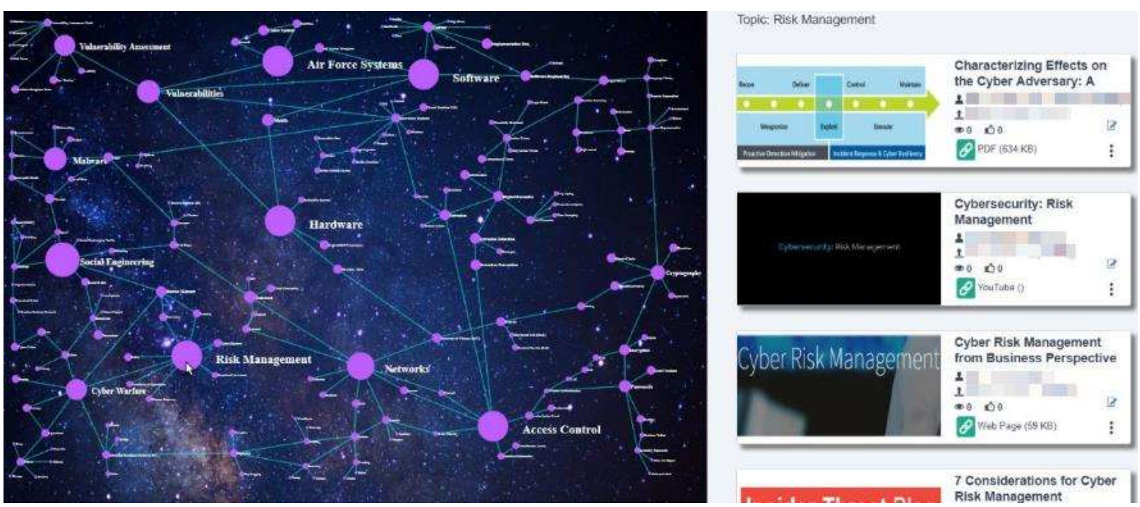


Figure 6. An early version of the Cyber Topic Map implementation ("Cyber Education Hub" 2019).

The Cyber Topic Map is our adaptation of the map game element. The primary functionality is when a user clicks a topic node, the associated content appears (Figure 6). The map allows the user to zoom in (Figure 7) and out to see specific subtopics and large topic areas. One advantage of mapping these topics and showing connections between content is that users can find content they are familiar with and orient themselves to related topics they have context to better understand. The Cyber Topic Map is designed to be dynamic. The community can edit and update the map to maintain currency and relevance of the topics and content better than one person or office could manage.

Although some game maps restrict areas to users until certain regions are unlocked, developers do not want to block any authorized user from content. A user spends a resource, their time, to discover content and explore the map. Anyone should be able to consume specific content without being required to complete a series of prerequisites, and thus the Topic Map lacks a heavy progression dynamic that many games present. To supplement the platform in this area, developers added the KSA Tree element, which poses progressively unlockable challenges to develop a user's KSAs. KSA Trees will be elaborated on in future work, but are referred to in the evaluation and experiment plan sections of this paper.

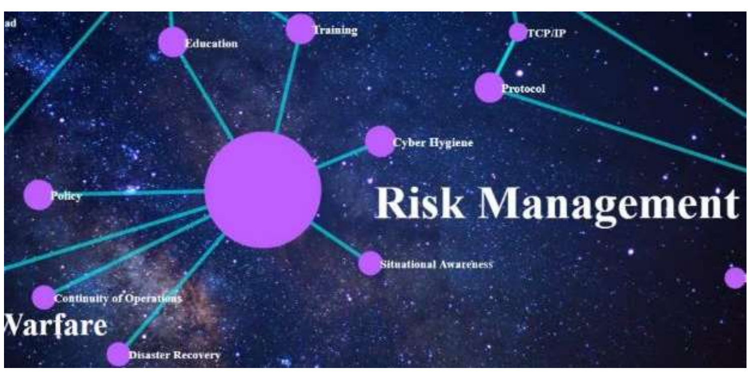


Figure 7. Zooming in on the Risk Management topic ("Cyber Education Hub" 2019).

The map encourages layman users to use the platform instead of fleeing in fear and can be useful for users of all skill levels. When considering the Dunning-Kruger effect (Dunning 2011), the map can also be useful for users at the peak near the left side of the competence axis (Figure 8). Imagine a user who just set up a home network with a custom firewall; this person may feel like a cyber-expert. The map shows the vastness of cyber, allowing the user to realize that they do not know everything. This realization is extremely beneficial because adults learn better when they realize how much, and what specifically, they do not know (Leh and Kapp 2018). This can also produce realistic expectations for people who think all young people who grew up with technology are cyber experts. As Dr. Yannakogeorgos mentions, “a perception exists that using a computer equates to knowing how it works” which is not a reality (Yannakogeorgos and Geis 2016).

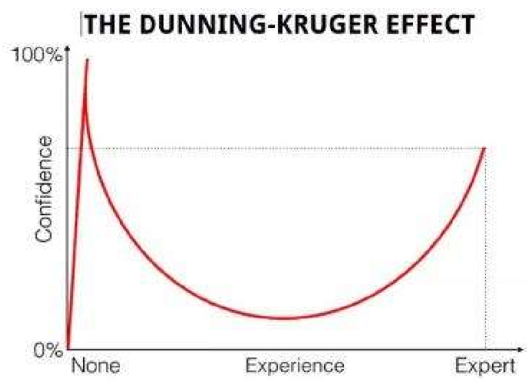


Figure 8. The Dunning-Kruger Effect (Dunning 2011).

A. Building the Cyber Topic Map

Building the initial Cyber Topic Map involved several considerations about the layout, theme, and content. Since the map will be dynamic, focus was on design principles that can be generalized and applied to other applications. Several layouts/themes for the map were considered and a space-themed node/edge layout was selected. The nodes look like planets, stars, and moons, and represent topics while the edges represent relations between topics. This style allows for constant expansion of the map and can show relatedness while not being too restrictive. A user can drag and edit the map to make their favorite topics central while maintaining the connections with other content around the edges. The disadvantages of this type of map are that a user could be completely uninterested in space or could feel there is not enough direction and feel overwhelmed. In future iterations, different themes could be available for user customization of the map and more direction could be given to beginners based on career field, and more advanced users can unlock more features.

The sources for topic nodes on the map included textbooks, articles on current issues in cyber warfare and security, and the Map of Computer Science (Walliman 2009). The primary textbooks used were (Andress and Winterfield 2014) and (Howard, LeBlanc, and Viega 2010). A list of topics was created, grouped, sorted into levels based on depth, given appropriate sizes based on subtopics, and connected. Developers iterated through multiple sorting categories to develop the initial connections.

Based on mind-mapping best practices, the important topics with many connections were placed in the center, and then connections to remaining nodes were added with an attempt to minimize crossing edges and minimize distance between closely related topics. After several iterations of a human researcher organizing the map, the data was inserted into Pajek (Batagelj and Mrvar 2011), a program for large network analysis that offered layouts based on (Kamada and Kawai 1989), (Fruchterman and Reingold 1991), and other force/energy-directed graph drawing algorithms. From a graph theory and algorithms perspective this is an NP-hard problem and took several iterations because of the many nodes (189) and edges (219). The Pajek results were combined with the human drawings to create the first complete version of the foundational Cyber Topic Map. Main adjustments involved spreading out the nodes to increase readability of topics and also further reducing some edge crossing.

B. Analysis of the Cyber Topic Map

The Octalysis framework was developed to analyze the gamification qualities of platforms based on the eight core drives (CD) of human motivation. The following is an analysis of the Cyber Topic Map featured on the CEH. There are several levels of Octalysis (Chou 2015), but we primarily utilize the first two below.

Level 1 Octalysis

This phase of analysis was applied in the early stages of design to evaluate where gaps exist, and how developers could better tailor the Topic Map element to increase motivation and engagement of users. The initial analysis of each Octalysis CD is below, followed by improvements to better attract more users.

- CD1: Epic Meaning & Calling is apparent because the map is a place to maintain currency and relevance of cyber content and topics for the community.
- CD2: Development & Accomplishment is weak. Users may be aware that they are developing themselves but the feedback is missing.
- CD3: Empowerment of Creativity & Feedback is very strong in the map; users have the freedom to traverse the map, add to the map, suggest edits, and explore without the constraints of a classroom or prerequisites.
- CD4: Ownership & Possession is present through manipulation of the map.
- CD5: Social Influence & Relatedness needs supplementing. In the early stages of the platform with a smaller user base it will be hard to feel a sense of community and the social aspects of the map. The relatedness portion may also be minor until more members from different communities are on the platform to help shape the map so that it is more inclusive and has something for every career field.

- CD6: Scarcity & Impatience is also weak in the map. The freedom of traversal basically eliminates making anything exclusive and once content is uploaded it remains on the platform until it is reported or deemed to have expired, so there is little impatience. This is a large difference in our map and the maps of most games, which are progressively unlockable.
- CD7: Unpredictability & Curiosity is naturally occurring due to the nature of not knowing what you will learn and what content exists.
- CD8: Loss & Avoidance of Punishment is weak with FOMO on content being the only contributing element.

Further Gamification of the Cyber Topic Map

The Octalysis Model of the Topic Map was lacking in several areas. Thus, there was potential to enhance the element to yield a more rounded Octalysis Model as seen in Figure 9. Although different people are motivated by different things and everyone would have their own unique Octalysis Model, additional research is needed. Although demographically the US Military is reflective of the American population (Segal and Wechsler 2004), what motivates this specific group may be different. The research study detailed later intends to help bridge this information gap. The following gamification features were applied to the Topic Map to make the Octalysis model rounder, appealing to more types of users.

Delivering feedback to the user based on activity in each topic node is represented via colors of nodes; this provides more feedback than YouTube's history list. To visually represent the decay of knowledge, nodes change in color, from green to yellow to red, for example, when a user neglects to revisit a topic and see new content or refresh on what they previously viewed, encouraging effective spaced repetition learning (Kelley and

Whatson 2013) and also motivating the user through CD8: Avoidance of Loss. A user's personal map showing their progress can be shared with peers for accountability purposes or to challenge one another. Maps can also be aggregated to create 'heat maps' that show where new content is being contributed and where the community is spending time exploring and consuming content. Personalizing the Topic Map view through custom backgrounds and colors increases the presence of the Ownership Core Drive. Scarcity through the 'artificial caps' game element is applied by only allowing users to view X number of content items while interacting with the map during a set time period. This parameter can be tuned and should be set at a number higher than what the average user would exceed. This drives content consumption closer to the artificial cap for the average user (Chou 2015) and can also bring them back more often so that they do not lose out on 'limited' opportunities to view content, which can further encourage spaced repetition learning technique (Kelley and Whatson 2013). Users that exceed the limit should have a way to increase the cap or unlock unlimited views.

- CD2 is supplemented by allowing users to track exploration via node coloring.
- CD4 is emphasized through map customization.
- CD5 is enhanced with the community heat map view and map sharing between peers.
- CD6 is increased by applying artificial caps.
- CD7 is further fortified by a 'Random Topic' button.
- CD8 is built in through the decay mechanism of the nodes, urging users to come back more often.

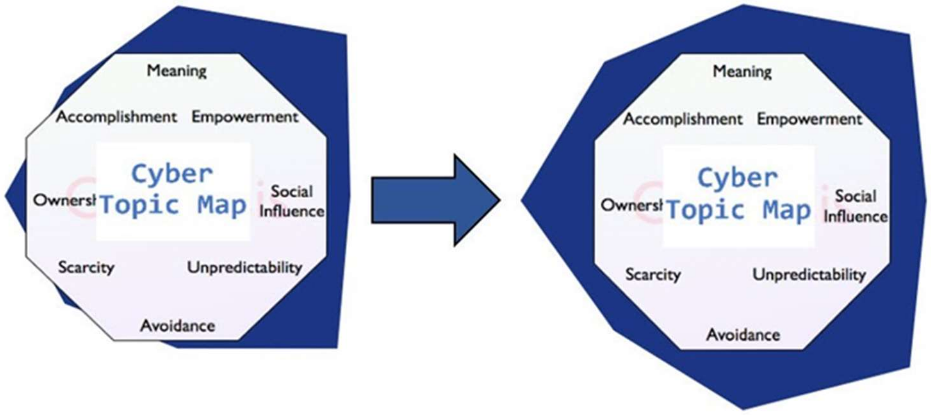


Figure 9. Early Octalysis model of Cyber Topic Map is unbalanced and strong in Intrinsic CDs. After further gamification the model is well-rounded.

Level 2 Octalysis

Level 2 Octalysis incorporates the four stages of the game into the design and analysis of the platform. Ensuring content for users at each stage of the game means that users can be attracted to the platform and stay for a long time without getting bored. The four stages of the 'game' are Discovery, Onboarding, Scaffolding, and Endgame (Chou 2015). A user can Discover the Cyber Topic Map via the top navigation bar of the platform or by hearing from a fellow user. At first, the Topic Map looks overwhelming and it could be hard to locate unfamiliar topics. Therefore, a text-search capability for the nodes of the map will be crucial. Also, a tutorial and suggested starting points can ease users into the experience and decrease frustration. When players discover topics that interest them and find content they enjoy and decide to keep playing, they are in the Onboarding Phase. The scaffolding phase is the main loop where users can find and consume content, comment, share, and contribute. Endgame users who are very familiar with the topics and content in the map will likely be those who keep the map up to date and mentor new users. Cyber changes so often that updating the map and creating fresh content will ensure that the game can never be complete for a user.

4. Knowledge, Skill, and Ability Trees

This section was not in the original publication and is therefore not included here. The supplemental material can be seen in *VI. Scholarly Article V. Supplement: KSA Trees* on page 85.

5. Experiment Plan

Although gamification has been shown to produce results in many applications, gamification must be applied carefully with close attention to desired outcomes. Simply copying from other successful platforms can lead to failure (Chou 2015). Our research seeks to evaluate the efficacy of gamification in online learning platforms and the military environment. While gamification has been considered throughout the design of the platform, the more unique elements, the Cyber Topic Map and KSA Trees, are our main focus for the human-subject research experiment.

The experiment involves recruiting volunteers to use the platform for three weeks. The platform is intended to be used voluntarily, so a typical lab setting would not yield realistic or desired results. The participants will be placed into one of two groups. The control group will have access only to a standard interface similar to YouTube or Netflix where the user can search content, see recommendations, and trending content in a series of columns and rows of different categories. The experimental group will also have access to the Cyber Topic Map and the Mobile Technology KSA Tree. The website requires a US DoD CAC to access, which limits the participants to US military members, civilians, and contractors.

There are four primary investigative research questions we seek to answer with this experiment:

- "How does a Topic Map and KSA Tree affect participants' engagement with online military education platforms?"
- "How does a Topic Map and KSA Tree affect participants' motivation to use the platform?"
- "How does a Topic Map and KSA Tree affect participants' motivation to pursue more cyber education?"

- "What does the Octalysis model of the participants look like? Are there significant variances between career fields, age groups, etc.?"

For the first investigative question the hypothesis is that experimental group participants will log-in more often, spend more time on the platform, and consume more content than the control group. Engagement will be measured by tracking when a participant logs in, the time spent on site, and the amount and type of consumed content. The number of log-ins along with the time data will allow the researchers to understand when the participants are using the platform, for how long, and if there is a significant difference between the control and experimental group.

The hypothesis for the second question is that experimental group participants will enjoy their experience more and be more motivated to continue using the platform than participants in the control group. Enjoyment and motivation are psychological phenomena that can be evaluated with a participant survey. The survey will allow the researchers to gauge the participant's feeling towards the Topic Map and KSA Tree and the platform in general.

The hypothesis for the third research question is that participants experimental group will be more motivated to continue pursuing outside cyber education more than participants in the control group. Again, the psychological phenomena can be measured with a survey. The survey will allow the researchers to record the participant's feeling towards cyber education and whether the participant is more motivated to pursue further education in the future as a result of the treatment (accessibility of Topic Map and KSA Tree). We will also ask the participants about their use of external information to assess if the treatment makes a difference in this area.

A survey can also be used to assess the final research question. By asking the participants about different core drives of motivation and how much each influences their choice of activities and what brings them joy, we can build an Octalysis models for each user or find common motivators for different career fields. For example, aircraft maintainers may be motivated by Ownership & Possession, pilots may be motivated by Achievement & Development, those who work with nuclear weapons may be motivated by Epic Meaning & Calling. This information can be used to tailor the CEH and other platforms to increase overall engagement, motivation, and enjoyment.

Although ensuring that actual learning takes place on the platform is crucial, we know that learning cannot occur without attendance. This research study primarily focuses on getting Airmen motivated to use and remain engaged with the platform. For now, we are assuming that exposure will yield learning. We also hypothesize that more learning will occur for users with access to the Cyber Topic Map and carefully developed KSA Trees. These elements allow the user to see the vastness of the cyber domain and can provide paths for users to learn foundational concepts and gradually become experts and critical thinkers. We encourage future work to focus on tweaking these elements to improve learning while maintaining or improving the current gamification design principles.

6. Generalization

Although tailored to address a specific problem within a specific community, we believe that the benefits of the platform, game elements, and experiment are generalizable to a wide range of applications. Other schools/units have reached out wishing to take the platform base and create their own instance of a Civil Engineering Hub, for example. Although improving cyber education and training is our specific end goal, the platform features of crowd-sourcing and utilizing the cloud as well as the Topic Map and KSA Trees can address issues in other educational fields, specific jobs, career fields, and more.

7. Limitations

The platform's first impression on Airmen will largely determine its success in the military environment. If users understand that it is part of research and is built to solve their problems, they should be able to accept unfamiliar features and frequent changes that come with development platforms. In contrast, if it is seen as a military training requirement it may turn users away. Furthermore, we do not know if the benefits of gamification and autonomy can lead to negative effects from the military perspective: can this lead to less respect for leadership directives, more distraction from operations, a lack of discipline, rejection of training requirements, etc?

Without enough users to get past the tipping point (Gladwell 2002) the platform may struggle as the crowd-sourcing of content is paramount to address the problems stated in (Reith et al. 2018). There are also limitations with eLearning including not having physical access to instructors. While the platform escapes the limitation of prescriptive education that tells students “let me guess what you need to know” by facilitating choice, formal education does have the benefit of standardization of learning objectives and certifications. A limitation of the experiment is that learning is not specifically tested. Although this is partially due to the ambiguity of the definition of ‘learning,’ we also realize the limitation that teaching concepts to a student does not always translate to changes in behavior, which is crucial in cybersecurity.

8. Conclusion

In summary, the authors believe that the CEH can avoid the pitfalls and shortcomings of other USAF cyber education/training platforms if it makes a good first-impression in users and inspires a grass-roots effort to make the platform truly user-owned and crowd-sourced. We hypothesize that the careful gamification design and especially the Cyber Topic Map and KSA Trees will lead to an increase in motivation and engagement with the platform. The human-subject research study focused on cyber education/training will ideally provide insight to the efficacy of the design and yield future recommendations for other education/training platforms and experiences.

VI. Scholarly Article V. Supplement: KSA Trees

Overview

The following material can be placed between §3: Topic Map and §5: Experiment Plan of *V. Scholarly Article: Applying Game Elements to Cyber eLearning: An Experimental Design*. This material roughly follows the same outline as §3:Topic Map and was left out of the aforementioned publication due to word count restrictions.

A. Skill Tree Introduction and Examples

The Elder Scrolls V: Skyrim presents the user with 18 different skills and associated skill trees to unlock new character abilities and skills. Characters may focus on their strengths and complete five or six skill trees or strive for breadth, touching the lower levels of many trees. One difference in the Skyrim skill trees and the CEH KSA Trees is that the former aids the character in meeting the game's increasingly difficult challenges (aligning with the principle of flow theory (Csikszentmihalyi 1990)), while the latter presents the user with challenges and tasks that result in real-life KSAs. Both are powerful minigames that increase user motivation and engagement with the game/platform.

Figure 10. Skyrim Skill Trees ("The Elder Scrolls | Skyrim" 2011).

A large number of games feature the skill tree game element. Role-Paying Games (RPGs) like Borderlands 2, Skyrim, Path of Exile, and Final Fantasy X have skill/talent trees for character building and development. Real-Time Strategy (RTS) games like Civilization V, XCOM: Enemy Unknown, and StarCraft 2 feature technology/resource trees for investment into one's society, military, etc. Each game has a unique tree layout and design. Some games have several separate trees while some have a 'forest' of interconnected trees. Below are several examples of these trees.

Figure 11. Borderlands 2 Skill Tree for Gaige (“Borderlands 2” 2012).

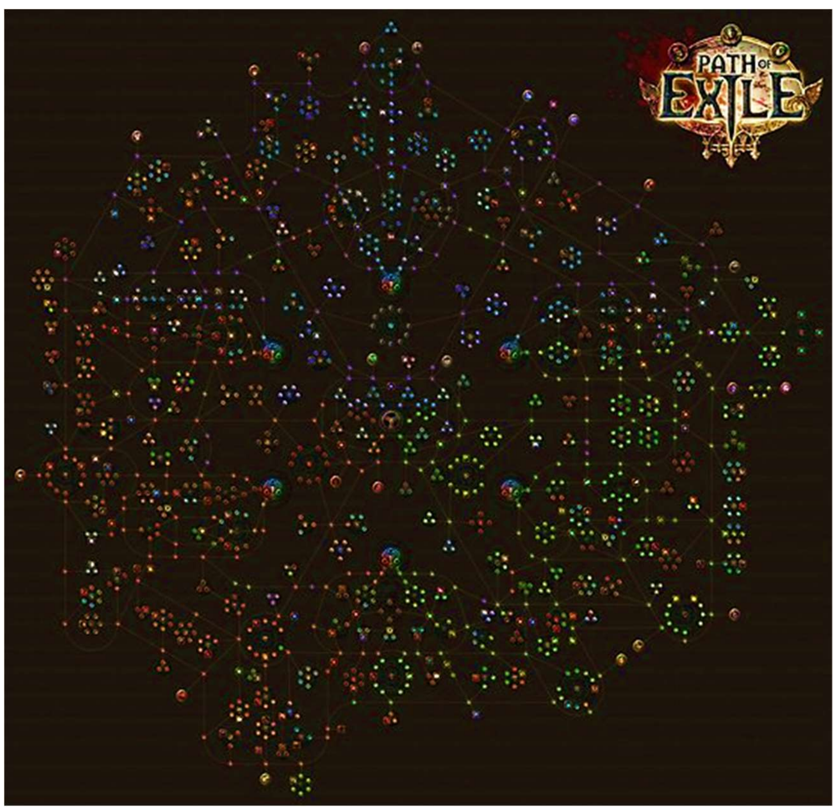


Figure 12. Path of Exile Skill Tree (forest style) (“Path of Exile” 2013).

Figure 13. Civilization V Tech Tree (partial) ("Civilization V" 2010).

Trees have also been previously applied to education and training in the US Air Force. These trees are primarily to show prerequisites and dependencies and allow for some choice. One example is the Course Flow for Computer Engineering (CompE) Majors at the United States Air Force Academy. There are required courses for all students, requirements for CompE majors, and color-coded elective courses that allow students to decide on a focus area or elect breadth. Arrows demonstrate pre-requisites and relationships between courses. These elements have potential to be further developed and used in more applications to increase engagement and motivation.

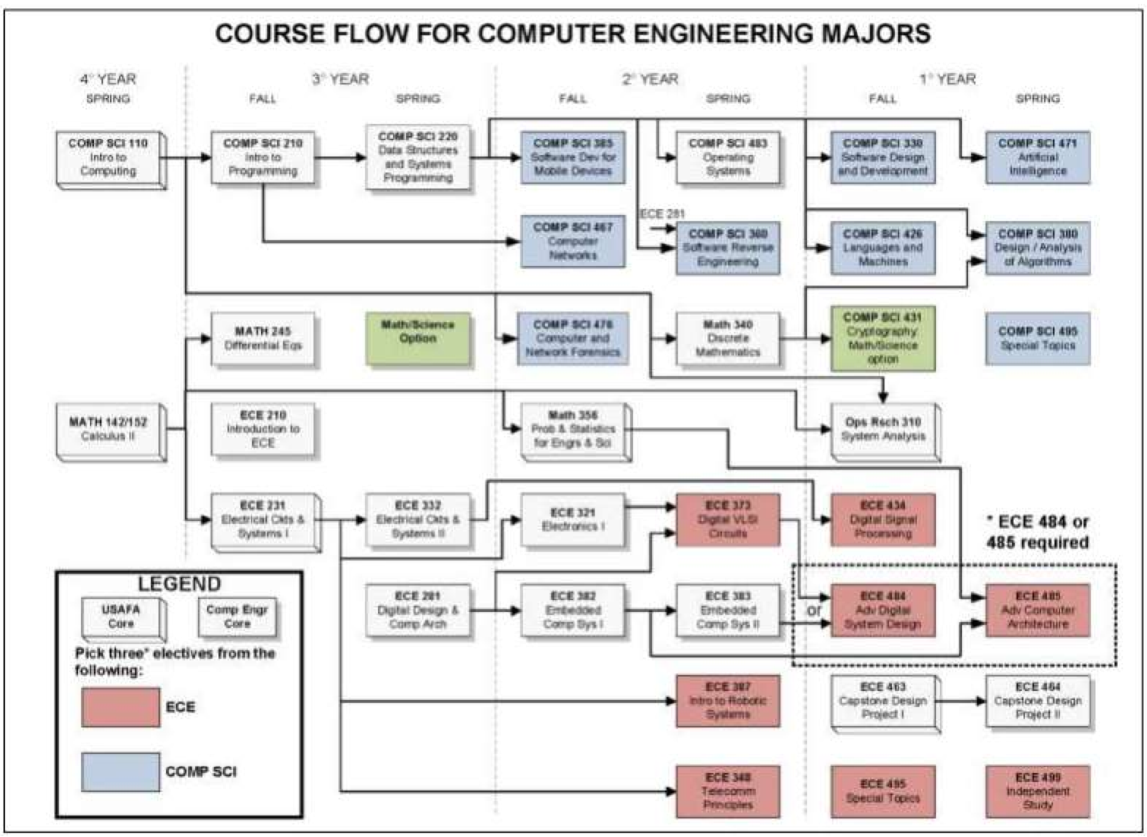


Figure 14. USAFA CompE course flow ("Computer Engineering Major" 2018).

B. Framework

In order to effectively adapt the game element of skill/talent/tech/resource trees and apply them in various educational and training courses, specific jobs, career paths, and more, we propose the following baseline framework for creating and using KSA Trees.

i. **Entry points**. Each KSAT needs at least one entry node where the user can begin their experience. For more autonomy, more entry points can be incorporated, but too many options and not enough direction can lead to overwhelming the user. There may be a few applications where it makes sense to make every KSA node an entry point, but these situations should be rare.

ii. **Progression**. One of the paramount features of the trees is the progression dynamic. This means that there are some unlockable nodes where something must be accomplished before the user can access these nodes. The entry points are the only unlocked nodes in the beginning and remaining nodes must be unlocked before they can be accessed and completed. This is one way to force prerequisites.

iii. **Connections**. Connections can represent prerequisites or related topics. With only a few connections there is little choice for the user and little reason for them to interact with the tree.

iv. **Choice**. Although there may be requirements before unlocking nodes, these unlocks should not always be limited to a single path. Highly desired nodes especially should be reachable from multiple avenues. Although some games like StarCraft and Civilization may not provide these options and can even require multiple prerequisites, games like Path of Exile provide the ability to get to any node from anywhere else. This allows the user to decide their own route and create their own unique character.

v. **Builds**. Not all trees in games themselves allow for different specializations within, but in those games, there are generally multiple trees that the user can choose to utilize and create a unique character that specializes in an area like healing, magic or damage. Specialized characters fit into what are commonly known as builds. In many RPGs there are different strategies to beat the game. Certain challenges will be easier for different builds and harder for others. Trees that facilitate several focus areas allows for different users to become proficient in different areas, just like in real life. Trees that facilitate several builds can be effective for team building. In an RPG, to defeat a boss you may want a Tank character to distract the boss and take damage, a Healer to keep the Tank alive, and a Mage to attack the boss from a distance. In a cyber unit you may want a Network Manager to manage systems and maintain connection, a Penetration Tester who finds holes in new systems, and a Defensive Specialist to repair vulnerabilities and be proactive with security.

vi. **Difficulty**. To ensure an effective progression dynamic for users at all stages of the game, flow theory (Csikszentmihalyi 1990) should be incorporated to the tree via different difficulty levels. Some nodes should be easy for the user (especially entry points). Other's should be slightly more challenging and a few should be only for experts. This ensures that users of all skills can take on challenges that are the appropriate level to facilitate flow instead of yielding frustration or boredom.

vii. **Depth**. Nodes can also have different depths within. For example, when within a KSA Tree for an introduction to programming course you may want users to get an overview of several types of languages including Python, C, C++, Java, and MATLAB. Thus, it would be a requirement to fulfill the base level challenge or task within each of these nodes, but there may be additional optional challenges of various difficulty within the nodes to allow a user to gain deeper KSAs in certain areas. Now instead of only writing a 'Hello World' program in each language, the user can get a taste of data structures, object-oriented programming, and threading specifically in C++. These depths could theoretically be included as different nodes altogether that are connected to the base C++ node, but it depends on the designer's decisions and the tree layout.

viii. **Achievement.** As users progress through the tree they should absorb knowledge, develop skills, and ultimately gain abilities. When appropriate, the lowest level nodes should be focused on knowledge, and skills should follow. Completing leaf nodes should be a representation that an ability has been gained and this achievement can be recognized with awards or other certificates so that the user is reassured that6 their time investment has paid off.

C. Building the Mobile Technology KSA Tree

The first KSA Tree was built to be used in the human subject research experiment described later in this paper, and also as an example for future KSA Trees. The target demographic for this experiment is USAF members with career fields or jobs related to cyber, although not everyone has to have a strict cyber background; some participants have intel backgrounds, while others have comm or flying backgrounds. Therefore, an ideal tree is related to cyber and has starting points for a wide range of participants. Mobile Technology was selected as the theme for the tree. Mobile Technology is current, relevant, and explores a wide variety of topics ranging from Cellular Networks to GPS Tracking to Bluetooth. The Cellular and Mobile Technology Knowledge Unit (from the NSA's mandatory program content for Cyber Operations CAEs), which was mapped to the NICE Framework Competencies, provided the foundational KSAs for the tree.

The tree was based on the forest or sphere grid layout used in games like Final Fantasy X (FFX) and Path of Exile (PoE). There are fewer connections in our KSA Tree than in the map, allowing for better organization and less overwhelming path choices. This tree type allows for specialization and depth but also balance in breadth. Different player types generally start in different areas of the tree and traverse in order to achieve their desired character build. The starting nodes for the Mobile Technology tree were selected to be Cellular Phones, Wi-Fi, Passwords, eMail, GPS, and Web Pages. One current weakness of the tree is that the layout is not as refined as those in PoE and FFX and therefore it is not as intuitive what to focus on for those who want to min-max character builds. The current version of this KSA Tree can be seen below.

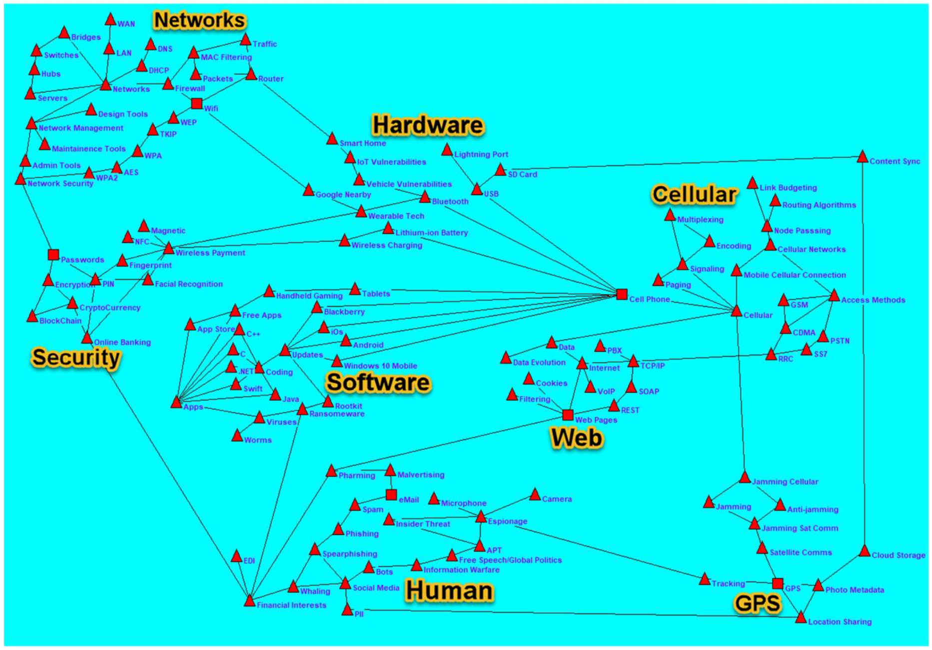


Figure 15. Mobile Technology KSA Tree

Another feature of the KSA Tree is the depth of each node. As mentioned previously, the nodes have associated challenges and tasks. At least one must be completed before the user can unlock connected nodes. In addition, each node has several levels of challenges, including Easy, Medium, Hard, Expert, and Master difficulties. These difficulties are loosely tied to Bloom's Taxonomy of Learning. Some Easy challenges ask the user to view a content item or play a game. Medium challenges have the user comment on other user's content to start a conversation. Hard challenges can vary from evaluating several content items within a topic or achieving a high level in a game to adding outside content sources into the CEH. Expert and Master challenges have the user contribute unique content (videos, documents, presentations, etc.) on a topic with the latest knowledge based on their perspective.

D. Further Gamification of the KSA Trees

The following gamification features may be applied to the CEH KSA Trees to better motivate and engage more types of users: providing useful feedback to users via node colors/icons that correspond to the locked/unlocked status of a node as well as the depth achieved within the node; more complete nodes are green or blue and minimally complete nodes are orange, for example. KSAs can also decay, causing a change in color. A ‘heat map’ view can show unit commanders aggregate trees where node color represents the number of users (and the average depth) at which each node has been completed to demonstrate the strong/weak KSAs in their unit. Impatience is emphasized by only allowing certain KSAs to be unlocked during certain times, encouraging users to complete them during those time windows. The ‘appointment dynamics’ element is different than a deadline because the window can open back up, contacting and alerting the user when it is available.

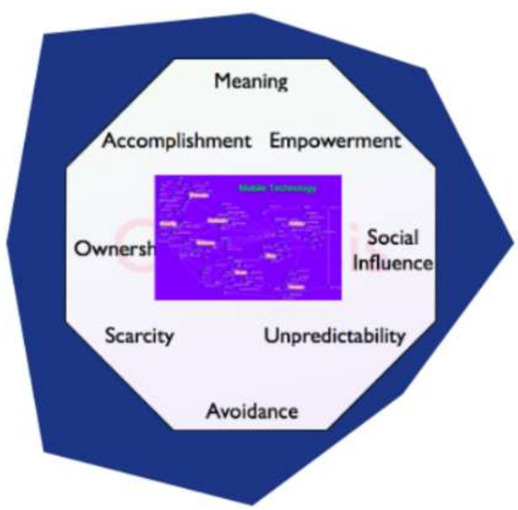


Figure 16. KSA Tree Octalysis

E. Level 1 Octalysis – KSA Tree

The KSA Tree Octalysis model, Figure 16, is also fairly balanced. However, it is more extrinsic focused (left three drives) whereas the Topic Map is more intrinsic focused (right three drives).

- CD1 is at a similar level as the map. The user has a more focused mission than they do with the map, and the tree is more about CD2 than it is about being a place to maintain currency and relevance of cyber topics, like the map.
- CD2 can be more easily tracked in the tree as there are specific challenges and depth to each node.
- CD3 is less powerful in the KSA Tree; there are still unique traversals and focus areas, but there are also more limits and more direction in the tree.
- CD4 is very similar to that of the map when it comes to customization if the user creates the tree themselves. With trees created by others, users will still own their personal knowledge, skill, and abilities demonstrated on the tree.
- CD5 will be less significant in the early days until the user base is substantial. Afterwards users can share their progress and accomplishments and with more friends, peer groups, and the larger community. Similar to the map, group trees and a heat map can also supplement this core drive.
- CD6 will be apparent due to some naturally occurring and intentionally forged KSA Tree nodes that will be rare due to difficulty or high specialization. Appointment dynamics will also fortify this core drive.
- CD7 is slightly limited in the tree game element. Although users cannot see the specific challenges in locked nodes (depending on the specific tree designer's decisions) they will still be able to see the name of the different nodes. This allows the users to have more information when choosing the area of the tree they want to focus on.
- CD8 is present due to KSAs decaying and expiring over time, similar to the map's feature.

VII. Unpublished Scholarly Article: Analyzing the relationship between Motivation and Engagement: Experimental Study Results and Analysis

2Lt Landon Tomcho
landon.tomcho@afit.edu

Air Force Institute of Technology
2950 Hobson Way, Wright-Patterson AFB, OH 45433

Abstract

Many modern education and training platforms focus on how to best teach concepts or how to present content to learners in more efficient ways. Before these challenges can be addressed, a platform must have a user-base; learners must be present before they can learn. Gamification is the practice of utilizing human-focused design to increase sustained user interaction with various platforms, such as a website. The surveys and experimental results discussed in this article utilize a common gamification framework to assess user data. This includes creating motivation models to compare against user activity on a fledgling, military, crowd-sourced education platform in an attempt to find gaps between what users desire and what the platform presents. Contributions of this paper include presenting a quantitative way to perform this data collection and analysis, discussions about software engineering practices that led to website and tracking database design decisions, a framework used for comparing user motivation and engagement data together to draw conclusions, and specific recommendations to designers that aim to increase motivation and engagement of target users with this platform.

1. Introduction

This paper discusses the human-subject study first presented in (Tomcho et al. 2019). The study consists of three portions: a base-survey to collect demographic and motivation data; an experiment where users interact with a website and tracking data is collected; and a post-experiment survey where the user provides feedback about the website. Participants are military members (from operational units and a classroom setting) and the focus website for the study is the Cyber Education Hub (CEH). This website was built off of the framework presented in (Reith et al. 2018) and further developed in (Tomcho and Reith 2019) and (Tomcho et al. 2019). The flow of the study can be seen in Figure 17.

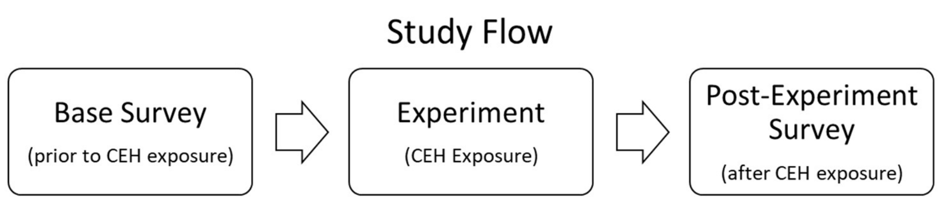


Figure 17 – Study Flow diagram.

The goal of this study is to test and evaluate the implementation of gamification in the military environment with the CEH as the primary vehicle for this investigation. The initial plan for the experimental portion of the study was to test differences in interaction with certain game elements between a control and experimental group. Due to logistics and further refinement of our research goals, the experiment portion of the study changed. Specific research questions are detailed below. The new plan was to let all users interact with all available website features and perform some tracking data comparisons between subgroups based on demographic and motivation data. Motivation and engagement are the core of gamification, and so the Octalysis Framework (Chou 2015) is a prominent tool used in the design and analysis of these surveys.

In this paper we present our research objectives then detail our methodology concerning: developing the two surveys, tracking engagement with the CEH website, and analyzing our resulting data. The results and analysis are presented afterward. Finally, conclusions about research questions are presented and future work is proposed. Major contributions of the article include the following:

- We show that statistically significant engagement differences can be attributed to differences in Motivation Levels using a novel framework.
- Recommendations are given to military leadership and other military platforms based on what core drives of motivation are the strongest and most commonly apparent in military members
- Research-based software design recommendations are given to CEH developers.

2. Methodology

A. Research Objectives

The four research questions stated in our previous work (Tomcho et al. 2019) were replaced with the four research questions below. Before our current research questions were fully formed, the previous four were split into several sub-questions which can be seen along with their answers in the Appendix on page IX-15. Answers to these questions can be seen on page 179.

- (How) can differences in engagement with a platform be attributed to differences in motivation? Which subgroups showed the greatest engagement with the CEH and why?
- By implementing modern design techniques such as gamification, do target users engage more with the CEH than they do with other platforms?
- (How) do unique game elements such as a Topic Map and KSA Trees have utility in the military environment?
- What differences exist between motivators that cause military members to act and motivators that military members enjoy in games/activities? Which should be prioritized when designing military platforms?

B. Survey Delivery

Two surveys, a base survey (Appendix A, page IX-52) and a post-experiment survey (Appendix B, page IX-58) were given to participants of two main groups: a classroom group, and an operational group. The survey was split to afford participants to take two (or at least one) 10-minute surveys rather than a single 20-minute survey. The questions that did not depend on use of the CEH platform were sectioned into the base survey, and questions that did depend on the use of the platform were placed into the post-experiment survey. Each part of the study (each survey and the experiment) was voluntary, and the majority could provide useful data without relying on the participant completing other parts of the study. The post-experiment survey required that the participant saw the CEH platform at least once, but they did not have to be an active user. The more parts of the study a participant volunteered for, the more complete and powerful the resulting data would be.

The 14 classroom participants were students in AFIT's (Air Force Institute of Technology) Introduction to Cyber Warfare graduate-level class. Volunteers in this study were exposed to the Cyber Education Hub (CEH) platform during the 10-week program and were able to utilize this website at their discretion. Multiple operational units participated in the study including 26 total members from the 88th Communications Squadron (88 CS), 33rd Network Warfare Squadron (33 NWS), and the 426th Network Warfare Squadron (426 NWS): the Reserve Associate Unit to the 33 NWS. The latter two are referred to as 33 NWS in this paper. Operational group participants were given the opportunity to access the site for at least 3 weeks before the post-experiment survey was distributed. 88 CS members were the only operational users with exposure to the site that met this threshold.

C. Base survey

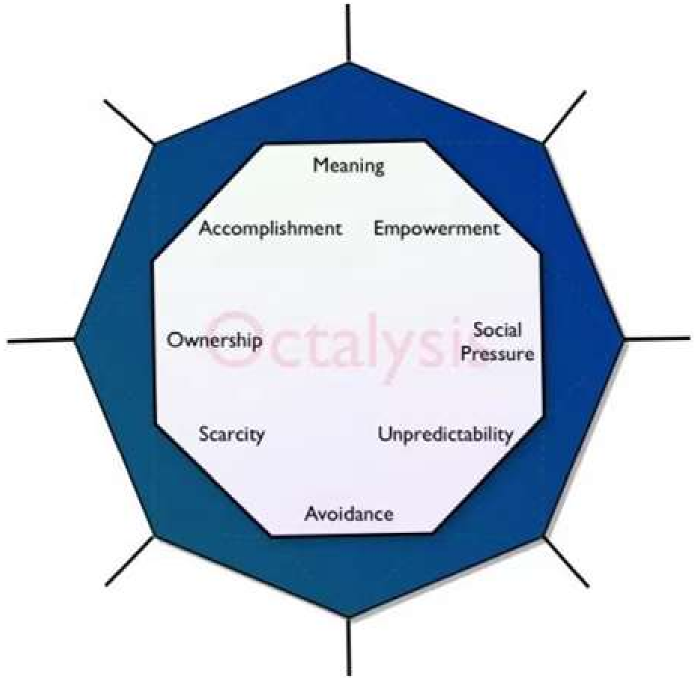
The base survey collected data about demographics, use of other Air Force platforms, and the Octalysis profile of the participant. These elements are broken-down and explained in more detail below. Only results pertaining to our primary research questions are presented in the main portion of this document, additional results can be found in the Appendix, beginning on page IX-67.

Career Information and Demographics

Demographic questions were used to understand how our results may generalize to larger populations and for readers to gain insight into the background of the participants. Some demographic questions were military specific, including determining if a participant was a civilian, officer, or enlisted. These questions are also leveraged to split participants into sub-groups for comparisons later in this paper.

Gauging the Octalysis Profile of the Participant

The Octalysis model is a gamification framework developed by Yukai-Chou that breaks human motivation and enjoyment down into eight CDs (shown in Figure 18). This model is summarized in our other works (Tomcho et al. 2019)(Tomcho and Reith 2019) and explained in detail in (Chou 2015). Based on a platform's Octalysis model, you can gauge how a user may feel when interacting with the platform, and how it will affect their motivation and engagement. The survey questions help us determine what motivates the participants to act and what they say they enjoy, and allows us to create an Octalysis model for each. Based on the findings, we can apply the Octalysis of the CEH website and compare it to what target users expect/desire. Any discovered gaps can then be addressed by the designers before releasing the platform to a larger audience.

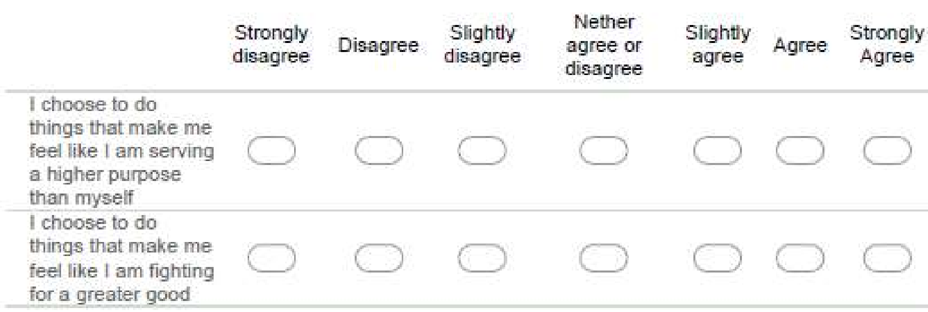


Core Drive	Name
CD1	Epic Meaning & Calling
CD2	Development & Accomplishment
CD3	Empowerment of Creativity & Feedback
CD4	Ownership & Possession
CD5	Social Influence & Relatedness
CD6	Scarcity & Impatience
CD7	Unpredictability & Curiosity
CD8	Loss & Avoidance

Figure 18 – Octalysis Framework and names of each Core Drive (Chou 2015).

The Octalysis Framework (Chou 2015) was used to develop survey questions intended to gauge which game elements motivated participants to act, and which they enjoyed. The survey decomposed each CD into two sub-questions. The questions were presented as statements that allowed the participant to respond on a 7-point Likert scale from ‘Strongly Disagree’ to ‘Strongly Agree.’ These 16 questions were delivered twice with the preface of “I do things that…” (changed to “I choose to do things that…” after the first iteration of the survey with the classroom participants) and “I enjoy games/activities that…” The purpose of this was to determine if there was a difference between which CDs one acts upon and which CDs one enjoys. The former group is hereby referred to as ‘CD: Act’ (Core Drive: Actions) questions and the latter will be called ‘CD: Game’ (Core Drive: Games/Activities) questions. In the post-experiment survey these questions are tailored to the CEH platform and are hereby referred to as ‘CD: CEH.’

The ordinal responses to the statements, based off of (Vagias 2006), were recorded as follows: 1-Strongly disagree, 2-disagree, 3-Slightly disagree, 4-Neither agree or disagree (sometimes referred to as ‘Neutral), 5-Slightly agree, 6-Agree, and 7-Strongly agree. In this paper, a ‘sub-question’ refers to one of the two questions of a CD. A sample of the survey format for this question type can be seen in Figure 19. The two sub-questions in Figure 19 are associated with CD: Act - CD1. This same Likert scale is also used in various question in the post-experiment survey.



	Strongly disagree	Disagree	Slightly disagree	Nether agree or disagree	Slightly agree	Agree	Strongly Agree
I choose to do things that make me feel like I am serving a higher purpose than myself							
I choose to do things that make me feel like I am fighting for a greater good							

Figure 19 – Sample question format with Likert-scale responses.

Yukai-Chou was contacted during the survey creation process about how to best ask survey questions to gauge the Octalysis model of a person. He was asked to specifically provide feedback on the proposed survey questions about how well they fit with each CD, but no response was received until after data collection and analysis were complete.

One of the messages I sent read "Hello, this is Landon Tomcho. I reached out a few months ago asking for advice on how to best ask survey questions to determine a person's Octalysis Model. I've since created the survey and wanted to share what I have with you." … "I asked people about "what they do", "what they enjoy" and "what the CEH has" based on 16 questions, two for each core drive. For context, the CEH is a website that I'm helping develop. The response options were on a Likert scale from Strongly disagree to strongly agree. Here are the 48 questions: …" Yu-kai Chou's response was "I like what you have here, but people are often too biased in self assessments" … "I think these would be better if they were trade off questions" and suggested that participants would have to choose between Core Drives in certain scenarios.

Sub-question statistics

The results of the 7-point Likert scale responses were mathematically evaluated through the following measures: average value, standard deviation, minimum and maximum of the responses to each sub-question; the absolute difference between the average of each sub-question within a CD; the visual representation of each CD within resulting Octalysis models for CD: Act and CD: Game questions; and the percent difference (based on the entire scale) in each CD between the CD: Act and CD: Game question groups. By performing this mathematical analysis, we are making several assumptions, but this allows us to more easily draw conclusions from the data than we could from a strictly categorical analysis. This ordinal to numerical translation assumes that the difference in all responses are close to equal: for example, a participant has the same difference in feeling between ‘Slightly agree’ and ‘Agree’ that they do between ‘Strongly disagree’ and ‘Disagree,’ and so on.

It should be noted that there has been controversy over whether ordinal data can be treated as interval data (Carifio and Perla 2008); some suggest that using means, standard deviations, and other parametric statistics may not give the best results and median values, frequencies, and other tests should be used instead (Jamieson 2004). Some have also argued that sample sizes must be large and the data should be normally distributed to use the metrics as we are using here such as mean and standard deviation (Norman 2010). However, (Norman 2010) has taken a deeper look at these criticisms, dissects them, and has determined that parametric tests can be used to analyze Likert scale responses.

In addition to reporting the standard deviation of the sub-questions, researchers considered reporting the coefficient of variance (CV) of each CD sub-question. The CV is calculated by dividing the standard deviation by the mean and reporting the result as a percentage. However, this metric did not add value to us because all responses used the same 1 to 7 scale, and so the CV did not show anything different than the standard deviation when comparing relative magnitudes. Also, (Neill 2017) tells us that a limitation of this metric is that it “is difficult to decide what constitutes high and low consensus based on CV values; therefore, application and interpretation of CV may be difficult.” Similarly, it is difficult to say what standard deviation values are significantly high or low. Some researchers have used a consensus metric where they define ‘agreement’ of participants to be when 80% or more of the responses are within the same category of importance. The article (Phillips et al. 2014) uses an 11-point Likert scale and separates these 11 values into 4 categories of importance. For this iteration of the study, we have decided to continue with the aforementioned metrics. Since each CD has two associated sub-questions, they may be referred to as CD1.1 and CD1.2, for example. CD1 would correspond to both sub-questions, either combined as an average or the difference (δ) between the two.

Equation (1), below, shows how percent difference is calculated in this paper. It is used to compare differences in sub-question statistics between different question groups within participant groups.

$$\%\ Difference = \frac{abs(x - y)}{Average(x, y)} \quad (1)$$

Scaling Response Values

In the analysis portion of the paper, the average values from sub-questions are taken and translated into a -10 to 10 Octalysis scale. The resulting values are input into the online Octalysis tool (“Octalysis / Gamification Building Developing Online Tool - by Yukai Chou” n.d.) which uses a 0 to 10 scale to create visual Octalysis Model representations. A value of 0 in the -10 to 10 scale corresponds directly with the baseline ‘Neutral’ category. We are primarily interested in which CDs do actually motivate the participants and what CDs they enjoy, only average responses past ‘Neutral’ toward the ‘Agree’ values are shown on the Octalysis Model visualizations in this paper; negative values appear the same as a ‘0’ would. Table 1, below, shows how different values are scaled using the following equation:

$$Output = Input * \frac{10}{3} + \left(-\frac{40}{3}\right) \quad (2)$$

Input (Likert)	Calculation	Output (Octalysis)
1	1*(10/3)+(-40/3)	-10
4	4*(10/3)+(-40/3)	0
7	7*(10/3)+(-40/3)	10

Table 1 – Example Likert-scale inputs converted to a -10 to 10 scale.

Air Force Education and Training Platform Use

A small group of questions asked participants how often they used other military training/education platforms This provided a baseline expectation for how members may use the CEH website that can be used for tailoring expectations when looking at the website tracking/engagement data. Participants were asked if they 'use the Air Force's Advanced Distributed Learning Service (ADLS),' 'take the DoD Cyber Awareness Challenge,' and 'use military education/training platforms such as milSuite, milTube, etc.' on a daily basis, never, or somewhere in between. The 'only when required' option was added to see if there is any voluntary use of these platforms or if it is strictly requirement driven.

D. Post-Experiment Survey

The post-experiment survey was delivered to the classroom group and the 88 CS participants of the operational group. To take this survey, the participant must have volunteered for the experiment portion of the study which consisted of signing up for and logging into the CEH website at least once. If the user did not interact with the website after their initial setup, the survey is designed to capture the reasoning behind that lack of engagement as much as it is setup to capture feedback from someone who logs on daily. Data collected includes how participants used and felt about the platform, general feedback on the website, and also feedback on the specific Topic Map (Tomcho and Reith 2019) (Tomcho et al. 2019) and KSA Tree features. CD: CEH questions were used to generate an Octalysis Model of the CEH for comparison with the Octalysis Models of the CD: Act and CD: Game questions.

Cyber Education Hub User Experience

This section asked the participant about how much they enjoyed the CEH website, if this website motivated them to consume educational content from other sources, and if they are motivated to use the website in the future; each question was presented as a statement with a 7-point Likert response scale based on agreement, just like the Likert questions in the base survey. The participants were then asked about their content consumption on and off of the CEH. The participants were also asked when/where they accessed the website. In early iterations of the survey, participants were asked to give open-ended responses to two questions: ‘Reasons why I used CEH,’ and ‘Reasons why I did not use CEH.’ After initial data was collected these were turned into checklist questions based on common answers. An option labeled ‘Other’ was provided for both in case they had a unique answer to add.

Gauging the Octalysis Profile of the Cyber Education Hub

The 16 questions in this section were nearly identical to the two sets of 16 used in the base survey. Instead of reading “I choose to do things that …” these questions read “Using the CEH makes me feel like…” or simply “The CEH …”. For example, the four questions below stem from CD4 and CD7. Two questions come from each, respectively.

- Using the CEH allows me to own content on the site.
- The CEH allows me to customize the site.
- The CEH has unpredictable elements when it comes to content.
- The CEH piques my curiosity.

Topic Map and KSA Tree User Experience

For both the Topic Map and KSAT, participants were asked if the element helped them consume more educational content and if the layout of each element was intuitive. This gives insight to the designers about what elements need improved and how the elements are being used. Participants were also asked to suggest improvements for both the Topic Map and KSATs. There was a question similar to the 'reasons why I used the CEH' question both of these elements. Again, the question was transformed from an open-ended response to a checklist based from common responses and developer insight.

Next, questions were based on different modes of content presentation or layouts, and the participants were asked about their preferences. They were also asked how familiar they were with each layout. This allowed the analysis to exclude the opinions of those who were unfamiliar with certain styles. The next few questions relate to each specific KSAT and how much the participant enjoyed them and how well they did at motivating the participant to learn about the respective topics. The end of the survey focuses on who should be in charge of the Topic Map and KSATs and if/where they would like to see them implemented, like AFIT Degree Programs, for example.

E. Website Tracking

As part of the study participants were given access to the CEH website, and their engagement with the platform was tracked over various lengths of time, depending on their associated unit/group. Classroom group engagement activity was tracked for 10 weeks over the span of the Introduction to Cyber Warfare program and for 4 additional weeks after the program's conclusion. These participants (14 of 15 students in the class) were able to optionally use the CEH website to accomplish certain class objectives including literature review and completing critical thinking checks (CTCs). CTCs could be accomplished by either responding to a question prepared for each lesson or by finding a recent article (document, presentation, video, news article, etc.) and sharing it with the class along with their comments on how it is relevant to class discussion for that lesson. The former option required an emailed response to the instructor and the latter option could be performed on the CEH website.

The 13 88 CS participants were tracked over a 12-week period. The 13 participants of the 33 NWS were tracked over a 2-week period. The cut-offs for these periods are on the same date, and the disparate start dates were due to logistical reasons with preparing for and accessing each unit. Computer networks used by these participants were whitelisted and able to connect to the CEH. The platform is also behind a military CAC authentication system, further limiting access to the platform. Participants were provided the opportunity to gain access to the CEH on their home networks, and 1 member from each group of 13 operational unit participants took advantage of this. In contrast, 12 of 14 classroom group participants requested home access. This difference in access should be noted when comparing the engagement numbers of each participant group.

Extraction Requirements

Requirements for the user engagement tracking data were based off of website features and user capabilities. Potentially noteworthy user actions such as viewing content, navigating to different pages, and utilizing unique website features were of specific interest to researchers. The requirements for what specific data was tracked and how it would be aggregated is detailed below. Only results pertaining to our primary research questions are presented in the main portion of this document, additional results can be found in the Appendix, beginning on page IX-80.

Query Timeframes and Experiment Groups

In order to pull collective data specific to user groups and specific times, it was a requirement to be able to query the tracking database for data within set timeframes and from different experiment groups. Since the initial intent was to have control and experimental groups, the ability to pull data from each of these sub-groups was also required. Within the set timeframe and experiment group, participant-specific results were required.

Participant-Specific Data

Participant ID

Each participant was assigned a unique Experiment ID (EID) number derived from a hash function on another individually-unique number stored in the website's database. This EID provided anonymity to the participants for their survey responses and tracking data results. As such, it was a requirement to deliver the participant-specific data with the EID. Some participants did not participate in the surveys and of those who did, some elected not to provide their EID and as a result their demographic and other survey data could not be paired with their engagement data.

User Activity

Other engagement data was tracked based on user activity on the CEH website. The activity groups and associated data can be seen in Table 2. Most of the activity groups are self-explanatory. 'Navigation' deals with the user switching between different displays such as 'Home', 'Help', 'Search', etc. For searches, the searched text was also collected. The 'Topic Map Activity' group included the collection of users clicking on nodes within the Topic Map to retrieve the associated content. Tracking data from each click included information about what node was clicked and what Topic Map was the source. For this experiment, there was only one Topic Map, but there may be more in the future.

Activity Group	Date and Time	Content Item Title	Content Item Type and Source	Comment Length	Topic Map Associated Nodes
Sessions	x				
Viewed Content	x	x	x		
Liked Content	x	x			
Comments	x	x		x	
Content Item Creation	x	x			x
Navigation	x				
Topic Map Activity	x				x
KSAT Activity	x				

Table 2 – Details associated with each tracked activity group for each participant.

The 'KSAT Activity' group encompassed numerous activities including opening a KSAT node, clicking on an activity with an embedded link, marking an activity as complete, and marking a completed activity as incomplete. Each of these included details about the source KSAT Tree, the node name, and the activity name, if applicable.

Testing

Researchers completed test runs with users not involved in the experiment to ensure that data was collected accurately and in the proper format. Three volunteers completed ordered tasks on the website while researchers recorded activity details such as time, content name, search details, etc. This information was compared against the database, corrections were made, and volunteers were iterated through until tracking was satisfactory. One unexpected issue, not discovered until post-experimental data collection, was a result of users viewing, commenting on, etc. content items that were later removed from the database. The tracking data still showed that users had taken action but the associated content name, type, etc. was absent. Thus, the data shown later in this paper does not reflect roughly 20 user actions due to deleted content items.

F. Tracking Extraction

Tracking data was extracted from the database in JavaScript Object Notation (JSON) format. JSON format was chosen for compatibility reasons, ease of human interpretation, and ease of parsing. Final tracking-data collection for all participants yielded a 1MB JSON file that was over 25,000 lines long. Researchers developed a roughly 2,000-line 73KB Python program to parse the data and create figures and tables to convey results. This software yields an engagement visualization capability that CEH developers would not have otherwise.

The JSON-parsing program is made up of four files. The primary file (Main) holds the high-level functions to be called by users. Main calls functions from a file which is responsible for creating the data plots and tables (Plotter). Plotter calls functions from the other two files which contain subfunctions (Subfunct) and helper routines

(Helper), respectively. The external libraries used by these files include: json, matplotlib and matplotlib.pyplot, sys, numpy, csv, and datetime.

The program is intended to be modified for future use by CEH engineers and potentially CEH users. As such, the code was developed with software engineering practices in mind to ensure an efficient handoff of the project. Practices used include functional decomposition, use of global variables where appropriate, 'self-commenting' function and variable names, and additional commenting where necessary. Function calls and parameters are setup to be easily modified for input via a GUI, to be developed in future work.

Figure Creation

The parsing program is able to output many different plots, each conveying different data or different combinations of data. Some of these plots are used to help answer current research questions and others may be used for future questions/purposes. Aggregate site activity, compiled from all specified users, can be displayed in one plot or separate plots for each of the following activities: logins, views, likes, comments, contributions (created content items), and searches. Each of these can be plotted over daily, weekly, or monthly intervals. User sessions can also be plotted where each user has their own timeline on the same plot, or where each user has a unique plot. User sessions can be plotted simply as logins, or also for additional activity including views, likes, comments, and contributions. Navigation activity is also plottable per user, showing user views, searches, home page accesses, Topic Map accesses, and KSAT accesses.

User interaction with the Topic Map and KSAT can also be plotted. Topic Map activity includes navigations to the Topic Map, clicks on topic nodes, and views from all sources. The KSAT activity plot shows views from all sources, KSAT accesses, KSAT node opens, activity links clicked, activity completions, and activity reversions.

Table Generation

The JSON-parsing program also creates comma separated value (csv) files that can be opened by spreadsheet utilities like Microsoft Excel to create data tables. These tables include more detailed but more aggregated information about the tracking data gathered from the website. The program extracts data from all users within specified groups to create tables that show aggregate information. The next sub-sections detail each table type. If you prefer not to explore the specifics, please navigate to page 120.

Sessions

The first table we will discuss presents session information for the group including the total number of logins, number of sessions binned based on duration, and the longest session in the group. The binned durations are as follows: short, 30 minutes or less; medium, 31-59 minutes; long, 60 minutes or more. Sessions automatically timeout after a half-hour of inactivity. Thus, users who login, check for new content and leave the site without any action are placed into the short session bin; users who view one or two videos and use the logout feature may also be placed into this bin.

Another function outputs each individual's session information. This table includes the user's first and last login date and time, number of logins, total session time, and average session time. Another Excel spreadsheet was built to take this information and calculate each user's activity span and the user's RoP (ratio of presence). These metrics are used when evaluating the addictiveness of online games, although span (number of days between first and last login) may be referred to as 'subscription period' and the RoP calculated in this paper is slightly different that the RoP metric defined in (Lou et al. 2012). In this paper RoP is the number of logins divided by the user's span on the website but (Lou et al. 2012) calculates RoP by dividing number of days with at least one login over the user's span, ensuring RoP is never greater than 1. The same Excel spreadsheet used for these calculations also calculates the group's average number of logins, average total time, average session duration, average span, and average RoP.

View Counts

The next two tables detail the total number of content item views by: type, subtype, raw sources, and refined sources. Content item types include web resources, videos, or files. Subtypes for each include webpage and YouTube; video; and .pdf, word document, PowerPoint, etc. Sources are the website page or feature from which the user found the content item such as 'Home-Trending' or 'Suggested Content-Recently Added.' The table with raw sources includes KSAT activities with their unique identifiers such as 'Activity-072881fb-…' and searches with full text included. The refined version combines all KSAT activities together and all searches together.

Comment Information

Comment information is binned and tabled based on number of characters. The bins are: very short, 100 or less characters; short, 101-205 characters; medium, 251-999 characters; long, 1000-2501 characters; and very long, 2501 or more characters. For example, "Wow! Great video, thank you!" is 28 characters and "I found this article when I was scrolling through my morning news feed and thought it might be relevant to some issues I've been seeing in the community. Anyone else care to share their perspective or thoughts about us implementing this policy at the unit level?" is 263 characters. The character count of the longest comment is also stored.

Classroom CTCs were expected to be 300-350 words long according to the syllabus. Since the average English word length is about 4.5 characters (Shannon 1951) the CTC comments were expected to be 1500 or more characters when accounting for spaces and punctuation. Thus, without individual inspection, we assume that comments binned as very short, short, or medium are not from CTCs.

Navigation

Tables with raw and refined navigation data can also be created. These tables show the count of user navigation to different pages on the website. Similar to the raw view data, searches with all text included are included in the raw navigation table; the difference is that this includes all searches, not only those that led to content views. The raw navigation table also includes information about the specific content items that users edited, like 'ContentItemProperties-469ecdd3-…' for example. Searches and edits are condensed to their respective collective counts in the refined table. Creating a separate table with search terms and counts is also an option.

Topic Map

Two separate tables related to the Topic Map feature can be generated by the program. The first is related to content contributions. The table presents the number of content items associated with each individual Topic Map node, the overall count of contributed items, and the count of contributed content with at least 1 Topic Map association. In the future, the content item type would be a useful piece of information to be stored in the JSON file with contributed content to easily count the types of contributed content. The second table presents each Topic Map node that was clicked and the associated total number of clicks. The node size information would also be useful to include in the JSON file in the future in to more easily analyze if the number of clicks or contributions correlated with the node's size.

KSAT Activity

KSAT activity is available in 5 separate tables, 3 for KSAT nodes and 2 for the activities within the nodes. The counts of node openings, links clicked, and net activity completions are presented for each KSAT node. Net activity completions specific to the activity itself is presented in a separate table. The final table breaks down net activity completion information by keyword. Some example keywords include verbs such as 'review' or 'skim,' and also difficulty levels such as 'easy,' 'hard,' or 'expert.'

Individual Activity by Number

The last table generated by the JSON-parsing program deals with individual user activity counts. This functionality was added after researchers found that survey data suggested that different demographic groups may be motivated by different things (detailed later in this paper). Since the website's JSON output was only specific to experimental groups and also did not have the survey demographic information, this table output was deemed necessary. By providing the EID along with the count of a user's logins, views, likes, comments, etc., researchers were able to pair EIDs from survey data with EIDs from engagement data. A separate Excel spreadsheet was used to place users into appropriate demographic groups and calculate average values, and more to determine if there were differences in engagement between groups.

G. Analysis Metrics

Data from tables output by the JSON-parsing program detailed above is presented in this paper and is evaluated with mixed methods. Only results pertaining to our primary research questions are presented in the main portion of this document, additional results can be found in the Appendix, beginning on page IX-80. Quantitative methods for these tables include average values and standard deviations (σ). Demographically-separated tracking data is presented and compared to across subgroups. The data presented are average values for each individual within the subgroup, standard deviations (σ), and maximum and minimum values within the subgroup.

The following metrics are compared between groups: logins, views, views per login, span, total time, average time, RoP, likes, comments, contributions, contributions w/ Topic Map associations, Topic Map contribution %, Topic Map clicks, KSAT node navigations (opens), KSAT links clicked, KSAT links per navigation, KSAT net activities completed, and KSAT activities completed per node. To determine where there may be significant differences between groups, we calculate the percentage difference between groups based on each average value via Equation (1).

If there was a 20% or greater difference in the averages between groups a single sided t-test was executed on the engagement data to determine if there was a statistically significant difference between groups. Significant results are reported based on a minimum 80% confidence value (an alpha value of 0.2). For our purposes, confidence value is 1 minus the p-value, which is the probability of falsely rejecting the null hypothesis of the t-test. The null hypothesis for each test is that 'there is *no* significant difference between participant demographic subgroups in the specified metric.' Thus, with an 80% confidence value, we can say that there is an 80% chance that we are correctly rejecting the null hypothesis and accepting the alternate hypothesis that 'there *is* significant difference between participant demographic subgroups in the specified metric.' The '% Confidence' values are calculated via Equation (3) where p is the p-value from the t-test. The specific t-test we used was a version of Welch's t-test (Welch 1947) which allows us to compare groups of unequal sample sizes.

$$\% \, Confidence = 1 - p \tag{3}$$

H. Group Comparisons

Question Groups

To analyze the differences in responses between the CD: Act, CD: Game, and CD: CEH questions, we look at the spread between each average response compared to the full response scale. Researchers considered using a conventional percentage difference equation like Equation (1), where the absolute value of the difference is divided by the average value to yield a percentage, but this yields different results based on how the values were scaled. Thus far, we have used a 1 to 7 scale and also a -10 to 10 scale, both based off of the ordinal Likert-scale responses. In either of these scales, a '1' and a '2' would yield a 66.7% difference, even though they represent significantly different responses on the different scales. In the 'spread' approach, a '1' and '7' would yield a 100% difference in the Likert scale and '-10' and '10' would yield the same difference relative to the Octalysis tool scale; everything else falls somewhere in between. Also, the resulting percentage value will be the same regardless of whether or not the values have been transferred to another scale (as long as the multiplier has also been adjusted). We decided to calculate all percent spread values based on the 1 to 7 scale, for simplicity and also because this is closer to the raw ordinal data. Equation (4) and Table 3 show this method.

$$Output = abs(x - y) * \frac{1}{6} \tag{4}$$

x	y	Calculation	Output
1	7	abs(1-7)*1/6	100%
1	4	abs(1-4)*1/6	50%
4	5	abs(4-5)*1/6	17%
4	6	abs(4-6)*1/6	33%
4	7	abs(4-7)*1/6	50%

Table 3 – Example inputs/outputs for Equation (4)

For further visualization and ease of analysis, we use a gradient color scale to demonstrate the magnitude of the scale-based differences. Figure 20 shows an example of the color scale based on the difference of a value from the scale baseline, which is a 'Neutral' response (also 'Neither agree or disagree'). An 'Agree' response would appear orange with a value of 33% when compared to a 'Neutral' response, for example. Values smaller than 17% are a fainter yellowish-white (although black Is used in the image below), and any value over 50% is the same intensity of red, but this would be a very extreme case. The scale below is only specific for values compared to a '4' value and in the case where $x = 6$ and $y = 7$, for example, we would get a result of 17% and a yellow cell.

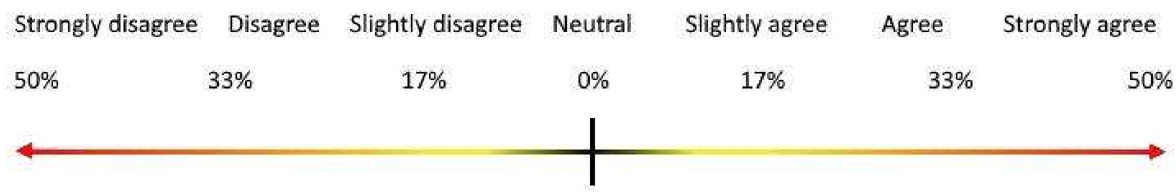


Figure 20 – Color gradient based on percent spread of responses

Demographic Groups

All responses are aggregated based on various demographic groupings and analyzed similarly to how they are within the classroom and operational groups. This data is also split into other demographic-based groups and compared, primarily via Octalysis Model representations and percent spreads.

I. Gauging Motivation Differences Between Groups for Specific Platforms (Framework)

During the study, researchers collected several datasets on participants via surveys and the website-based experiment. Before we proceed with detailing how we approached the analysis on this data, we present several assumptions that have been made in this unique study:

- The Octalysis Framework can be used to compare participant survey answers to their platform activity based on game elements within respective core drives.
- The survey questions and survey format accurately and precisely captured participant motivation and enjoyment levels for respective core drives.
- Differences between demographic groups in terms of motivation and/or enjoyment can lead to differences in platform engagement based on featured game elements and their associated core drives.

- Platform features can be extracted into a small subset of core drives, and differences between groups with respect to their engagement with those features will be reflected as differences in their Motivation Levels with these Core Drives.

Along with these assumptions, there were other questions. Researchers had no prior work to refer to when considering what exactly may be responsible for differences in activity when comparing data between groups, for example:

- Will participants with the highest motivation with CDx seek out features with CDx on the platform more than other users regardless of the presence of CDx on the platform?
- Will participants with the least difference between their motivation and their thoughts of the platform respective to CDx interact with CDx features on the website more than participants with a larger difference?
- What CDs most influence total use of the platform (e.g. Number of logins, total time spent on site, etc.)?

Researchers sought a way to reduce these assumptions and begin to answer some of these questions. Figure 21 shows the data that was collected during the study that will be used to align data and draw conclusions about the aforementioned research questions and listed assumptions/inquiries.

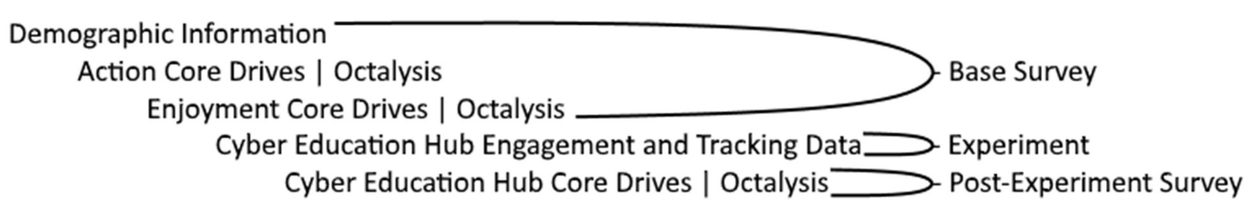


Figure 21 – Data collected during the study that is used to analyze differences in subgroup behavior and respective sources.

In theory, researchers with ample understanding of the Octalysis Framework could inspect all the data listed in Figure 21, find significant differences, and report the findings strictly based on individual or collective qualitative analysis. When dealing with research that may be unique or the first of its kind, perhaps this would be appropriate, but it is not as generalizable as we would prefer.

In this case, we are interested in seeing how the survey data listed in Figure 21 can be combined to draw conclusions about whether the differences in motivation models based on the Octalysis Framework can be aligned with differences in engagement data. Essentially, "do participant responses about what they do, what they enjoy, and how they feel about a platform actually relate to how they used the platform."

Instead of inspecting and categorizing the data haphazardly, researchers developed a framework, that we hope can be generalized. The framework shown in Figure 22 takes user motivation data and user-reported platform data, and outputs a Motivation Level that summarizes how that user (or group of users) may respond to the platform, specific to a certain motivator. In this case, the motivators we are inspecting are the Octalysis CDs.

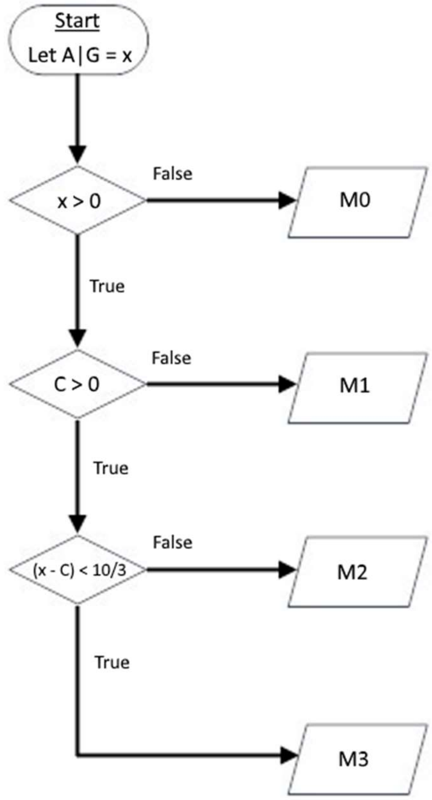


Figure 22 – Framework used for determining Motivation Level toward a specific platform.

The process used to take the survey data listed in Figure 21 and prepare it for input into the framework in Figure 22 was:

- Translate each individual participant's 7-point Likert scale responses into values ranging from -10 to 10. Via Equation (2).
- Combine responses at the individual level to represent their average for each Octalysis CD for each of the base-survey question groups (Action questions and Game/Activity Enjoyment questions in this study) and post-experiment survey question groups (CEH questions in this study).
- Determine if there are differences in motivators between demographic groups.
- Split data based on demographic group and question group and calculate averages on the -10 to 10 scale.

Next, we will walk through the framework in Figure 22 using our study's dataset as an example. In order to begin, we input two small pieces of data: a piece of user motivation data, and a corresponding piece of user-reported data about the platform. When we say 'corresponding' we mean that the data is related to the same motivator. In this example, we will input the average participant Action question CDX value for demographic subgroup A; and the average participant CEH question CDX value for demographic subgroup A. The conditions are that the averages are on the same scale, -10 to 10 in this case; the averages come from the same demographic group, A=A; and the motivators are the same, CDX = CDX.

In Figure 22, 'A' represents a value calculated from the 'Action' survey question group, 'G' represents a value calculated from the 'Game/Activity Enjoyment' survey question group, and 'C' comes from the 'CEH' question group. These question groups are referred to as 'CD: Act,' 'CD: Game,' and 'CD: CEH' questions in this paper, respectively. Essentially, the 'x' value is data about the participant, and the 'C' value represents how the participant (or participants) feels about the target platform.

In the first decision node, we check if the participants feel positively about the corresponding motivator. If the 'x' value is 0 or less, then the average participant responded that they feel 'neutral' about the motivator, or even 'disagree.' If this is the case, we have reached Motivation Level 0: "User Doesn't Care" (M0). In this scenario, the user (or demographic subgroup) is not necessarily interested in this motivator, so they may not be affected by corresponding game elements on the platform.

If the first comparison returns as true, we check if the user reported that the motivator is present on the platform. If the 'C' value is 0 or less, we remain at Motivation

Level 1: “Motivator Not There” (M1). In this situation, the user has expressed an interest in the motivator, but according to them, it is not apparent on the platform. Such users may search the platform for anything that resembles a feature with this motivator, testing out the platform for a short period of time, and ultimately give up.

The final decision node checks if the motivator’s presence on the platform is ‘close enough’ to the value that corresponds to the average user in the group. If the platform’s motivator is within 10/3 on the -10 to 10 scale (1 Likert value on the 7-point scale), or if ‘C’ is greater than ‘x’, then we enter Motivation Level 3: “Fair” (M3). Otherwise, the motivator is present within the platform, but is not quite at the level the average user in the subgroup desires and we stop in Motivation Level 2: “Needs Some Care” (M2).

After we have categorized the relationship between each motivator relative to the user and the platform, we can begin to compare them between subgroups and get one step closer to explaining differences in user engagement/activity. The group with a higher total number of motivators at M3 could likely correlate to the same group uses the platform more in terms of logging in, overall time on the website, etc. Another metric that may provide insight could be the total number of higher classified motivators between groups. In simple situations, different groups will have distinct Motivation Levels for corresponding motivators, but this may not always be the case. Also, when looking at the more specific elements within a platform, related motivators and their Motivation Level differences between groups may account for differences in engagement. Thus, for comparing the number of higher values, or when looking at specific elements it may prove useful to have ‘tiebreakers’ if the motivators are at the same Motivation Level in both groups.

We propose two tie-breakers to be used in case motivators are at the same Motivation Level, other than M0. If the Motivation Level relative to that motivator has been deemed ‘not present’ (M0), it does not make much sense to see which is ‘less’ non-present. The first tiebreaker can be used when two motivators are at M1. In this scenario, both user groups are potentially driven by the motivator, but both have responded that it is not present on the platform. Here, we expect the subgroup with the higher ‘x’ value to be more motivated and will likely engage more with elements on the platform that at all resemble this motivator.

- M1 Tie-breaker: Higher ‘x’ value wins.

The second tie-breaker is to be used when both Motivation Levels are at M2 or when both are at M3. In this scenario, participants from both groups are driven by this motivator and have responded that it is present in the site. We expect that the less difference there is between the motivator’s presence on the site (or greater surplus), the more that respective group will engage. So, for this comparison, the lower resulting value when calculating ‘x-C’ is deemed the winner. In cases of negatives (which can only occur at M3), it is awarded to the more negative value, but could be deemed a tie as long as the evaluator is consistent for all comparisons of this type.

- M2 | M3 Tie-breaker: Calculate ‘x-C.’ Lower resulting value wins.

3. Results and Analysis

In this section we will look at the results of the base survey and post-experiment survey results for both the classroom group and operational unit participants. Additionally, these groups are combined and we look at the overall results as well as comparisons between sub-groups within the entire participant pool. Results are derived by taking the raw survey responses and using the methods described in §2 to knead out more digestible data. Only results pertaining to our primary research questions are presented in the main portion of this document, additional results can be found in the Appendix, beginning on page IX-67 for survey results and IX-80 for engagement tracking results.

A. Classroom Base Survey

The post-experiment survey was delivered 7 weeks after participants' first exposure to the CEH website. The class consisted of 15 college graduates. All but one student volunteered for the experiment, 12 participated in the base survey, and 7 participated in the post-experiment survey.

Participant Demographics

The survey volunteers included one enlisted member, one government civilian, and 10 Air Force and Army officers. The officers were primarily company-grade officers with two-thirds of the participants having less than 6 years of service. 8 of 12 volunteers also had no cybersecurity work experience. 17% of the participants had cybersecurity certifications such as CompTIA's Security+. The career fields of the volunteers are: Operations Research Systems Analysts (2), Computer Scientist (1), Cyber Transport Systems (1), and Cyberspace Operators (8).

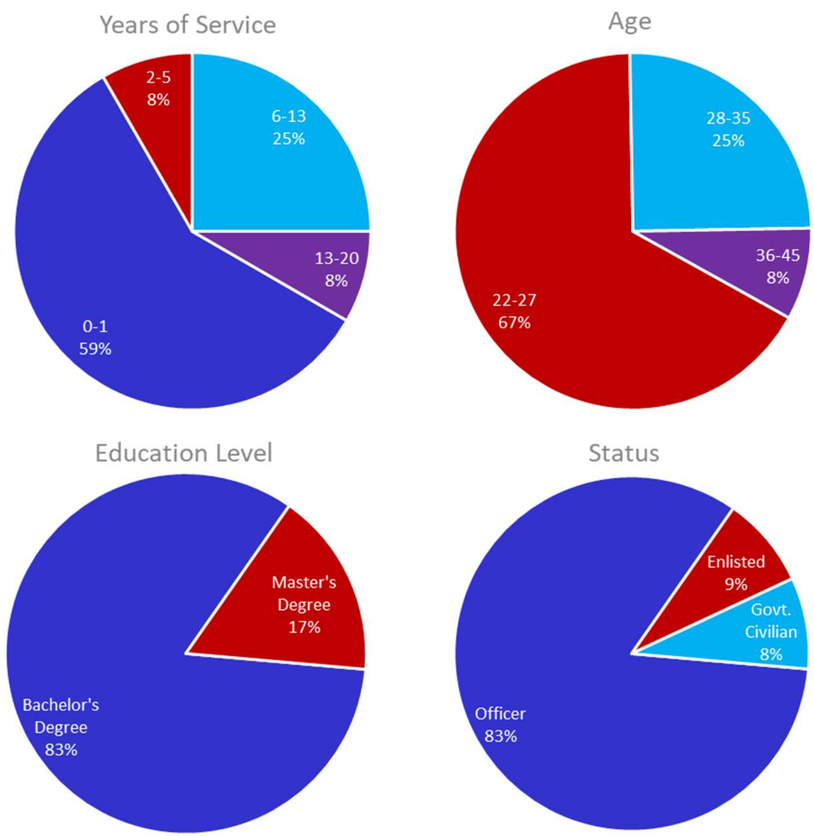


Figure 23 – Some demographic information about the AFIT classroom survey participants

Gauging the Octalysis Model of the Participants

Core Drive Questions

Table 25 (Appendix C) shows the results for the CD: Act questions from the classroom participants and Table 26 (Appendix C) shows the results for the CD: Game questions. When considering what responses may be the most important/urgent for consideration of integration to the CEH website, designers may want to focus on sub-questions with smaller σ values, with the intent of adding design features that consistently appeal across more users. For the purposes of this paper we will not focus as much on these statistics, especially within each participant group.

Generating the Octalysis Models

Table 4 shows the average and scaled response values for each CD for both question sets. This includes the data of all participants in the survey from the classroom group. Note that the averages are that of both sub-questions within each CD.

	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8
CD: Act | Likert	6.13	6.13	5.33	5.33	5.79	4.63	5.58	5.38
CD: Act | Scaled	7.08	7.08	4.44	4.44	5.97	2.08	5.28	4.58
CD: Game | Likert	4.50	6.00	5.13	5.38	5.88	5.50	6.25	3.63
CD: Game | Scaled	1.67	6.67	3.75	4.58	6.25	5.00	7.50	-1.25

Table 4 – Average Likert and Scaled values for each CD.

Below in Figure 24 are the visualizations of the scaled CD: Act and CD: Game Octalysis Models for all participants from CSCE 525. The Octalysis Model for the CD: Game questions is somewhat different, meaning that what participants do is, on average, different than what participants enjoy in games, at least relative to some CDs. This Octalysis Model for CD: Game questions is weak in CD1 and CD8, but is fairly balanced otherwise.

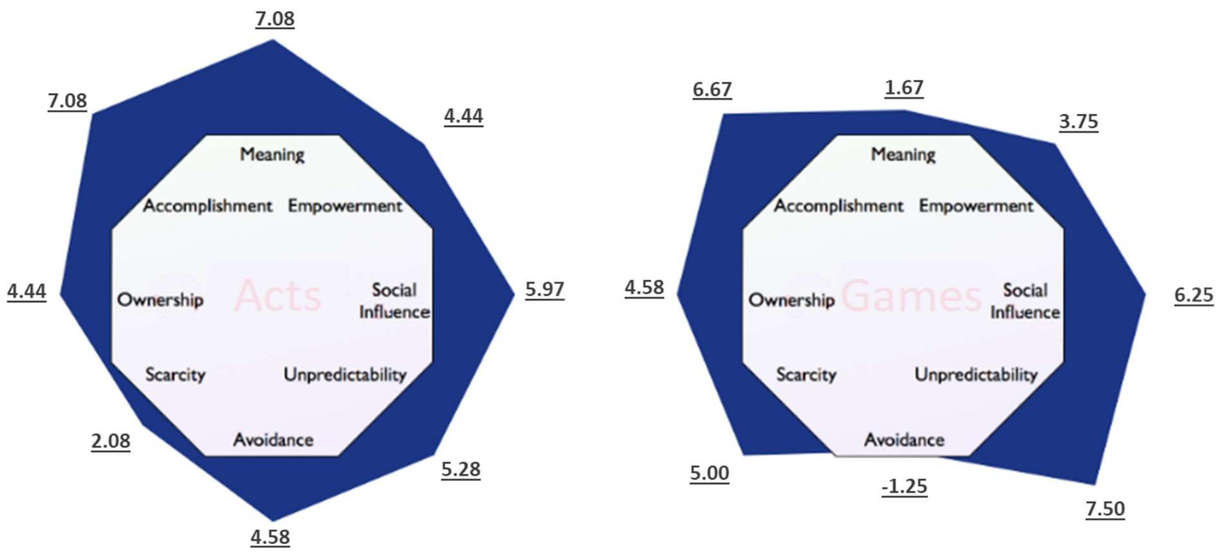


Figure 24 – Octalysis Models for CD: Act and CD: Game statements for classroom group participants.

From the data presented above, we can see that on average, this group of participants felt differently about some of the CDs involved in ‘what they do’ and ‘what they enjoy in games/activities.’ In actions they take, participants are highly influenced by Epic Meaning & Calling and Development & Accomplishment. In the games/activities they enjoy, Development & Accomplishment, Social Influence & Relatedness, and Unpredictability & Curiosity are the most apparent.

The only responses that were below “Neutral” (toward the “Disagree” statements) were the statements about CD8: Avoidance & Loss with respect to what the participants enjoy in games/activities. When making future design decisions about the CEH website based on this data, developers should take this information into account and prioritize game elements and features that align with the statements above and avoid, if possible, elements that align with the following two statements:

- “I enjoy games that I have to play in order to avoid losing my progress”
- “I enjoy games that I have to play in order to avoid missing opportunities”

We hope that this sample size is representative of the population of target users for the CEH website, but it may only be representative of those who take Introduction to Cyber Warfare at AFIT. If the latter is true, then this data is still quite useful when considering how to motivate students in a classroom to engage with the lesson and activities.

Use of Air Force Education/Training Platforms

The data shows that of those participants who had ever used the Air Force's Advanced Distributed Learning Service or taken the DoD's Cyber Awareness Challenge, 100% of participants used them "Only when required." As for milSuite, milTube, etc. most users responded that they never use it, but some use it monthly and one answered 'weekly.' Since the intent is that the CEH this is a voluntary-use based platform (Reith et al. 2018), it will be a good gauge to see how often users log-in to the website. If the developers have successfully implemented gamification techniques and ensured that the platform is motivating for users to interact with, we predict users should engage more than they do with these other platforms. As rough metric of success would be bringing typical 'monthly' users of training platforms back weekly and 'rarely' or 'never' users back at least monthly.

Electronic Device Use

The questions about electronic device usage should be removed from future iterations of this survey. This change is already reflected in the most current version of the survey (Appendix A, page IX-52). This survey borrowed some of these questions from a previous researcher's pre-experiment questionnaire but no useful data was collected from this portion of questions.

B. Classroom Post-Experiment Survey

7 of the 15 students in the class submitted a completed post-experiment survey, and all 7 were among the 12 that took the base survey. All 7 participants in this portion of the survey were military officers, and two-thirds had one year of service or less. This half of the survey focused on gathering participant-reported data about how they used the platform, what elements they did and did not enjoy, and gauging the Octalysis profile of the platform itself.

Use of the Cyber Education Hub Website

All 7 participants stated that they were able to access the website from home with Common Access Cards and white-listed IP addresses. 29% said "Over the course of the class I consumed more content from outside sources than I did on the Cyber Education Hub" and 57% affirmed that "Over the course of the class I consumed more content on the Cyber Education Hub than I did from outside sources." This is likely partly because the pre-class readings/videos were all hosted on the CEH website, but it should also be noted that the links to each were also provided separately, allowing students to have no dependence on the CEH website. To demonstrate this non-dependence, we reiterate that only 14 of 15 students registered to use the website; one of those 14 only logged in on six days over the 10-week, 20-lesson program; no students failed the class. It should also be noted that course objectives and grading were the same as in previous offerings that did not include the opportunity to participate in this study.

Based on the same 7-point Likert scale mentioned previously, translated into numerical values, we present data from questions based on user enjoyment of the Topic Map, KSAT, and the platform overall. Table 5 shows that the average resulting values were between 'Slightly agree' and 'Agree.' Only one individual responded 'Slightly disagree' once among all 3 questions and there were no other responses below 'Neutral.' This suggests that while the CEH website is not perfect (to be expected from a platform still under development) it did not seemingly turn anyone off with regard to future use. The beta website was delivered in a state where it was usable and valuable (to some), but improvements can certainly be made. It may also be important to note that the average participant 'Slightly agree[s]' that the platform motivated them to seek more outside educational/training than they would have otherwise; an effect that may otherwise go unnoticed.

Statement	Average	σ	Min	Max
I enjoy using the Cyber Education Hub	5.29	1.03	4	7
Using the Cyber Education Hub motivated me to consume more outside educational/training content	5.14	0.64	4	6
I am motivated to continue using the Cyber Education Hub in the future	5.29	1.39	3	7

Table 5 – Classroom group responses relating to overall feelings about the CEH

Topic Map and KSA Tree Questions

The average response to many of the statements about the Topic Map feature and the KSAT feature were between 4 ('Neutral') and 5 ('Slightly agree'). The most positive and least varying results are that of the questions relating to the intuitive use of the Topic Map and KSAT; the intuitiveness of the layout of each follows closely behind. The only other statement with an average response value close to 5 ('Slightly agree') is that the KSAT and Topic Map influence more content consumption on the CEH. These two statements' responses had some of the highest standard deviations, along with the questions about preference of the Topic Map, KSAT, pre-requisite style, and YouTube-style content layouts. The variety in responses suggests that each participant has their own preferences. What may be most meaningful is that the CEH's unique elements, the Cyber Topic Map and KSATs, were preferred over other more common layouts by some participants; this alone warrants the existence of these elements.

Statement	Avg.	σ	Min	Max
I consumed more content on the Cyber Education Hub than I would have if I did not have access to the Topic Map	4.86	1.73	2	7
I consumed more educational content OUTSIDE of the Cyber Education Hub than I would have if I did not have access to the Topic Map	4.14	0.99	3	6
The layout of the Topic Map was intuitive	4.71	1.39	3	7
The use of the Topic Map was intuitive	5.29	0.88	4	7
I prefer the Topic Map layout to the YouTube-style layout of content	4.00	1.51	1	6
I prefer the YouTube-style layout to the Topic Map layout of content	4.50	1.76	2	7
I prefer the Topic Map layout to the prerequisite-style layout of content	3.71	1.48	1	6
I prefer the prerequisite-style layout to the Topic Map layout of content	4.71	1.75	2	7

Statement	Avg.	σ	Min	Max
I consumed more content on the Cyber Education Hub than I would have if I did not have access to the KSA Tree	5.14	2.36	1	7
I consumed more educational content OUTSIDE of the Cyber Education Hub than I would have if I did not have access to the KSA Tree	4.14	0.99	3	6
The layout of the KSA Tree was intuitive	5.14	1.25	3	7
The use of the KSA Tree was intuitive	5.71	0.70	4	7
I prefer the KSA Tree layout to the YouTube-style layout of content	4.33	1.83	1	7
I prefer the YouTube-style layout to the KSA Tree layout of content	4.17	2.10	1	7
I prefer the KSA Tree layout to the prerequisite-style layout of content	4.29	1.48	1	6
I prefer the prerequisite-style layout to the KSA Tree layout of content	4.43	1.84	2	7

Table 6 – Statistics for classroom group responses to questions about the Topic Map and KSAT features on the CEH.

The data in Table 7 suggests that the participants had mixed feelings about KSATs if you do not take the separate topics into account. The participants in the group we are currently discussing were all in the CSCE 525 class, and thus were either required by AFIT to take the class or took it optionally, so they had some investment in the topic. As we can see below, the average response for each question about the CSCE 525 KSAT was about 1.5 "categories" higher on the 7-point Likert scale than the responses about the

Mobile Technology KSAT. The overarching Topic of Mobile Technology was chosen as the comparison topic because it encapsulates many focus areas of cyber, including networks, software, hardware, adversaries, and vulnerabilities. The idea was that there would be something for anyone on that KSAT. With the CSCE 525 KSAT the participants felt 'Neutral' at worst about the enjoyment of the topic and the specific nodes' challenges and tasks. The good news is that some members enjoyed it, and overall, the participants 'Slightly agree' with statements about the CSCE 525 KSAT.

Statement	Average	Std dev	Min	Max
I enjoyed the CSCE 525 KSA Tree challenges and tasks	5.57	1.05	4	7
I enjoyed the KSA Tree topic of CSCE 525	5.14	0.99	4	6
I was interested in CSCE 525 before using the KSA Tree	5.71	0.88	4	7
The skill tree motivated me to learn about CSCE 525	4.57	1.40	3	7
I enjoyed the Mobile Technology KSA Tree challenges and tasks	3.43	0.73	2	4
I enjoyed the KSA Tree topic of Mobile Technology	3.67	0.47	3	4
I was interested in Mobile Technology before using the KSA Tree	3.67	0.47	3	4
The KSA Tree motivated me to learn about Mobile Technology	3.17	1.07	1	4

Table 7 – Results relating to each KSAT topic available to classroom group participants

Overall, the classroom participants showed mixed feelings towards the KSATs and Topic Map. Some users preferred them over other more typical content layouts and the tools motivated them and fostered enjoyment. Others were not entertained and would rather stick to what they know. The open-ended responses/suggestions about the Topic Map had few responses, and the findings are summarized as follows:

- Improve Topic Map scaling/zooming
- Add search feature
- Useful for content contribution
- Lacking some topics

Contrasting responses to open-ended questions can further demonstrate the point that participants had mixed responses to the KSAT feature.

Question: “Suggestions for improvement of the KSA Tree:”

- Response A: “Align it with the CSCE 525 syllabus better”
- Response B: “No suggestions, I thought it was very helpful and encouraged me to learn more. I would have probably explored more if it was *not* paired with the structure of the 525 class.”

The responses above about the KSATs and Topic Map demonstrate that these features are valuable to some users, but could also use improvement. Different users like different things, and one of the main principles of the CEH is the idea of crowd-sourcing (Reith et al. 2018). When asked who should generate Topic Maps and KSATs, participants responded with many different answers. Therefore, instead of focusing on creating the perfect KSAT or Topic Map for all users, developers should focus on creating the tools to allow Airmen to create and influence their own (or community) Topic Maps and KSATs. 79% of participants said that they would like to see Topic Maps or KSATs implemented in other applications.

Cyber Education Hub Octalysis Model

The same approach used to gauge the Octalysis Profiles of participants earlier (2 groups of 16 questions and a 7-point Likert Scale) was used to determine the model of the CEH. The average resulting values shown in Table 27 (from CD: CEH question responses) are compared with those of the CD: Act and CD: Game questions later in this section.

Table 8 shows the average responses when combining the two sub-questions from each CD. Just as before, the data is shown on the 1 to 7 Likert scale as well as the -10 to 10 scale used to create the Octalysis Model visualization, seen in Figure 25. The resulting Octalysis Model based on the participants' responses was fairly weak compared to the information we gathered about what motivates users to act and what they enjoy. Keep in mind that this user group had at least 7 weeks to access the platform and had the opportunity to become quite familiar with it. In the future it may be interesting to see how user responses to these 16 questions change based on exposure time to the website.

	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8
CD: CEH | Average	4.00	5.57	3.57	4.79	4.86	4.14	4.71	3.86
CD: CEH | Scaled	0.00	5.24	-1.43	2.62	2.86	0.48	2.38	-0.48

Table 8 – Average Likert and Scaled values for each CD for CD: CEH questions.

Something that immediately stands out is that this model is not quite as positive as either of the two previous models. In fact, the values are smaller or more negative on nearly every CD. Before asserting any negative statements about the CEH, we must consider that these questions are focused on a specific platform, whereas the earlier question sets were more abstracted and asked the user what they do, in general, and what they enjoy, in general.

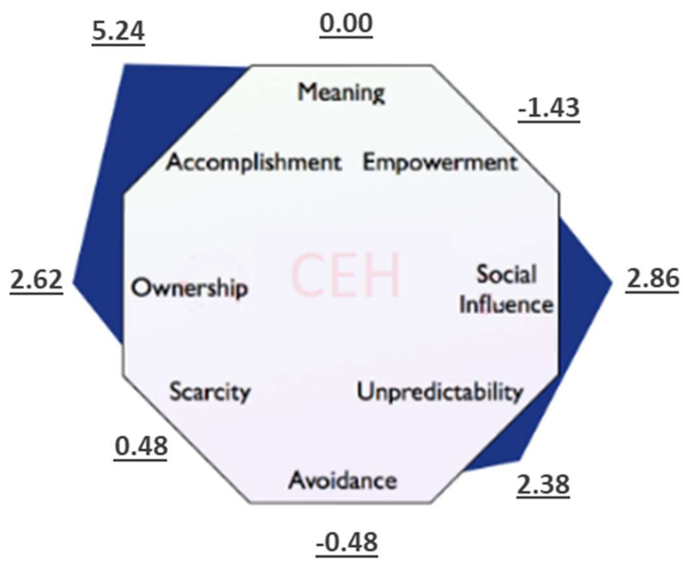


Figure 25 – Octalysis Model for CD: CEH questions

Comparing Responses for each Octalysis Question Group

When considering what features/enhancements to prioritize, researchers should not necessarily look to bolster the CDs with the lowest response values first. Perhaps, what is more important is the difference in what users do/enjoy and what the CEH currently looks like. If using the first technique, researchers may elect to strengthen CD3 through milestone unlocks and meaningful choices and then move on to enhance the presence of CD8: via countdown timers and loss of progress. But using the second method, researchers would prioritize CD6 and CD7 after CD3 (based on the difference in CD: Games and CD: CEH, as we will see later in Table 9). A framework is later presented that helps us further explore this idea.

Table 9, below, shows the spreads between the CD averages from each 16-question group among all participants in CSCE 525. Later in this paper we separate the data into several different groups and analyze the results. The Octalysis Models for each of the question groups are placed side-by-side in Figure 26, for further visual comparison of the data.

% Spread	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts - Games	27%	2%	3%	1%	1%	15%	11%	29%	11%
Games - CEH	8%	7%	26%	10%	17%	23%	26%	4%	15%
Acts - CEH	35%	9%	29%	9%	16%	8%	14%	25%	18%

Table 9 – Percent spread for each CD compared to each question group for classroom participants.

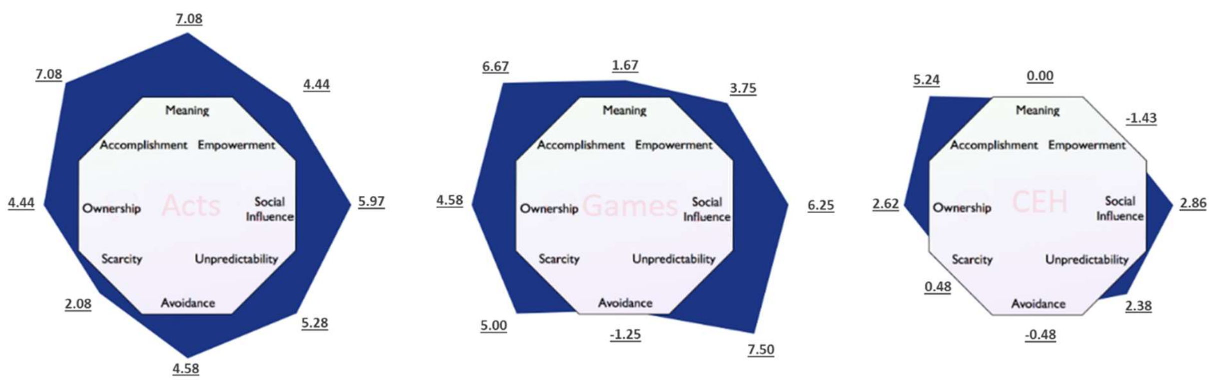


Figure 26 – Side-by-side comparison of each Octalysis Model based on the classroom participants' responses

Here we will focus on the comparisons between CD: CEH questions and the remaining two Octalysis Models. The largest spread is between CD1: Acts – CEH. However, we can also see that there is a large value for CD1: Acts – Games and also a small value for CD1: Games – CEH. So, what does this mean? It means that there is a large variation between 'what participants do' and 'how participants feel about the CEH' but, there is not much of a difference in 'what participants enjoy in games' and 'how participants feel about the CEH' for 'Epic Meaning & Calling.' Since we do not yet have data on which CDs, and which of the 'do' versus 'enjoy' formats, most correlate with a participant's use of the CEH we will focus less on large spreads between CD: Act and

CD: Game and instead focus on CDs with small spreads in this row (top row of Table 9) and large spreads in the other two rows. Secondarily, we will focus on large spreads between CD: Games and CD: CEH responses; perhaps enjoyment is more important.

The most substantial data point above is the lack of CD3 in the participants evaluation of the CEH. This finding is very surprising, given that some developers considered that this was the strongest natural CD for the CEH. CD3 relates to being able to be creative and receiving feedback on how you are doing. The CEH allows users to import whatever content they choose, or build their own content, and share it with the community who can give feedback on content items via liking and/or commenting.

We can see that there are also significant variations in the responses for CD7: Unpredictability & Curiosity and also CD6: Scarcity & Impatience. These are both black hat core drives that can inspire engagement, but should not be overpowering, or users can be driven away from the platform due to fatigue (Chou 2015). On a more positive note, the CEH shows smaller differences in responses compared to the Act/Game questions for both CD2 and CD4, suggesting that participants felt that the CEH allowed them to develop themselves, track accomplishments, and own things near as much as they do/enjoy elsewhere.

Future Use of the Cyber Education Hub Website

CSCE 525 was the primary reason that participants were introduced to the CEH. When responding to the question about why they were or were not motivated to continue using the CEH website after the class was over, some participants responded:

- “I want to see what other people have found useful/interesting to learn about cyber. Also, I think it will be a useful tool for use across the Air Force.”
- “It is crucial to stay up to date with current cyber events and findings.”

- “Definitely a good platform - the problem is I don't reference the CEH when looking for articles. I find outside sources.”
- “Not enough useful training / CE content”
- “If the training was more tangible, I'd be more interested”
- “With the hub, I can keep up to date on cyber news and topics, without needing to search through other news sources.”

In responses to several different questions, participants noted that lack of time was a primary reason for not using the website more often:

- “Time required to sign in and the responsiveness of the web interface.”
- “Didn't have time due to other coursework.”
- “I was very busy during this period of time, so that limited the amount of time I could spend exploring the Topic Map and KSA Tree.”
- “Not enough time during the busy quarter.”

C. Operational Units Base Survey

As mentioned above, the survey was also delivered to volunteers from various operational units. From the 88 CS, 13 members volunteered and participated in the experiment and 13 members of the 33 NWS volunteered and participated in the experiment for a total of 26 operational personnel. Of these 26, only 9 had taken the base survey at the time this data was analyzed. At least 14 days have passed between distribution and this analysis. While the survey for the classroom-based study was delivered on Google Forms, this survey was delivered to operational participants via milSurvey, a tool available on milSuite. Certain DoD networks block Google Forms, so the CAC-protected alternative was used. There are likely many reasons that the response rate was so much different for the operational units compared to the classroom group. Perhaps the milSurvey site is not as easy to access from home since a CAC reader is required; while at work on DoD networks, access times may be slow; first-time users of milSuite must wait for a profile to build; operational units may have other priorities and do not see research as urgent/important as graduate students. Due to the smaller sample size, we will not perform as much of an in-depth analysis on this group, and instead look further into the aggregate results of all participants (classroom and operational groups together).

Participant Demographics

The operational unit population has a much higher percentage of enlisted members than AFIT, and this is reflected in the participant pool. The participants from the operational units also had a higher average age and more years of service than those in the AFIT classroom. The volunteers from the operational units all had at least an associate's degree and several had bachelor's degrees. This population group also had more (over 66%) cybersecurity certifications and work experience than those at AFIT (less than 33%) which is to be expected given the number of first-assignment students at AFIT. Figure 27, below, shows more detail.

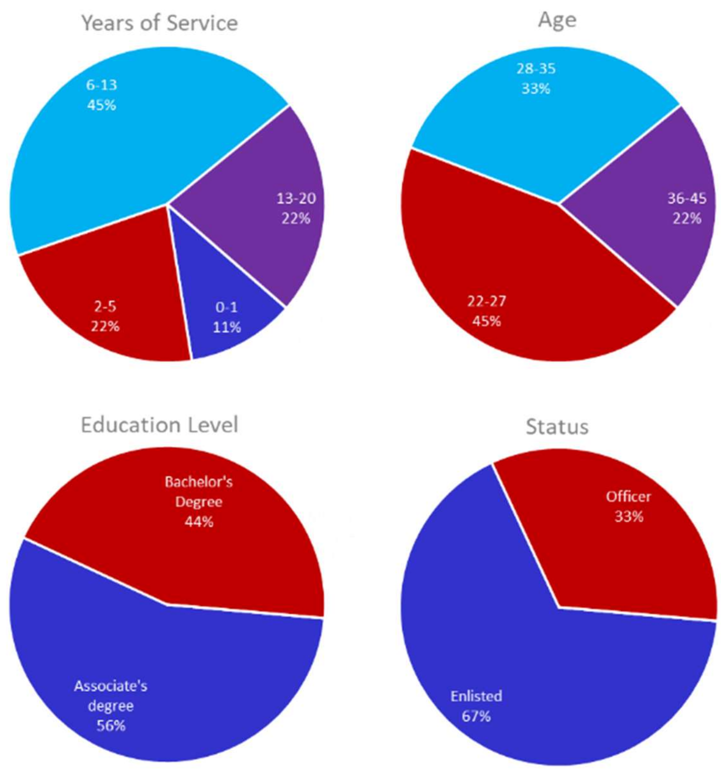


Figure 27 – Demographic data for operational unit survey volunteers.

Octalysis Models

For the operational units, gauging the Octalysis Models of the participants was conducted in a very similar manner of what was used for the AFIT classroom group. 2 groups of 16 questions each were presented to be responded to on a 7-point Likert scale. Each Octalysis CD had two associated questions in each group, previously named CD: Act questions and CD: Game questions. One change, mentioned before, is that the "I do things that …" statements were reworded to say "I choose to do things that …"

Core Drive Questions

The same methods shown in §2.C were used to gather and report the data shown in Table 29 (Appendix C) and Table 30 (Appendix C). Due to the smaller sample size of 9 participants, less analysis is performed until later in the paper when both the classroom and operational groups are combined and comparisons are performed based off of various demographic features.

Model Results

The results shown below in Table 10 are the averages taken from the responses to each CD's associated statements and are then scaled based on Equation (2). Similar to the results from the AFIT classroom, we can see that CD: Game CD1 and CD8 are the only responses that are on the 'Disagree' side of the Likert scale and the negative side of the Octalysis tool scale.

Question Group	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8
CD: Act | Likert	5.22	5.78	5.67	6.17	5.56	5.11	5.50	5.17
CD: Act | Scaled	4.07	5.93	5.56	7.22	5.19	3.70	5.00	3.89
CD: Game | Likert	3.94	6.00	5.67	5.72	5.17	5.61	5.61	2.83
CD: Game | Scaled	-0.19	6.67	5.56	5.74	3.89	5.37	5.37	-3.89

Table 10 – Average Likert and Scaled values for each CD.

The Octalysis Models for both the CD: Act and CD: Game questions can be seen below in Figure 28. These models are quite similar in shape to those in Figure 24. The most noticeable differences are that this group has a stronger draw toward CD4 and less from CD1 with respect to actions. Overall, we can see that the CD: Act Octalysis Model is quite well-rounded with all responses stronger than 'Slightly agree' responses. In addition, just like in the previous group, CD1 and CD8 see a very significant drop-off from CD: Act to CD: Game responses. The percent spread values calculated via Equation (4) can be seen in Table 11 along with the coloring scheme from Figure 20.

Figure 28 – Octalysis Models for CD: Act and CD: Game statements for operational unit participants.

% Spread	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts - Games	21%	4%	0%	7%	6%	8%	2%	39%	12%

Table 11 – Percent spread for each CD compared to each question group for operational unit participants.

Post-experiment Survey Results

As mentioned above there was only one response to the post-experiment survey for operational participants. The survey was not yet delivered to members of the 33 NWS due to timing issues. The one respondent was from the 88 CS. This may be due to the timing of the survey delivery (21 December and 28 December) or other previously hypothesized issues. The survey was delivered four weeks after users were first exposed to the website, and perhaps by this time they had lost interest or felt had not used the platform enough or recent enough to contribute. The response rate may be an interesting data point in itself. Nonetheless, the single response may prove valuable; its analysis is included in Appendix C.

D. General Engagement Results

Sessions

In general, the most platform use was from the group of participants in the classroom setting. These students were actively involved in a course directly related to cyber with their peers, and were able to optionally use the CEH website to complete certain class goals. However, after the class was completed, usage of the platform decreased below the average level of the operational unit users. This data can be seen in Table 12. Use of the platform may be directly related to an individual's personal feeling about the topic of cyber education, or the platform itself. Users in the classroom may have volunteered because they were interested in testing out a new platform, or perhaps they wanted to leverage its utility in the program. Thus, motivation to use the website may have dropped after the utility of the platform decreased.

In contrast, operational unit users did not have a significant reason to change behavior throughout the experiment period. Increasing the presence of operational units as well as increasing utility by making the CEH relate to more individuals, perhaps via additional KSA Trees, may help get the CEH past the tipping point (Gladwell 2002) where more and more users jump on the CEH bandwagon.

	All	CSCE 525 Class	CSCE 525 Post-Class	All Operational	88CS	33NWS
Participant Count	40	14	14	26	13	13
Exposure (# of Weeks)	x	10	4	x	12	2
Total Sessions	440	364	10	66	49	17
Short Sessions	52	46	0	6	5	1
Medium Sessions	270	214	9	47	34	13
Long Sessions	118	104	1	13	10	3
Longest Session	178	178	61	160	160	68
Average:						
Logins	11.0	26.0	0.3	2.6	3.8	1.3
σ	15.3	16.4	x	2.8	3.5	0.5
Span	36.0	69.9	x	12.7	23.8	1.6
σ	37.3	4.7	x	20.9	25.3	1.7
Total Time	577.5	1393.1	x	123.1	191.4	54.8
σ	804.6	828.0	x	168.6	220.4	23.4
Time per Session	48.1	54.6	x	44.6	46.7	42.5
σ	13.2	10.5	x	13.3	12.2	14.5
Ratio of Presence	0.6	0.4	x	0.7	0.5	0.9
σ	0.4	0.2	x	0.4	0.5	0.2

Table 12 – User session data.

The amount of long sessions (over 60 minutes) is promising. Some users recorded over two consecutive hours, even within the operational group. On average, user sessions are 45-55 minutes long. When accounting for the 30-minute session timeout this suggests that the average sessions consists of viewing a few content items and leaving. Standard deviations for the number of logins per user suggest that there is a high variation in platform use between individuals; some users rarely log in after registration and some users check back once or twice per week. Users that only log in once inflate the average RoP values also. When you have 1 login over 1 day and someone else has 20 logins over 60 days, the former will have a RoP that 3x greater. Thus, RoP is not always the best measure of usage or platform addictiveness, as noted in (Lou et al. 2012).

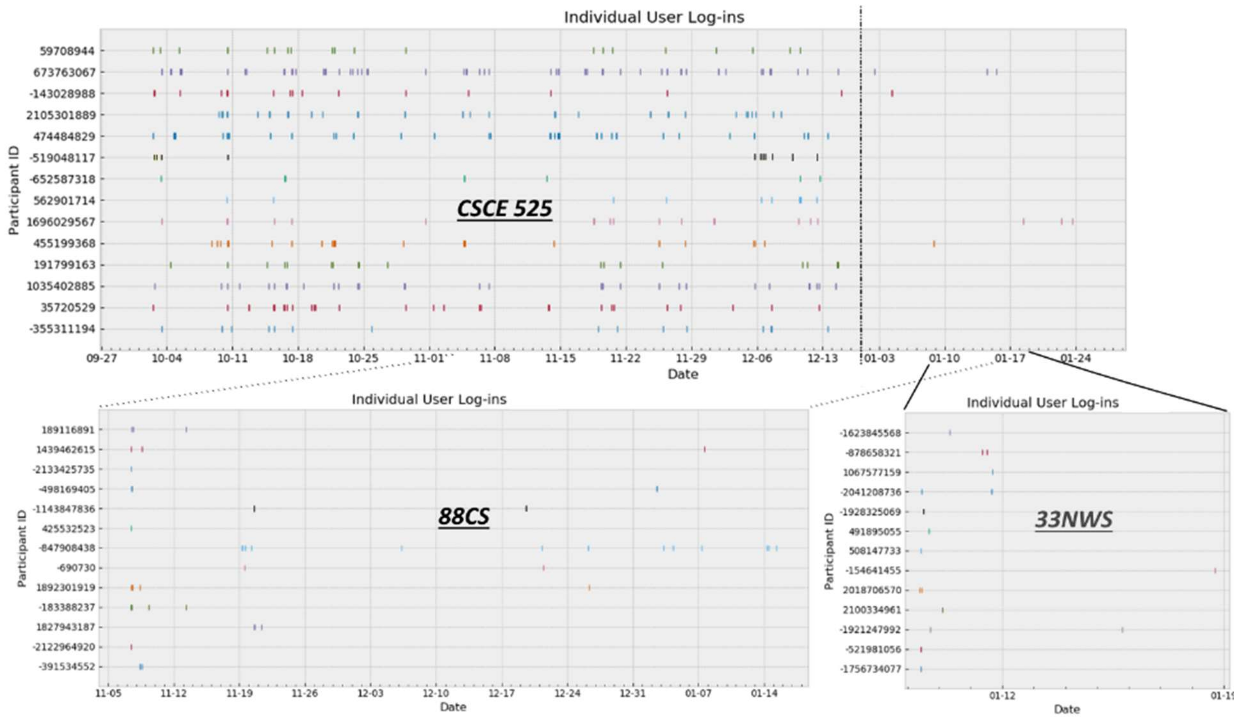


Figure 29 – User sessions for each experiment group.

Figure 29 shows a plot of each user session against the time axis. There was no user activity from December 16th to January 1st for the classroom group, so this time was cut from the plot, demonstrated by the dotted vertical line. Some operational users from 88 CS logged into the platform 3 or more weeks apart, suggesting that we may expect similar behavior from 33 NWS users in the future. Due to time restrictions, only 2 weeks of data was collected from 33 NWS participants.

For the purposes of comparing against other platforms, we characterize platform use for each group except the 33 NWS due to limited data. For the 88 CS, out of 13 participants, there was 1 weekly user, 5 monthly users, and 2 participants who only signed on once; the remaining 5 are considered rare users. For the classroom group during the 10-week class period 6 of 14 participants are classified as daily users, 5 are considered weekly, and the remaining 3 used the platform monthly. After the class was over, 4 participants logged on during the one-month period.

Figure 30 shows the total and average session duration for classroom participants during the 10-week class period and the data's associated trendlines. The average session time was fairly consistent, but dips near November 8th when no users logged in for a 5-day span caused the average times to be 0 minutes on those days.

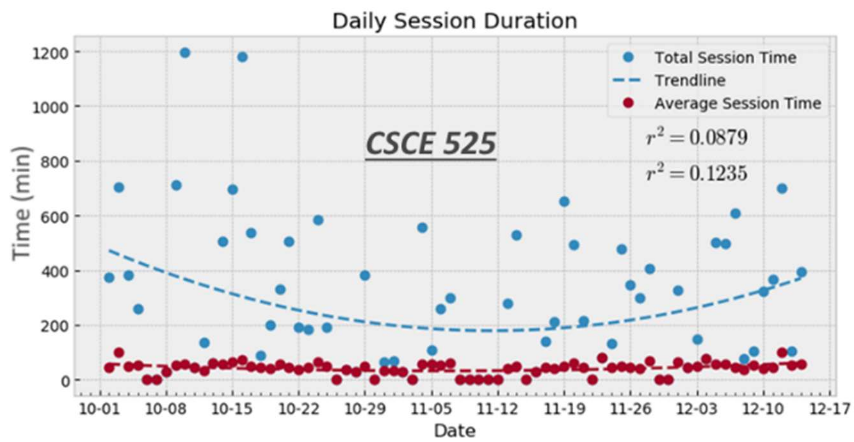


Figure 30 – Daily session duration plot for classroom participants.

Navigation

The final data we present in this subsection relates to user navigation on the CEH website. This information can be seen in Table 13 for each user group. This table shows that other than the home page, users were mostly contributing or editing content items ('ContentItemProperties' page). This table does not include data about viewing content items, so we cannot say that 'users were contributing content more often than they were consuming it.' We can compare the total view count of 754 with the total count of items contributed (111) and the number of navigations to the 'ContentItemProperties' (361) and conclude that the aforementioned statement is unlikely. As we assumed, most users did not use the ‘Logout’ feature and about 30 minutes can be subtracted from each session duration when we consider active time of use. However, we reiterate that time spent on external websites or using downloaded files is not accounted for.

	All	CSCE 525 Class	CSCE 525 Post-Class	All Operational	88CS	33NWS
Home	709	549	12	148	117	31
ContentItemProperties	361	318	6	37	32	5
KSAT-user	190	172	4	14	7	7
KSAT-ksat2	135	124	3	8	2	6
Search	79	71	8	0	0	0
MyContent-mycontent	72	68	1	3	2	1
TopicMap	52	30	6	16	11	5
SubmitFeedback	16	5	0	11	8	3
MyContent-Recently Added	14	7	0	7	6	1
User	11	3	1	7	3	4
KSAT-ksat1	10	3	0	7	3	4
Logout	9	5	0	4	3	1
Help	6	0	0	6	6	0
SubmitFeedback-Form	3	0	3	0	0	0
MyContent-Protect and Defend	3	0	0	3	1	2
MyContent-Strategy	1	1	0	0	0	0

Table 13 – User navigation data.

An interesting piece of data from this table is the disparity between the number of times participants navigated to the ‘KSAT-user’ page, which is where a user selects a KSAT to enter, and the combined number of navigations to specific KSATs (‘KSAT-ksat1,’ and ‘KSAT-ksat2’). 45 times, a user was either confused or lost motivation at the ‘KSAT-user’ page (‘KSAT-user’ minus ‘KSAT-ksat1’ minus ‘KSAT-ksat2’). We predicted that this was because the selection tool was not obvious enough and so the dropdown list was changed, as we can see below in Figure 31.

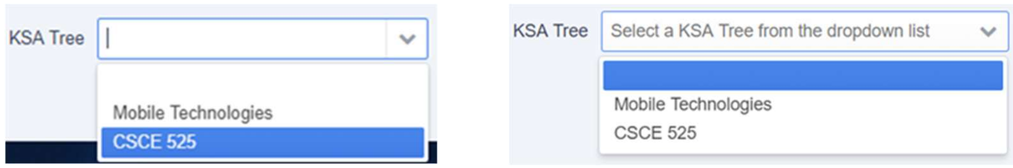


Figure 31 – Instruction added to dropdown list due to data from Table 13 (“Cyber Education Hub” 2019).

E. Overall Survey Results and Other Observations

We will now combine and analyze the data from the classroom group and the operational units. In total there were 21 participants in the base survey, 12 from the classroom group and 9 from operational units. Combining groups allows us to draw more generalizable conclusions and perform comparisons between larger demographical subgroups.

Participant Demographics

The participants in this survey were all affiliated with the Air Force as either an officer, enlisted person, or government civilian. Since the CEH is focused on cyber topics, Air Force cyber-related units and the Introduction to Cyber Warfare class at AFIT were targets for the study described in (Tomcho and Reith 2019) which encompasses this survey. Consequently, the career-fields of those involved are cyber-related. Due to the majority of participants coming from the AFIT classroom group, there is a greater percentage of young first-assignment officers than one would see in a typical unit.

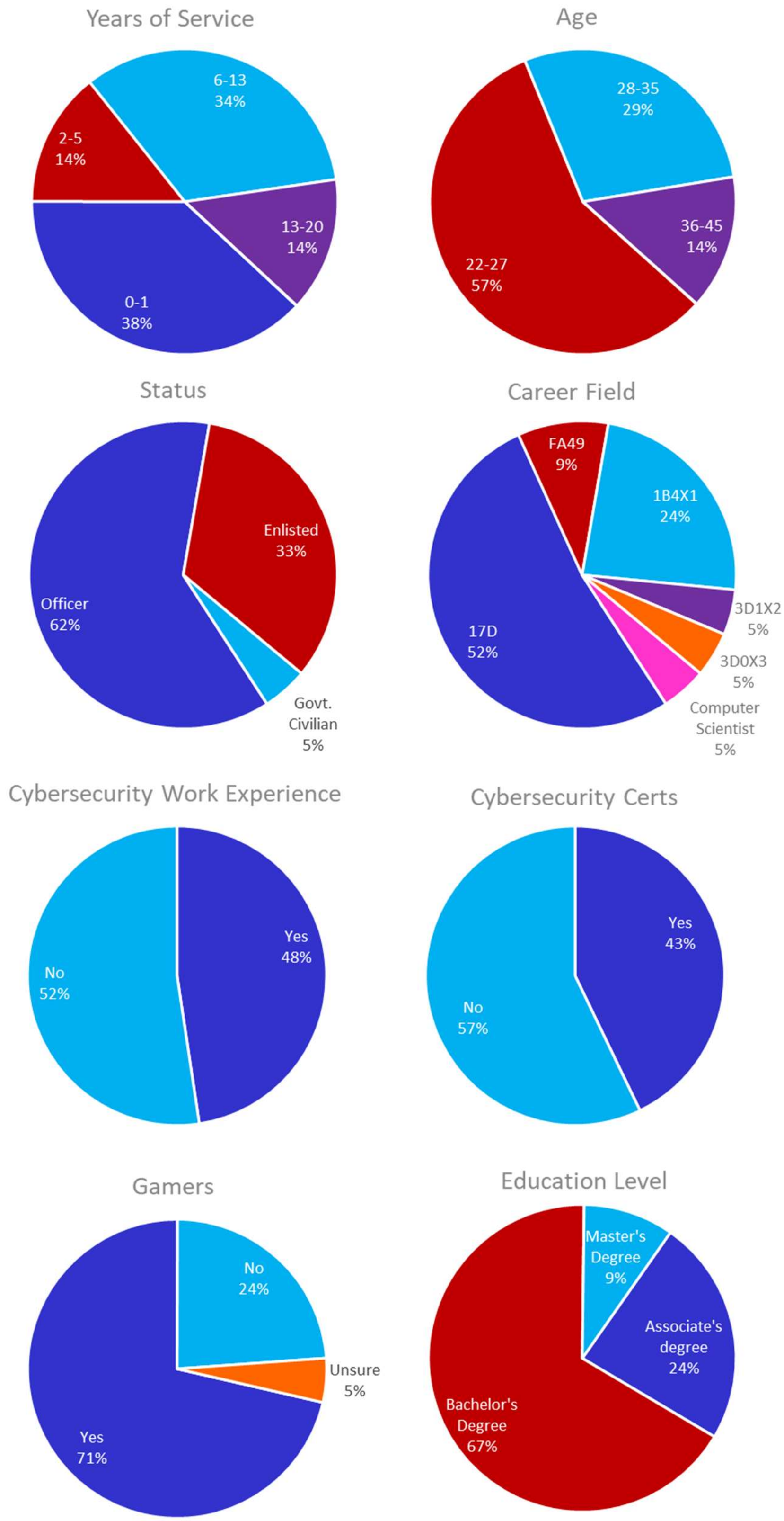


Figure 32 – Demographic information for all participants

Octalysis Models

This subsection will focus on determining the average Octalysis Models for all participants for both the CD: Act and CD: Game question groups. The methods used are more detailed in §2.

Sub-questions

The following bulleted list represents the sub-questions with the largest gaps between what motivates users (to act and what they enjoy) and what the CEH currently has to offer. These statements are presented in decreasing order beginning with the two largest gaps. While there are other gaps, they are not as considerable.

- CD5.2 "The CEH allows me to interact with peers/friends"
- CD3.2 "The CEH gives me feedback"
- CD4.2 "The CEH allows me to customize the site"
- CD6.2 "The CEH has difficult challenges to strive towards"
- CD3.1 "Using the CEH makes me feel empowered to use my creativity"
- CD7.1 "The CEH has unpredictable elements when it comes to content"

The following five sub-questions also showed significant gaps, but only in between CD: CEH and CD: Act (denoted 'A') or CD: CEH and CD: Game (denoted 'G'), but not both. Decisions about increasing CD8 should be very careful as many participants responded negatively toward these statements relative to enjoyment in games/activities.

- CD6.1G "The CEH has limited/exclusive elements"
- CD1A – serving a higher purpose, acting for a greater good
- CD8A – have to use the platform in order to avoid losing progress or missing out on opportunities

Overall Model Results

When the groups are combined and the sub-questions are averaged, no response value is greater than 6.00, which directly corresponds to an 'agree' response. However, all but one CD (CD: Game - CD8) averaged out to be on the positive side of the Octalysis tool scale.

Question Group	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8
CD: Act - Average	5.74	5.98	5.48	5.69	5.69	4.83	5.55	5.29
CD: Act - Scaled	5.79	6.59	4.92	5.63	5.63	2.78	5.16	4.29
CD: Game - Average	4.26	6.00	5.36	5.52	5.57	5.55	5.98	3.29
CD: Game - Scaled	0.87	6.67	4.52	5.08	5.24	5.16	6.59	-2.38

Table 14 – Average Likert and Scaled values for each CD.

The resulting Octalysis Models (Figure 33) are similar to those in Figure 24 and Figure 28, but more balanced when all participants' responses are combined. Spreads can be seen in Table 15. Changes between question groups (CD: Act and CD: Game) with respect to each CD group (white-hat, black-hat, extrinsic, and intrinsic) are not consistent.

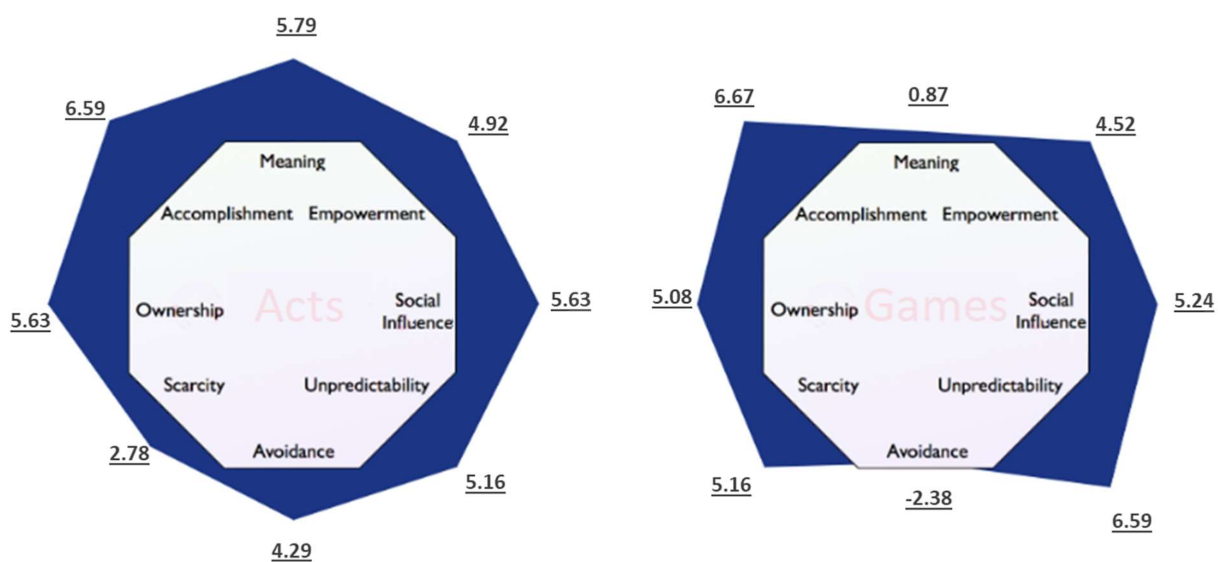


Figure 33 – Octalysis Models for CD: Act and CD: Game questions for all participants.

% Spread	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts - Games	25%	0%	2%	3%	2%	12%	7%	33%	11%

Table 15 – Percent spread for each CD compared to each question group for all participants.

Comparing Subgroups

Analysis of Variance (ANOVA) statistics were not calculated due to inequalities in the sample sizes of each group and more importantly that this data was already translated from ordinal to numerical. Performing this type of analysis does not necessarily add value when applied to categorical responses. However, we will compare the average response values for each CD and report the differences in spread between groups. The greatest differences were seen between Gamers and Non-Gamers and we will discuss these two groups here. To see the differences in Junior v Senior and Younger v Older, please see Appendix C.

Gamer v Non-Gamer Octalysis Models

This sub-group comparison was performed based on participants' responses to the question "would you consider yourself a gamer?" 15 participants (71%) affirmed and were placed into the Gamer group, while the remaining 6 (29%) were placed in the Non-Gamer group. The resulting Octalysis Models and spreads can be seen in Figure 34 and Table 16, respectively.

Spreads that are near 2 Likert categories different (33%) appear in CD6 for both CD: Act and CD: Game and CD4 of the CD: Game group, all of which show decreases. Spreads near 17%, or 1 Likert category are evident in CD: Act CD3, CD4, CD5, and CD7 and also CD: Game CD2 and CD3. All of these spreads also demonstrate a decrease in average value. One may consider "of course non-gamers enjoy certain game elements less in games than gamers," but note that the survey instructs participants "if you don't play/enjoy games, replace the word 'games' with 'activities' and answer the questions accordingly."

The individual CD spreads help account for the largest average spreads we have seen between demographic groups, 13% and 14%, for CD: Act and CD: Game questions, respectively. This suggests that what motivates and fosters enjoyment between gamers and non-gamers may be more critical than any other demographic comparison including Officer v Enlisted, Younger v Older, and Junior v Senior (although the latter reaches the closest exaggeration to Gamers v Non-gamers).

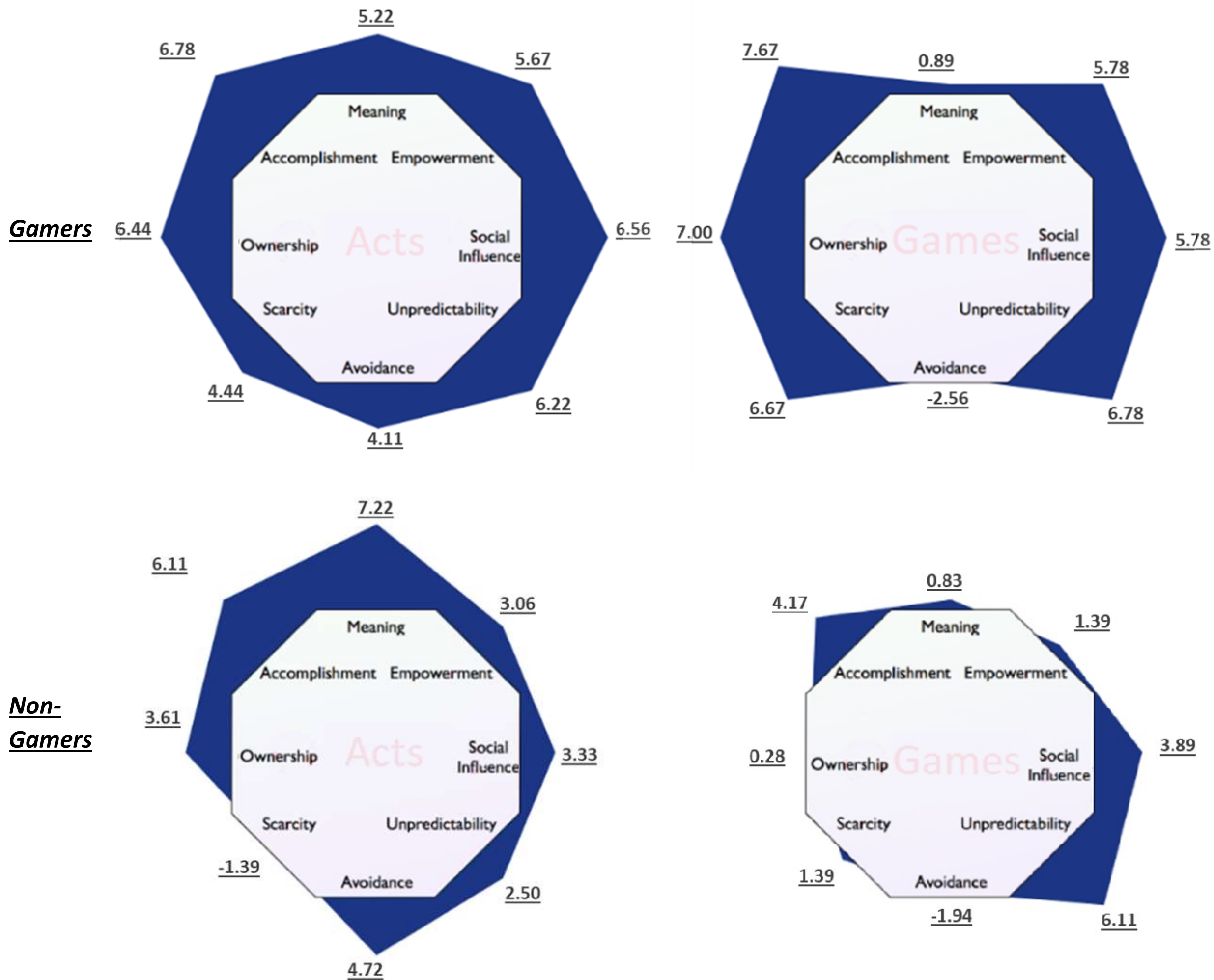


Figure 34 – Octalysis Models for Gamer and Non-Gamer participants.

Gamers v Non-Gamers	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	10%	3%	13%	14%	16%	29%	19%	3%	13%
Games	0%	18%	22%	34%	9%	26%	3%	3%	14%

Table 16 – Spreads for CDs of Gamer v Non-Gamer subgroups.

Gaming

The average Gamer was 25.7 years old with a σ of 3.5 years. Non-gamers had an average age of 33.3 with a σ of 6.3 years. Gaming activity data was captured in the base survey. Each self-identified Gamer plays on at least one of the listed platforms at least weekly, except one self-identified gamer who only plays on TV/Game console monthly. For those who did not identify as gamers, most partake in gaming monthly at most, except for 2 participants who partake in mobile gaming weekly. Essentially, we found that the self-reporting aligned with participants' gaming behavior.

Gamer v Non-Gamer Post-Experimental Survey Results

Since the data above suggests that the Octalysis Models of gamers and non-gamers are quite different, we will take another look at the post-experimental survey results for the classroom group through this lens. 57% of post-experiment survey participants identified as gamers in the base survey, so there was a near-even number of gamers and non-gamers accounted for in the data below. Overall, as we can see from the results of question 1, there was no large difference in feelings toward the CEH platform between these groups. The statements are taken from the survey version delivered to the classroom group. Some questions were modified in the most recent version which is shown in Appendix B on page IX-58. The question numbers in Table 17 represent the corresponding question in the most recent version.

Question #	Statement	Gamers	Non-Gamers	% Spread
1.1	I enjoy using the Cyber Education Hub	5.3	5.3	1%
1.2	Using the Cyber Education Hub motivated me to consume more outside educational/training content	5.0	5.3	6%
1.3	I am motivated to continue using the Cyber Education Hub in the future	5.3	5.3	1%
12.1	I consumed more content on the Cyber Education Hub than I would have if I did not have access to the Topic Map	4.3	5.7	24%
12.2	I consumed more educational content OUTSIDE of the Cyber Education Hub than I would have if I did not have access to the Topic Map	4.0	4.3	6%
12.3	The layout of the Topic Map was intuitive	4.3	5.3	18%
12.4	The use of the Topic Map was intuitive	5.3	5.3	1%
16.1	I consumed more content on the Cyber Education Hub than I would have if I did not have access to the KSA Tree	4.0	6.7	44%
16.2	I consumed more educational content OUTSIDE of the Cyber Education Hub than I would have if I did not have access to the KSA Tree	4.0	4.3	6%
16.3	The layout of the KSA Tree was intuitive	4.5	6.0	25%
16.4	The use of the KSA Tree was intuitive	5.5	6.0	8%
20.1	I am familiar with the YouTube layout, search and recommendation features	6.8	5.0	29%
21	I prefer the Topic Map layout to the YouTube-style layout of content	3.3	5.0	29%
21	I prefer the YouTube-style layout to the Topic Map layout of content	5.5	3.0	42%
22	I prefer the Topic Map layout to the prerequisite-style layout of content	3.5	4.0	8%
22	I prefer the prerequisite-style layout to the Topic Map layout of content	5.3	4.0	21%
23	I prefer the KSA Tree layout to the YouTube-style layout of content	3.8	5.0	21%
23	I prefer the YouTube-style layout to the KSA Tree layout of content	5.0	3.0	33%
24	I prefer the KSA Tree layout to the prerequisite-style layout of content	3.5	5.3	31%
24	I prefer the prerequisite-style layout to the KSA Tree layout of content	5.3	3.3	32%

Table 17 – Percent spreads for classroom group post-experiment survey responses between gamers and non-gamers.

From questions 12 and 16 we can see that the Topic Map and KSAT features were more intuitive for non-gamers. The responses also show that these features helped non-gamers consume more educational content on the Cyber Education Hub. Gamers were more familiar with the You-Tube style layout than non-gamers. Gamers also prefer both the You-Tube style layout and pre-requisite style layout to both the Topic-Map and KSAT layouts, with large spreads compared to non-gamers.

As mentioned above, the results of the survey showed that the 3 largest differences in Octalysis Models between groups were from the following subgroups, from greatest to least: Gamers v Non-Gamers, Junior v Senior, and Younger v Older. Below we will compare the aforementioned subgroups based on the average user engagement with different website elements and perform statistical tests.

F. Subgroup Analysis

18 different raw and derived tracking measures were compared between demographic subgroups to determine if there were any statistically significant differences in engagement with different platform elements or with the platform overall. Since demographic data and engagement data was mostly available for the classroom group, the data within the 10-week class period was used for this analysis. Since there were widely varying degrees of use between the classroom group and the operational group, and also between the 10-week program timeline and the 4-week post-program timeframe, the data used was refined to be as consistent as possible to increase the likelihood that observed engagement differences were related to demographic subgroups and their related motivation models, and not from other factors.

Only 12 of 14 participants in the classroom group experiment provided base-survey responses that allowed researchers to determine what demographic sub-groups they each belonged to. Thus, only these same 12 participants' engagement data is used for the following analysis. In addition, when motivator (Octalysis CD) values were calculated for input into the framework from Figure 22, only values calculated from subgroups of these 12 participants were utilized.

Significant Differences in Subgroup Engagement

First, in Table 18, we present the subgroup data with the four metrics mentioned in §2:

Methodology: average, standard deviation (σ), maximum value, and minimum value for each of the 18 tracking measures. Next, we calculated the percent differences in each average between groups using Equation (1). Researchers took raw individual data from the measures where resulting differences were greater than 20% and used a Welch's t-test (Welch 1947) to determine if there was a statistically significant difference between groups.

Class: Gamers	Logins	Views	Views per Login	Span	Total Time (min)	Average Time (min)	Ratio of Presence	Likes	Comments	Contributions	Contributions w/ Topic Map	Topic Map Contribution %	Topic Map Clicks	KSAT Node Navs	KSAT Links Clicked	KSAT Links per Nav	Net KSAT Activities Complete	KSAT Completes per Nav
Average	24.0	49.5	2.9	77.4	1306.9	55.3	0.3	0.1	6.1	7.5	2.8	0.6	1.5	28.4	15.3	0.6	49.1	1.4
σ	11.9	32.6	2.8	16.1	612.9	11.7	0.2	0.3	6.1	6.6	3.1	0.4	2.4	23.6	12.6	0.5	47.4	1.0
Max	48.0	119.0	8.8	113.0	2350.0	78.2	0.8	1.0	15.0	18.0	9.0	1.0	7.0	67.0	39.0	1.5	152.0	3.1
Min	10.0	19.0	0.7	61.0	514.0	42.8	0.2	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Class: Non-Gamers																		
Average	39.0	70.5	1.6	86.3	1957.8	49.6	0.4	0.3	7.0	8.3	6.8	0.8	0.3	44.8	47.5	1.2	55.0	1.2
σ	18.9	48.5	0.5	14.2	987.2	1.7	0.2	0.4	6.5	5.1	5.0	0.3	0.4	19.9	12.3	0.3	24.0	0.1
Max	69.0	144.0	2.2	105.0	3524.0	51.1	0.7	1.0	18.0	17.0	15.0	1.0	1.0	65.0	63.0	1.5	83.0	1.3
Min	18.0	22.0	1.1	72.0	842.0	46.8	0.2	0.0	1.0	4.0	2.0	0.3	0.0	22.0	30.0	0.8	26.0	1.2

Class: Junior	Logins	Views	Views per Login	Span	Total Time (min)	Average Time (min)	Ratio of Presence	Likes	Comments	Contributions	Contributions w/ Topic Map	Topic Map Contribution %	Topic Map Clicks	KSAT Node Navs	KSAT Links Clicked	KSAT Links per Nav	Net KSAT Activities Complete	KSAT Completes per Nav
Average	31.2	53.8	1.9	81.3	1657.9	54.1	0.4	0.1	7.9	9.1	4.2	0.6	1.0	34.3	23.4	0.7	54.8	1.3
σ	16.8	42.3	1.6	16.9	837.1	11.2	0.2	0.3	6.5	6.4	4.8	0.3	2.2	24.0	20.3	0.5	44.9	0.9
Max	69.0	144.0	6.3	113.0	3524.0	78.2	0.8	1.0	18.0	18.0	15.0	1.0	7.0	67.0	63.0	1.5	152.0	3.1
Min	12.0	19.0	0.7	61.0	514.0	42.8	0.2	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Class: Senior																		
Average	22.3	64.7	4.1	77.3	1121.7	51.2	0.3	0.3	2.0	3.7	3.7	1.0	1.3	32.3	33.7	1.3	40.0	1.4
σ	12.2	30.2	3.4	13.0	599.8	4.2	0.2	0.5	2.2	2.1	2.1	0.0	1.9	22.8	15.3	0.3	24.2	0.3
Max	39.0	88.0	8.8	95.0	1955.0	56.8	0.5	1.0	5.0	6.0	6.0	1.0	4.0	64.0	54.0	1.5	74.0	1.8
Min	10.0	22.0	1.2	64.0	568.0	46.8	0.2	0.0	0.0	1.0	1.0	1.0	0.0	11.0	17.0	0.8	20.0	1.2

Class: Younger	Logins	Views	Views per Login	Span	Total Time (min)	Average Time (min)	Ratio of Presence	Likes	Comments	Contributions	Contributions w/ Topic Map	Topic Map Contribution %	Topic Map Clicks	KSAT Node Navs	KSAT Links Clicked	KSAT Links per Nav	Net KSAT Activities Complete	KSAT Completes per Nav
Average	26.5	42.5	1.9	78.4	1424.6	54.5	0.4	0.0	6.6	8.1	2.9	0.5	1.0	30.5	18.5	0.6	51.3	1.3
σ	10.8	29.4	1.7	15.5	546.5	11.8	0.2	0.0	5.7	6.1	3.1	0.3	2.3	22.7	15.6	0.5	46.5	1.0
Max	48.0	119.0	6.3	113.0	2350.0	78.2	0.8	0.0	15.0	18.0	9.0	1.0	7.0	67.0	43.0	1.5	152.0	3.1
Min	12.0	19.0	0.7	61.0	514.0	42.8	0.2	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Class: Older																		
Average	34.0	84.5	3.6	84.3	1722.3	51.2	0.4	0.5	6.0	7.0	6.5	1.0	1.3	40.5	41.0	1.2	50.8	1.4
σ	22.8	43.2	3.0	16.5	1162.7	3.6	0.2	0.5	7.2	6.0	5.2	0.1	1.6	24.3	18.4	0.3	28.0	0.3
Max	69.0	144.0	8.8	105.0	3524.0	56.8	0.7	1.0	18.0	17.0	15.0	1.0	4.0	65.0	63.0	1.5	83.0	1.8
Min	10.0	22.0	1.2	64.0	568.0	46.8	0.2	0.0	0.0	1.0	1.0	0.9	0.0	11.0	17.0	0.8	20.0	1.2

Table 18 – Engagement data compared between subgroups for the 10-week classroom group dataset.

After performing the Welch's t-test, we found that 13 of 18 engagement measures showed statistically significant differences between at least 1 pair of subgroups. Two measures were found to be statistically significant between all 3 pairs of subgroups: Topic Map contribution percentage (number of content items with at least one association to the Topic Map), and KSAT links per nav (number of activity links clicked within KSAT nodes divided by the number of times a user opened KSAT nodes). In Table 19, the statistically significant differences are shown, with insignificant results (either less than 20% difference or less than 80% confidence) grayed out. Percent confidence was calculated using Equation (3).

		Logins	Views	Views per Login	Comments	Total Time (min)	Ratio of Presence	Contributions	Contributions w/ Topic Map	Topic Map Contribution %	Topic Map Clicks	KSAT Node Navs	KSAT Links Clicked	KSAT Links per Nav
Gamers v Non-Gamers	> Group	Non-Gamers	Non-Gamers	Gamers	<%20	Non-Gamers	Non-Gamers	<%20	Non-Gamers	Non-Gamers	Gamers	Non-Gamers	Non-Gamers	Non-Gamers
	% Difference	48%	35%	55%	13%	40%	31%	10%	84%	30%	143%	45%	103%	58%
	% Confidence	86%	73%	85%	N/A	83%	82%	N/A	87%	82%	89%	85%	99.5%	97%
Junior v Senior	> Group	Junior	<%20	Senior	Junior	Junior	Junior	Junior	<%20	Senior	Senior	<%20	Senior	Senior
	% Difference	33%	18%	72%	119%	39%	29%	85%	14%	58%	29%	6%	36%	59%
	% Confidence	78%	N/A	76%	97%	82%	74%	96%	N/A	99.7%	58%	N/A	76%	96%
Younger v Older	> Group	Older	Older	Older	<%20	<%20	<%20	<%20	Older	Older	Older	Older	Older	Older
	% Difference	25%	66%	61%	10%	19%	7%	15%	77%	62%	22%	28%	76%	58%
	% Confidence	69%	90%	79%	N/A	N/A	N/A	N/A	84%	99.6%	57%	72%	94%	97%

Table 19 – Statistically significant differences in engagement data between subgroups.

Subgroup Motivation Levels

Next, researchers used participant survey responses to determine Motivation Levels of each subgroup relative to the CEH platform via the framework presented in in §2.I and Figure 22. The results are presented below in Table 20. In the interest of horizontal page area, the Motivation Levels M0, M1, M2, and M3 are represented by their corresponding digit only. The Motivation Level that is higher between compared subgroups is bold and underlined. The asterisks represent that the higher level was determined by a tiebreaker.

	CD 1		CD 2		CD 3		CD 4		CD 5		CD 6		CD 7		CD 8		Count of 3s	# Higher (of 16)
	AC	GC	AC	GC	AC	GC	AC	GC	AC	GC	AC	GC	AC	GC	AC	GC	AC+GC	AC+GC
Gamers	**2**	**3**	3	2	**1***	**1**	3	**2**	2	2	**2**	2	2	2	1	0	3	6
Non-Gamers	1	1	**3***	**3**	1	0	**3***	0	**3**	**3**	0	**3**	**3**	**3**	**2**	0	6	9
Junior	**2**	**3**	**3***	3	**1***	**1***	**3**	**3**	**2**	**2**	**3**	**2**	3	2	1	0	6	10
Senior	1	1	3	**3***	1	1	1	1	1	1	1	1	**3***	**2***	**3**	0	2	5
Younger	**2**	**3**	3	3	**1***	**1***	**3***	3	2	2	**1**	1	2	2	1	0	4	6
Older	1	1	**3***	**3***	1	1	3	**3***	**3**	**3**	0	**3**	**3**	**2***	**3**	0	6	9

Table 20 – Motivation Levels of each subgroup with respect to Octalysis Core Drive motivators according to the framework shown in Figure 22.

Recommendations to Developers

In Table 21 we present a summary of the results of Table 20 to guide CEH developers in design decisions based on each Octalysis CD. Each CD is presented based on the majority Motivator Level that appears for the subgroups in the aforementioned table. M1 and M2 CDs need attention and supplementing on the platform, while M0 and M3 categorization suggests that these CDs be placed on the backburner. We take this data a step further, providing some specific game elements for each CD that can be incorporated into the website design.

We also provide general recommendations for Air Force (or military) leadership or other platform developers based on our findings. In every motivation model we looked at for each subgroup, the values of CD2, CD5, and CD7 were always positive (above neutral). On average, in 'A' questions, CD2, CD1, CD5, and CD4 motivators were the strongest among participants. Similarly, CD2, CD7, CD5, and CD6 were the strongest on average for 'G' questions. Thus, we recommend that experiences designed for Airmen incorporate, at a minimum, elements that appeal to the three CDs that are most apparent in our military members (CD2, CD5, and CD7).

Target Audience	Core Drive Priorities	Recommended Elements
Cyber Education Hub Team	Empowerment of Creativity & Feedback	Website/profile customization, Ability to create KSA Trees, Custom channels
	Scarcity & Impatience	Artificial caps on user actions, rare unlocks
	Social Influence & Relatedness	Content sharing, direct messaging, notification system
	Curiosity & Unpredictability	Random content button, random unlocks
Military Leaders and Other Platforms	Development & Accomplishment	Provide progression and give users smaller/accomplishable goals to achieve, allow users to see their progress
	Social Influence & Relatedness	Allow users to interact with and compete against peers, relate education/training to the user (career field, for example)
	Curiosity & Unpredictability	Avoid stagnation to maintain attention

Table 21 – Summary of recommendations based on motivation data.

Comparing Significant Engagement Differences and Motivation Levels

The next 3 tables show the statistically significant differences in engagement between subgroups and the associated Octalysis CDs related to each website feature. Engagement measures are aligned with the subgroup that showed more activity with the respective feature or measure. Some features primarily align with only a couple CDs while some measures are relative to the overall user motivation with the platform. For example, participants that login or spend more time on a website are likely more satisfied with it overall than someone who does not engage, but when looking at engagement with a specific feature on the website such as commenting, we can narrow this feature down to CD5: Social Influence & Relatedness.

The CDs associated with the CEH features and engagement measures are presented with each significant difference. These numbers are bold and underlined if the

Motivation Level for both AC and GC are higher than the other subgroup and only underlined if one group was higher on AC and the other was higher on GC. This only happens in one scenario in the tables below: with CD2 between the Junior and Senior subgroups. If numbers are not bold or underlined then they are lower in both AC and GC Motivation Levels compared to the other subgroup.

Gamers v Non-Gamers

Table 22 combines the results shown in Table 19 and Table 20 for the Gamers and Non-Gamers along with the CDs associated with each tracking measure. 'Views per Login' (V/L) and 'Topic Map Clicks' were both determined to be primarily related to CD3 and CD7. Users that view more content each time they login seem to be exploring more content items per session. The empowerment given to participants to view content at their will and discover the platform along with the unpredictable elements of the available content led researchers to attribute these CDs with the V/L measure. Use of the Topic Map and its related Octalysis Model has already been laid out in previous work (Tomcho et al. 2019). CD3 and CD7 were determined to be the highest motivators associated to the use of this feature. Similar to the V/L measure, users that make more clicks on the Topic Map seem to be exploring the nodes of the map.

Gamers	Views per Login	Topic Map Clicks
% Difference	55%	143%
% Confidence	85%	89%
Associated CDs	3,7	3,7

Non-Gamers	Logins	Total Time (min)	Ratio of Presence	Contributions w/ Topic Map	Topic Map Contribution %	KSAT Node Navs	KSAT Links Clicked	KSAT Links per Nav
% Difference	48%	40%	31%	84%	30%	45%	103%	58%
% Confidence	86%	83%	82%	87%	82%	85%	99.5%	97%
Associated CDs	**Overall**	**Overall**	**Overall**	2,5	2,5	2,7	2,7	7,8

Table 22 – Gamer v Non-Gamer: significant differences in engagement with associated CDs and Motivation Level insights.

The Motivation Level for Gamers was higher for CD7, but lower for CD3. This discrepancy may mean that one of the assumptions listed in the beginning of §2.I does not hold, our assignment of CDs to these engagement measures should be re-evaluated, the framework in Figure 22 may produce weak results in some cases, or we need more data to determine if the differences in these demographic subgroups remains significant when including more participants.

Non-Gamers were more involved with the CEH platform on average. This can be seen in the number of logins, total time spent on the platform, and how often users logged into the platform between their first and last sessions (RoP). As we can see in Table 20, Non-Gamers had twice as many Motivation Level 3s as Gamers and of the 16 compared motivators, 9 were higher for Non-Gamers. This suggests that the framework in Figure 22 may produce reliable results when considering all Octalysis CDs.

We also see positive results for the remaining 5 significant engagement differences between Gamers and Non-Gamers. The first 4 of these 5 are associated with CD2. Topic Map contribution measures are also associated with CD5. When contributing content items to the CEH, a user has the option to associate that item with nodes on the

Topic Map. The idea is that when users explore the Topic Map, they will be able to find your content easier, and you will consequently gain exposure to more users, get more views, and earn more likes by associating your content item with the Topic Map.

The ‘KSAT navigations’ and ‘links clicked’ measures are associated with CD2 and CD7. Use of the KSAT itself is primarily aligned with accomplishing goals, hence CD2. In addition, CD7 is present through navigations because this shows that the user is opening up more nodes to see the activities inside. Clicking links implies that users are viewing the content items associated with activities/challenges within the node to check these boxes, or find out what content item is hiding behind the hyperlink. Many activity descriptions are as simple as “Easy – view” and do not provide other details. Curious users may be more likely to click the links and view the associated content. When looking at the links clicked per node opening, researchers determined that this was mostly due to the element of curiosity (CD7); the user may have been impatient and wanted to see all the linked content ASAP. Impatience is a common trait among Black Hat core drives, and is why CD8 is also associated here, specifically for the fear of missing out (FOMO) element that may be present.

Junior v Senior

Junior	Comments	Contributions	Total Time (min)		Senior	Topic Map Contribution %	KSAT Links per Nav
% Difference	119%	85%	39%		% Difference	58%	59%
% Confidence	97%	96%	82%		% Confidence	99.7%	96%
Associated CDs	**5**	**1,2,3,4,5**	**Overall**		Associated CDs	**2,5**	**7,8**

Table 23 – Junior v Senior: significant differences in engagement with associated CDs and Motivation Level insights.

The framework also proved to be an accurate gauge of overall use ('Total Time') between the Junior and Senior subgroups when considering total number of Motivation Level 3s and total higher number of levels between subgroups. Comments are associated with CD5 and for the classroom setting, comments for CTCs could be closely linked to the number of contributions. But when looking at the averages for each group (Table 18) the Juniors had 1.2 less comments than contributions and the Seniors had 1.7 less comments than contributions.

4 of 5 of the CDs attributed to contributing content items were higher for the junior group, and 1 of 5 was split between AC and GC. Contributing content is the main driver of the crowd-sourced CEH platform. CD1 is apparent because users are sharing knowledge and increasing exposure to educational material. Users can experience CD2 and CD5 when they earn views and likes from sharing content with peers, or even competing for views with other content providers. Users can upload and receive feedback on whatever content they choose and are empowered to share their own thoughts (CD3)

or take an expert’s ideas and discuss how they may be applied in a new scenario. CD4 is present in that the content is ‘owned’ by whoever uploads the content.

The engagement measures ‘Topic Map Contribution %’ and ‘KSAT Links per Nav’ were detailed in the previous subsection. The Motivation Levels and the Associated CDs yield expected results again in this case.

Younger v Older

Younger	N/A		Older	Views	Contributions w/ TopicMap	Topic Map Contribution %	KSAT Links Clicked	KSAT Links per Nav
% Difference	N/A		% Difference	66%	77%	62%	76%	58%
% Confidence	N/A		% Confidence	90%	84%	99.6%	94%	97%
Associated CDs	N/A		Associated CDs	2,7	2,5	2,5	2,7	7,8

Table 24 – Younger v Older: significant differences in engagement with associated CDs and Motivation Level insights.

From Table 24 we can see that there were no engagement measures that were statistically significant and higher for the Younger participants. All 5 significant measures that are greater for the Older participants align with the Motivation Levels and associated CDs, showing once more the value of the framework of Figure 22. 'Views' is the only metric for which we have not discussed the associated CDs so far. CD7 is naturally occurring in viewing content; curiosity and the unknown are highly associated with this action. Most views stemmed from the KSAT activities and recently added content. The KSAT heavily features CD2, where users feel that they are developing themselves and accomplishing goals. Alongside contributing content, viewing content is a core engagement loop of the CEH. Viewing content for the purposes of learning and bettering yourself directly relates to CD2.

4. Conclusions and Future Work

These surveys should be further refined to be more effective and efficient in the future. After insight from the first iteration, many changes were made to reduce confusion from questions, improve the analysis quality of certain questions, and remove unnecessary or uninsightful questions. As more data is collected, the survey should be improved to get the most useful information possible in the most compact format to decrease participant time-cost and increase participation. Determining the levers in user responses could significantly help this effort.

In addition to shaping the survey, the results should be used to tailor development of the CEH website. This survey can also be adapted to other platforms to gather similar insight. Based on the success of using survey insights in CEH design, the CEH's model for incorporating user motivations and feedback into the platform could be expanded to additional training domains. Collecting survey data from additional domains and larger sample sizes would be highly desirable. Determining what demographic or career factors influence a person's Octalysis model could prove useful. Overall, the surveys provided useful results for CEH designers as well as insight that can be generalized for other interested parties.

The experiment and total study analysis provided useful results for CEH designers as well as insight that may be generalized for other interested parties. While this paper demonstrates advances forward in terms of our research questions, the opportunity for future progress still exists. Contributions, Future Work, and Conclusions presented below may also include deductions gained from information that is located in the Appendix.

A. Contributions

- Presented a way to quantitatively assess Octalysis Profiles of users, specific platforms, and gaps between the two via survey.
- Software engineering discussions related to this type of experiment
    - Discussion of driving requirements for the platform to facilitate the engagement experiment and the data to be stored tracking database
    - Testing the engagement tracking for the platform.
    - Development of a program that manipulates and presents tracking data in useful plots and tables.
- Discussion of what metrics were potentially noteworthy and which ended up resulting in statistically significant differences between subgroups. Discussion of what CDs correspond to some website features / tracking metrics and why
- Generalizable framework that demonstrates why differences in engagement are apparent in different subgroups based on differences in motivation relative to a specific platform.
- Suggestions presented to developers related to which CDs need the most attention during future design decisions.

B. Future Work

- Place Likert-scale responses in categorical bins and perform consensus-type analysis
- Group users based on their categorically-binned responses and perform statistical tests against engagement data.
- Determine what percent of users access additional various platforms comparable to the CEH at certain binned rates.
- Offer classroom participants different types of KSA Trees (linear, forest, branching, etc.) with same/similar nodes and see which is most popular.
- Continue to improve the Topic Map and KSATs. KSAT generation should be opened up to CEH users.

- Users reported that they were motivated to use the CEH more in the future, but were too busy. Can we find a way to better integrate the CEH into the workplace or somehow better showcase its utility?
- Users asked for a tutorial on how to use the site. Consider building a CEH Tutorial in KSA format.
- Collect data from more volunteers from more diverse groups. Continue to modify the survey based off of recommendations throughout the paper.
- Continue assessing and developing the presented Motivation Level framework.
- Run more experiments with larger groups of participants once platform has been improved based on current recommendations.
- Develop a GUI for ease-of-use of the tracking data parsing program.
- Make changes to some engagement data stored in the tracking database.
    - Associate content item type with contributions and Topic Map clicks.
    - Include node size information with Topic Map contributions and Topic Map clicks.
    - Re-test Topic Map view tracking.

C. Conclusions

Conclusions here are drawn from each subsection of data collected in presented in this paper, as well as the information in the Appendix. Answers that are direct to our research questions are presented in the next sub-section.

- We should expect user engagement at least monthly to consider the CEH a successful platform compared to other military education/training platforms.
- Common results based on the Octalysis Models include:
    - CD8 is consistently the lowest motivator followed by CD1, as reported by participants
    - CD6 and CD7 vary the most among participants.
    - Reponses about enjoyment in games/activities vary more across participants than responses about actions.

    - Black Hat CDs are more prevalent in enjoyment than actions.
- Post-experiment survey response rate was much lower. Asking people in person (base survey) increases responses.
- CEH was a useful classroom tool that can be useful operationally with improvements and a larger user-base.
    - CEH motivated users to seek more outside educational content than they would have otherwise.
    - KSAT and Topic Map are preferred by some users to the YouTube-style and prerequisite-style content layouts, warranting their existence on the platform. Users should still be given their choice due to the deviations of responses.
    - Participants used KSATs related to topics they had a vested interest in. KSATs should have a tutorial to show users how they work and their utility.
- Classroom users had a much higher level of use than operational unit participants. The students' job is to study and the CEH was a tool for this while operational participants have other jobs and the CEH was more of an opportunity to explore something new on their own time.
- Usage levels are comparable to that of other military education platforms and have a chance to grow much higher with a larger user base and more platform improvement.
- Many long sessions on CEH, educational content is being consumed.
    - Several sessions over 2 hours long.
    - 118 Sessions over 1 hour long. 105 from classroom participants, 13 from operational participants.
- Average time per session suggests that users watch a couple short videos and leave, download an article and leave, or navigate to another webpage without returning. We cannot account for the additional time spent on the latter two.
    - Most views come from the 'Recently Added' website features and the KSATs.

- There were 111 total content items contributed by participants.
    - 34 items and 14 were not class CTC related.
    - 4 items contributed from operational participants and 2 comments (about the MDC2 card game).
- The Topic Map generated 115 contributed content associations and 40 clicks but no original content views according to the database.
    - Nodes of all sizes were clicked. Nodes near 'Cyber Warfare' region were primarily contributed to.
- KSAT nodes near the beginning node were most accessed and completed.
    - All class-related activities were completed at least once.
    - Some optional activities and nodes were also completed, meaning that bonus learning occurred!
- The help page (User Guide) was utilized several times by operational participants.
- Using the Framework born in this paper, we were able to attribute differences in engagement between subgroups with differences in Motivation Levels.
    - Statistical differences in engagement and differences in motivation were seen between several demographic subgroups.
        - Motivation Levels of Non-Gamers were higher than Gamers and so was their overall engagement with the platform.
        - Motivation Levels of Junior members were higher than Senior members and so was their overall engagement with the platform.
        - Motivation Levels of Older members were higher than younger members, and while there was no significant difference in overall use, all significant differences in specific feature use were higher for Older members.

D. Research Questions and Answers

- (How) can differences in engagement with a platform be attributed to differences in motivation? Which subgroups showed the greatest engagement with the CEH and why?

    - User subgroups with statistically higher engagement also had higher Motivation Levels relative to the associated CEH website elements.
    - Non-Gamers showed the most interaction with the CEH yet only represent 29% of the population. CEH met expectations/desires of these users the best. Need to improve CEH design (Empowerment of Creativity & Feedback; Scarcity & Impatience; Social Influence & Relatedness; Curiosity & Unpredictability) in order to capture motivation and engage the other 71%.
- By implementing modern design techniques such as gamification, do target users engage more with the CEH than they do with other platforms?
    - CEH yields more user engagement than the Cyber Awareness Challenge (DISA 2018) and ADLS.
    - Research-backed design improvements can likely increase engagement in specific subgroups and overall with the CEH. By showing that meeting user desires is related to more engagement we can reasonably expect engagement to increase if we emphasize design to meet desires of more users.
- (How) do unique game elements such as a Topic Map and KSA Trees have utility in the military environment?
    - The Topic Map and KSA Trees were enjoyed by and increased educational content consumption for some users; variation was high.
    - KSA Trees are more effective when users are already interested in the topic; they can motivate these users to learn more. Additional targeted KSA Trees may be useful.

- What differences exist between motivators that cause military members to act and motivators that military members enjoy in games/activities? Which should be prioritized when designing military platforms?
    - The major differences in what motivates users to act and what users enjoy are apparent in CD1 Epic Meaning & Calling and CD8 Avoidance & Loss.
    - CD2: Development & Accomplishment, CD5: Social Influence & Relatedness, and CD7: Unpredictability & Curiosity should be prioritized when developing platforms for military users.

VIII. Final Conclusions

1. Summary and Significance

In this thesis we explored the problem domain, performed analysis of preexisting implementations, and then designed our own approach. The application of Topic Map and KSA Tree elements along with the creation of surveys and an experimental design led to insight about the implementation of gamification and how engagement can be attributed to motivation data. Although gamification is still a developing field of study it has shown effectiveness when carefully implemented. Through our experimentation and analysis, we found consistencies through frameworks and produced design decisions to ultimately improve user motivation and engagement with our target platform, the CEH.

This research advances the field of software engineering specifically with driving design requirements centered around the human user by showing that user engagement with a platform can be attributed to motivation attributed to design features of the platform. The application of this research to advance the state of cyber education and training should not be overlooked as this domain has a critical interest at getting Airmen up to speed with cyber. This research is an application of industry practice in a unique setting and has led to generalizable methods and results that can be applied in other fields and applications. More specific contributions can be seen throughout the thesis and listed in Introduction §4 on page 21. Below, we present a review of findings, this is partially inclusive of our research questions and answers presented on page 179.

2. Review of Findings

- Current education and training approaches are insufficient in several Core Drive motivators, according to Octalysis.
- Experiment participants were motivated to use the CEH in the future; said they didn't use it more during experiment because they were too busy. Improvements can be made on the CEH to likely increase user motivation with the platform.
- CEH was heavily utilized by classroom participants (daily/weekly use). Need to add more utility for operational users (monthly use).
    - Classroom participants also utilized the CEH for extracurricular activity
    - CEH also increased consumption of outside educational material.
- There are differences between what motivates users to act and what users enjoy in terms of Octalysis Core Drives.
    - Our participants showed that CD2: Development & Accomplishment was the most influential CD for both action and enjoyment. CD5 and CD7 were also strongly apparent motivators, on average.
    - KSAT was used more often than the Topic Map. When the Topic Map was utilized, it was during content contribution, not exploration. This can likely be attributed to the high levels of CD2 in the users.
    - If someone is already interested in a topic, and associated KSAT can increase their level of interest, but if they do not care, the KSAT does not increase interest.
- Different types of users engaged with the platform differently. Some users prefer the Topic Map and KSATs to You-Tube and prerequisite-style content layouts.
- Users whose Octalysis Model better aligns with the Octalysis Model of the platform engaged more with the platform.

3. Recommendations for Future Work

- Explore learning and behavior change; does an increase in engagement relate to an increase in the others? Test what types of content and what presentation styles are the best at keeping certain individuals engaged and teaching concepts based on personality traits, demographics, etc.
- Continue to improve the CEH.
    - Sustain continual assessment of the platform and also perform deeper Octalysis Levels II, III, etc.
    - Improve Topic Map and KSA Trees based on research results. Produce KSA Trees that can be utilized by more users.
    - Investigate the best approach to roll out the CEH platform to a wider user-base to maximize initial participation and long-term engagement.
    - Implement design changes on CEH; run an experiment where two groups engage w/ current version and new version of CEH; compare results; were these design recommendations effective?
    - Refine and reiterate experiments and surveys with larger and more diverse groups. provide data for the design of platforms tailored to specific communities. Consider gamifying the experiment; give the users a mission to accomplish.

.

IX. Appendix

Literature Review Supplement: Air Force Cyber Education and Octalysis Level I

This Appendix section discusses information that is relevant, but not critical to this thesis effort. Different types of Air Force cyber education and training is discussed, followed by analysis using the Octalysis Framework for each strand.

1. Air Force Cyber Education

The Air Force has several departments that deal with cyber and overall as a service currently offers several levels of cyber education to Airmen including short courses, annual training, tech school, and Masters/Ph.D. programs specific to cyber. Some examples are: the Advanced Cyber Education Course offered by the Air Force Institute of Technology, which is a four week course offered to Academy and ROTC cadets as a summer program; the mandatory annual Air Force training called the 'Cyber Awareness Challenge' (DISA 2018); tech school for Air Force cyber operators; and the studies delivered by the Center for Cyber Research (CCR) at the Air Force Institute of Technology (AFIT).

In one research presentation titled *Cyber Education & Training,* Airmen are placed into 4 classes: All Airmen, Airmen with AFSCs (Air Force Specialty Codes) most affected by cyber, cyber operators, and Air Force leadership (Dacus 2018).

All Airmen: It is difficult to refine 'All Airmen' into a specific category or categories based on their roles. Since cyber is naturally integrated into Air Force systems many different cyber tasks are performed by different types of Airmen. One important note is that all Airmen are involved in cyber due to the interconnectivity between daily operations and cyber components such as computers, electronic weapons, everyday communications, the Internet, and also secure government networks (Maybury 2015).

AFSCs Most Affected by Cyber: Airmen with AFSCs most affected by Cyber includes many specialties including the Acquisitions, Legal, Developmental Engineer, and Intelligence career fields. Airmen in this group are involved in developing and acquiring cyber technologies used by the Air Force as well as making critical decisions based on information obtained in this domain. A deeper understanding of cyber and cyberspace is critical to ensure that Airmen with these AFSCs perform their jobs and provide the Air Force with the best advantage in cyberspace operations.

Cyber Operators: The Air Force specialty description of Cyber Operators, or Cyberspace Operations Officers, is "Executes cyberspace operations and information operations functions and activities. Plans, organizes, directs and executes cyberspace and information operations such as, Defensive Cyber Operations (DCO), Offensive Cyber Operations (OCO), Department of Defense (DoD) Information Network (DoDIN) Operations and Mission Assurance for Air Force weapons systems and platforms. Such operations cover the spectrum of mission areas within the cyberspace domain" ("AFSC 17X Cyberspace Operations Officer Career Field Education and Training Plan" 2015). Cyber Operators are obviously very heavily involved in this domain. The success of cyber operations directly hinges on the knowledge and education of Air Force Cyber Operators.

Challenges

Like any major career field, or war-fighting domain especially, there are associated difficult and complex challenges. Cyber is no different, even in the Air Force.

Among these challenges are the problems of having undermanned, under-educated, and under-motivated forces to face the current cyber threats. Many of these threats stem from the fact that the struggle between international actors has spilled over from conventional warfare/politics into the cyber domain. As cyberspace grows and continues to expand, many more people from various disciplines naturally become part of this conflict and war-fighting domain whether they realize it or not. Many recognize that the dedicated cyber forces of the Air Force are likely too small to adequately defend the vast information resources in play. Moreover, the natural aptitude and passion to be an effective cyber operator are exhibited by a relatively small percentage of the population, which begs the question of how to produce more of these types of people. As stated previously, every Airman is involved in cyber in some way, and therefore better education and training may be one possible avenue to help answer this question.

The United States Air Force must find a way to educate and train people who may not have a particularly strong desire to understand. Not everyone shares the same excitement and passion to learn about cyber as those in the cyber career field may. However, even these under-motivated Airmen are still connected to the cyber domain. To better understand how to make future education and training programs more appealing to a broader range of users and more successful overall, those in charge of development first need to learn and understand why other education programs and platforms are not achieving the desired effects, and how to avoid the same pitfalls. Evaluating current and past education related to the groups of Airmen specified above will provide a good starting point. The concept of gamification will be used as the primary tool for this analysis.

2. Level I Octalysis

Now that the Octalysis Framework has been described, it can be utilized for several levels of analysis. Level I analysis based on the Octalysis Framework consists of taking a product or experience it and assessing its strengths and weaknesses with respect to motivation via the eight core drives described above (Chou 2015). This analysis will be applied to different aspects of current Air Force Cyber education and training. Performing this analysis may give insight as to why some of these practices are not inspiring more user motivation among other desired outcomes. Highlighting these areas will hopefully show future Air Force Cyber Education designers' examples of what has not worked in the past and also provide insight into alternate designs that may be more successful. It should be noted that there are over four levels of Octalysis analysis which go into more depth based on the type of user involved in the product/experience among other factors not specifically considered in level I analysis.

Education for All Airmen

As explained earlier in this paper, all Air Force Airmen (one of the four previously specified groups) are involved in cyber and some form of cyber education. One low hanging fruit involving all airmen is the annual training (for all DoD) called the 'Cyber Awareness Challenge.' Right away, one might notice the incorporation of 'challenge' into the name (Core Drive 2: Development & Accomplishment). This training is in fact an explicit 'serious game,' whether this design was intended or not. This is evident due to the animation, storyline, and other noticeable game elements. This method already interferes with the definition of gamification provided by the authors of (Werbach and Hunter 2012) since it is a 'game' context. Another aspect of this training that interferes with gamification is the fact that it is mandatory. The author of (Chou 2015) asserts several times that in order for something to be a game, it has to be a choice. It is necessary, but unfortunate, that this serious game must be mandatory. Nonetheless, the Octalysis Framework can be used to evaluate this educational training. Some of the observations below are taken from (Dacus 2018), and others are observations based on personal experience as well as conversations with Airmen about the 'Cyber Awareness Challenge.'

Epic Meaning & Calling: Right out of the gate it is apparent that the challenge tries to incorporate Core Drive 1 into the experience. The user is told that he/she is very important and they need to "maintain cybersecurity situational awareness" because there are a high number of attacks, "so make sure to do your part to secure information" (DISA 2018). This is an attempt to incorporate higher meaning into the task at hand.

Development & Accomplishment: As for Core Drive 2, there are several checkpoints in the game and also a list of levels which need to be completed. There are also points which are awarded or taken away for every good/bad decision that is made. Trophies for each level are also awarded for perfect scores. There are many different game elements applied from this core drive. However, most DoD employees may admit that the only accomplishment in the game that is important to them is finishing the training so that they will not have to do it again for a year.

Empowerment of Creativity & Feedback: The 'Cyber Awareness Challenge' is very lacking when it comes to Core Drive 3. There is only one way to win, users are dictated through each part of the game, and most importantly, they were forced to 'play' the game without making their own choice. Users are given feedback on how they do in terms of the points, but it really can't help them until the next year or unless they restart the training to immediately do better.

Ownership & Possession: This core drive is also not very apparent in this training. Users are not inclined to feel ownership over any aspect of the training. Even with a low score, users only have to achieve the bare minimum to pass and no one will know how well they did, but rather see that the training was completed. If an airman fails the challenge, they can simply retry until succeeding, only losing their time in the process.

Social Influence & Relatedness: This is yet another area where this educational training is deficient. Unlike real world cyber, you are all alone in the challenge. While there are virtual characters, you are still the only real person in the game. There is also a lack of competition and pride in one's training between players due to the previously explained pass/fail/retry system. The relatedness piece of this core drive also does not positively motivate the user. Airmen likely have negative memories of this training rather than positive, and this can increase the negative feeling toward the training even before actually logging in for the yearly session.

Scarcity & Impatience: The challenge is not actually much of a challenge after the user has completed it before. Since nothing changes, the questions and answers are always the same. Therefore, it is almost trivial to get a high score and even less challenging just to complete the training; a certificate of completion for this training is not hard to obtain. In addition, although it is required yearly, it is always accessible to complete, but there are likely very few airmen who access this training for fun (notice how 'fun' implies doing something that is a choice). The only element which touches on this core drive is that some levels are locked until prerequisites are completed.

Unpredictability & Curiosity: The very first time a user participates in this training, they will experience this core drive. However, since the training module is rarely updated and never seems to be any different, this core drive is completely lost after the initial training; everyone knows that the man in the coffee shop is going to steal the BlackBerry.

Loss & Avoidance: Unlike many other drives, Core Drive 8 is readily apparent in this training. Although not motivated by much else, users will still engage in this training because it is mandatory and they prefer to avoid punishment for not being up-to-date on their records.

Figure 35. Octalysis Model for Cyber Awareness Challenge

Summary: The overall Octalysis summary for this form of education can be seen in Figure 35. Overall, this training scores very low with only about 3 core drives significantly apparent. There is a good mix of White Hat and Black Hat core drives.

There is also a skew towards the Left-Brain core drives thanks to the points and trophies integrated into the game. In summation, this game suffers from being mandatory, but the annual training could be significantly improved to create a better experience for and to motivate the users.

AFSCs Most Related to Cyber

Air Force Specialty Codes (AFSCs) most related to cyber may include but is not limited to the Acquisitions, Legal, Developmental Engineer, and Intelligence career fields. According to (Dacus 2018), the education and training provided to this group is not standardized and there are some instances of duplicated efforts. Due to the lack of standardization, it is difficult to assess the education of this entire group through a single module/experience. Although there is no specific educational product or experience to evaluate, the Octalysis Framework can be hypothetically considered. Based on the fact that these career-fields are more involved in cyber, one could assume that the core drives Ownership & Possession as well as Development & Accomplishment would be more incorporated. Specific educational experiences could possibly neglect these core drives, however.

Cyber Operators

The Cyber Operators group is better defined and therefore, the training and education associated with this group is more specified and standardized. Before evaluating the specific training and education offered by the Air Force, it may be important to note that only 35 percent of those accessed into the Air Force cyber workforce and 11 percent of 17D (Cyberspace Operations Officers) have a cyber-related bachelor's degree (computer science, electrical engineering, or computer engineering) (Yannakogeorgos and Geis 2016). Due to the fact that Air Force Officers in the 17D career field gain their bachelor's degrees from hundreds of different universities, it would be extremely tasking to evaluate each different education experience at this level. One common element between all Cyber Operators, however, is the initial Air Force's Undergraduate Cyber Training ("AFSC 17X Cyberspace Operations Officer Career Field Education and Training Plan" 2015). The evaluations below are based off of interviews with Airmen who have personally went through the training as well as some observations in (Dacus 2018) and (Yannakogeorgos and Geis 2016).

Epic Meaning & Calling: Epic Meaning and Calling is an apparent core drive throughout Undergraduate Cyber Training. The importance of the mission and each individual's role is emphasized often. During the higher levels of this education and training, it becomes more and more evident that what is being studied has substantial meaning in relation to modern national security issues.

Development & Accomplishment: Although these Airmen may be inspired to finish the training and get a certificate of completion, this core drive could definitely be implemented better. This education does not allow Airmen to know how they are doing or check their progress on a daily basis and larger, somewhat unpredictable tests are relied upon to test knowledge of sections. Another issue is that the training records are often incomplete and do not implement a standard way of recording the specific skills each Airman acquired (Dacus 2018). Without progress checks and incomplete overall feedback on what was accomplished, motivation through this core drive is low.

Empowerment of Creativity & Feedback: Due to the methodology of this education, the core drive of Empowerment of Creativity & Feedback is virtually non-existent. With the exception of small coding exercises, which allow users some freedom in developing a solution, there is only one way to reach the 'win state.' There is not much, if any, choice over what material is learned due to the strict lesson schedule and standardization between classes.

Ownership & Possession: Although the 'players' in this experience are learning about material directly related to their careers, the Ownership & Possession core drive is lacking. Some Airmen have heard from others that at their next assignment they "will be retrained on what they actually need to know" and "a lot of what is learned will not be useful at their next assignment." Furthermore, such a specialized career field having short-term assignments can impede learning and retention rates due to the fact that the next assignment is likely outside of that operational unit (Yannakogeorgos and Geis 2016)(Dacus 2018). This may inspire a lack of investment because what is learned may not be useful for some period of time and upon return, they will be re-taught anyway.

Social Influence & Relatedness: This education is given to classes of multiple Airmen and thus, the Social Influence portion of Core Drive 5 is present. Another positive element involving this core drive is the class' group progress checks although they are somewhat uncommon. Competition in the class is not a motivator because the program is pass/fail and also the number of distinguished graduates is not set and so this achievement is mostly determined by one's own effort. Mentorship is present in the class, but based on the instructor the effects can be positive or negative. Since the instructors have a wide range of personalities and skill levels, the motivation of each student may change based on the instructor.

Scarcity & Impatience: As stated earlier, this training is mandatory for these Cyber Operators, and thus much of this core drive is lost. However, a portion of the training is Secret and exclusive. Also, the distinguished graduate achievement is a rare accomplishment that serves to motivate students to put forth effort and excel.

Unpredictability & Curiosity: There are several factors of this education/training that involve the core drive of Unpredictability & Curiosity. The negatives include: students with a bachelor's degree relating to cyber are generally required to sit through the basic knowledge portion that they already received in undergrad instead of being able to test out and make more effective use of their time (Yannakogeorgos and Geis 2016); cyber range technology is infrequently updated due to the fact that it interferes with student scheduling, which then leads to a range representing non-current functionality (Yannakogeorgos and Geis 2016); and also the slow adaptation of curriculum and textbooks in such a rapidly advancing field (Dacus 2018). The positive element of this training is that a good portion after Phase 1 is Secret level knowledge and therefore the students are motivated by the uncertainty of what they will learn.

Loss & Avoidance: The core drive Loss & Avoidance is also a strong motivator in this case. If Airmen fail or do not complete the course, they can face consequences such as losing their career or facing other repercussions.

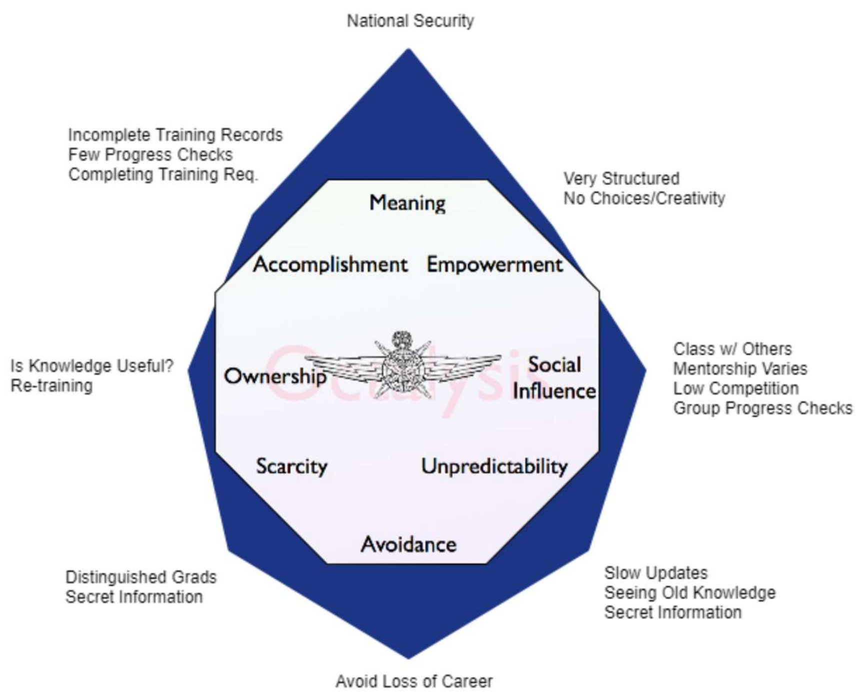


Figure 36. Octalysis Model for Undergraduate Cyber Training

Summary: The overall Octalysis summary for this form of education can be seen in Figure 36. Overall, this training scores well with about half of the core drives significantly represented. There is a good balance of White Hat and Black Hat motivation. There is also a decent balance between the Left-Brain and Right-Brain core drives. In summation, this education/training has some positive aspects, but still could be improved to create a better experience for those involved and to better motivate the students.

Leadership

Currently, the Air Force is trying to incorporate more and more cyber knowledge into its professional military education (PME) programs. Since the cyber domain is still relatively young, it has been a challenge to ensure that adequate cyber education has been provided throughout the career of those who are now leaders at the strategic level. Thus, there is a scarcity of faculty expertise for the PME courses which may be accounted for in the future through the creation of additional cyber strategy certificates or degrees. Without a standard specific course given to Air Force Leadership it is difficult to use the Octalysis Framework to assess the education and training of this group. However, one could imagine that such a course has a high focus on National Security, conveyed through a high focus on Epic Meaning & Calling. Ownership & Possession as well as Social Influence & Relatedness are likely to be present when a senior leader is learning about a topic that will strongly impact his/her forces and will also be a topic in which they must convey their competence about when addressing their troops. Many of the other factors are up to the educational designer to consider when creating such a platform or educational experience.

<u>Intermittent Research Questions and Answers</u>

The research questions below were refined from the four questions presented in (Tomcho et al. 2019). Answering some of these questions helped us refine and answer our final research questions. We present these intermittent questions and answers below in hopes that it may help or provide insight to other researchers. Additionally, some of these questions/answers could be further explored in future work.

1. "How does a Topic Map and KSA Tree affect participants' engagement with online military education platforms?"
    a. "How often do participants engage with a voluntary-use online military education platform and how do they engage with the features within the site?" <u>We saw most typical users logging into the site on a monthly basis, comparable to other military education platforms. Classroom participants typically logged in on a daily or weekly basis. Different types of users engaged with different on-site features.</u>
    b. "How do participants engage with a Topic Map within an online military education platform? <u>The Topic Map was used more for contributed content associations than as a source for viewing content. Participants wanted their content to be found via the Topic Map, but were not often utilizing it themselves.</u>
    c. "How do participants engage with a KSA Tree within an online military education platform? <u>Participants that are already interested in the overarching KSAT themes utilize the element to view content and accomplish goals. Completion rate decreased as the nodes got further from the beginning node.</u>

d. "How can user tracking data be displayed in figures and tables to provide insight about the overall use of the platform?" Figures and Tables can be displayed to give many insights into the website's engagement data. There were several findings in this paper that led to design changes, including making the KSAT navigation page clearer.

2. "How does a Topic Map and KSA Tree affect participants' motivation to use the platform?" Since all users had access to these features, we cannot be sure if there is a relationship; saved for future work.
   a. "Did participants enjoy using the CEH and are they motivated to continue using the platform in the future?" Overall, yes. On average, users 'slightly agree' that they are motivated to continue using the platform in the future.
   b. "Did participants consume more educational content on the CEH or from outside sources over the experiment timeframe?" 57% of participants (that participated in the post-experiment survey) consumed more content on the CEH while 29% consumed more from outside sources.
   c. "Why did participants use or not use the CEH website? In what settings did participants access the CEH?" Many participants saw the CEH as a place to stay updated on cyber news and relevant technology within the community, but some found the website to be too empty or clunky. The most common response about not using the platform was being 'too busy.' Perhaps why participant count was low overall. Most CEH access came from home networks.
3. "How does a Topic Map and KSA Tree affect participants' motivation to pursue more cyber education?"

a. "How did the KSA Tree effect participants' consumption of educational content both on and outside of the CEH?" The KSATs primarily increased the consumption of content on the CEH, mostly with Non-Gamers.

b. "How did the Topic Map effect participants' consumption of educational content both on and outside of the CEH?" The Topic Map primarily increased the consumption of content on the CEH. Again, mostly with Non-Gamers, but not as much as the KSATs.

c. "Which content-presentation styles do users prefer among the Topic Map, KSA Tree, YouTube-style, and prerequisite-style layouts?" Different users prefer different layouts, but some users did prefer the Topic Map and KSAT over the other two layouts.

d. "How did different KSA Trees effect participants' interest in their respective topic areas?" KSATs increased participants' interests in topics, but only if they were already interested in that topic beforehand.

e. "Would users like to see Topic Maps and/or KSA Trees implemented in other applications?" Yes, and they would like to see them in different areas with different people in charge of the design/implementation.

4. "What does the Octalysis model of the participants look like? Are there significant variances between career fields, age groups, etc.?"

a. "What Core Drives motivate users to act? What does the resulting Octalysis model tell us?" Users are motivated to act by all 8 CDs. CD2 was the strongest and CD6 was the weakest.

b. "What Core Drives do users enjoy in games/activities? What does the resulting Octalysis model tell us?" Users enjoy all CDs except CD1 and CD8. The strongest are CD2 and CD7.

c. “What does the Octalysis Model of the CEH look like according to participants’ responses?" Users reported that many CDs on the CEH were not present or were at least lacking compared to what motivates them and what they enjoy. Specific suggestions for improvement are presented in the main document.

d. “What are the demographics of the participants? Can the group be split into demographic sub-groups to compare the respective Octalysis Models against each other to see differences in motivation/enjoyment? Participants were all military affiliated, primarily with the Air Force. The two demographic subgroups that showed the largest differences were Gamers and Non-Gamers. The next greatest difference was seen in Junior and Senior members.

e. “Do self-identified gamers behave differently than non-gamers? Are there differences between the Octalysis Models of these groups? Do these groups feel differently about the CEH website?” Yes, especially with regard to gaming activity. Gamers had higher response values for all CDs except CD1 and CD8. Overall, the groups felt similarly about the CEH website. When looking at specific elements, Non-Gamers enjoyed the Topic Map and KSAT more than Gamers.

f. “Do different demographic groups engage differently with the online military education platform?” Yes, there were significant differences found between the use of the platform by demographic subgroups. Non-Gamers used the site more than Gamers; this gap was the largest between all tested groups.

g. “What do the survey results and engagement data suggest about motivation and engagement when combined?” Using the framework presented in this paper, researchers were able to show that differences in engagement data could, in most cases, be directly attributed to differences in motivation between groups.

Unpublished Scholarly Article: Complex Optimization Algorithm Design Project: Minimal Steiner Tree in Graphs Variant

2Lt Landon Tomcho
landon.tomcho@afit.edu

Gary B. Lamont, Ph.D.
Gary.Lamont@afit.edu

Air Force Institute of Technology
2950 Hobson Way, Wright-Patterson AFB, OH 45433

Abstract

The Steiner Tree in graphs (STG) is a well-known NP-Hard optimization problem ("Steiner Tree Problem" 2018)(Chlebík and Chlebíková 2002). This problem relates directly to my thesis research effort and can be used to implement an efficient algorithm on the Cyber Education Hub (CEH) Website (Eddins 2018). In this paper, different approaches are taken to solve this problem and then compared. The general approaches fall under the categories of deterministic search, stochastic search, and local search. The techniques used to solve this problem include an explicit problem domain specification, selection and integration of appropriate search elements, modification and adaptation of algorithm templates, and implementation in code for testing and evaluation. Three different algorithms are presented and compared including a deterministic depth-first search (DFS), a stochastic search (Genetic Algorithm), and a local search (Tabu). For our examples, the performance from best to worst is DFS, Tabu, then Genetic Algorithm.

Keywords

Steiner Tree, optimization, algorithm, genetic algorithm, Tabu search, deterministic search

1. Introduction / Problem Selection

My thesis research deals with the gamification of the Cyber Education Hub (CEH). The CEH is an online-based 21st century education platform where users can voluntarily learn about cyber and enhance their skills to become more competent and more competitive in their careers (Reith et al. 2018)(Tomcho and Reith 2019)(Eddins 2018). Being voluntary means that the platform must be easy and also 'fun' to use. This is where gamification comes into play. One aspect of gamification is taking common game elements and applying them to your platform to create a more engaging and motivating experience for the user. The CEH is a platform that holds user-uploaded content to ensure that knowledge remains up-to-date and practical, but an issue is that this content can be hard to navigate for cyber novices and experts alike. If a user wanted direction and also autonomy in their educational experience, a topic map or skill tree is the perfect game element to apply (Tomcho and Reith 2019).

Skill trees and maps are used in games to show a user's progress, show where the user can eventually go, and provide different paths to get there although the route is ultimately up to the user. On the CEH the tree is a graph of nodes and edges which represent topics and connections, respectively. Content that is tagged with a certain topic is viewable when a user is within the node. One functionality we would like to have on the CEH is to allow the user to select certain topics they want to learn about and then based on what the user already knows, we can show them the most efficient route to build off of what they know to learn what they need and ultimately what they want.

This problem is a variant of the minimal Steiner Tree in graphs (MSTG) problem ("Steiner Tree Problem" 2018)(Prömel and Steger 2002). One difference is that the nodes, and not the edges, have associated weights. An example Steiner Tree on a basic cyber skill tree can be seen in Figure 37 below. The green nodes represent the goal nodes, what the user wants to know. The golden nodes represent the Steiner nodes, the nodes required to make a connected subgraph, they represent the knowledge the user needs as a foundation. The key optimization is to create a tree with minimal weight so that the user can learn what is desired at minimum cost (most likely time). Note that the weights associated with each vertex is not present in the figure below. The weights could be the total length of all content associated with that node/topic, the average length of time required until a user can pass a quiz at 75% competency, etc. This is not as important now, if we can efficiently solve the Steiner Tree problem detailed in ("Steiner Tree Problem" 2018), we can be concerned with the weights of this graph later.

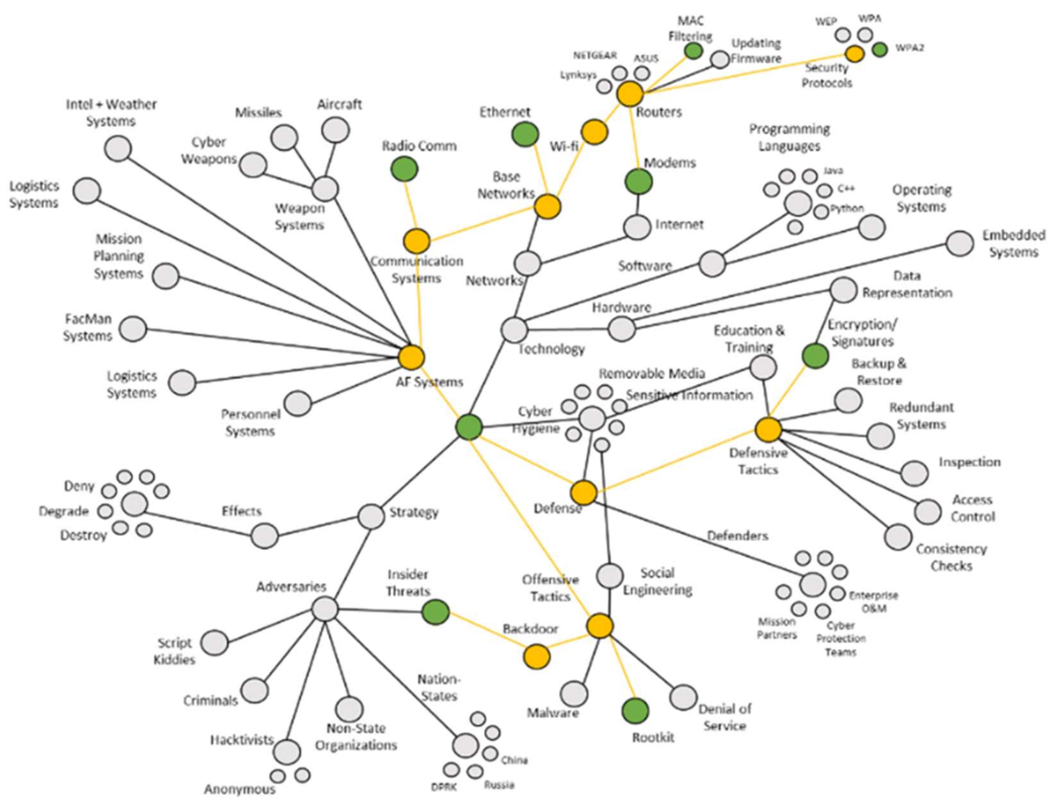


Figure 37. An example Steiner Tree on a basic Cyber Topic Map.

There have been many papers written on the Steiner Tree problem and also on the Steiner Tree problem in Graphs. Just a few are (Klein and Ravi 1995), (Mehlhorn 1988), and (Zelikovsky 1993).

2. Problem Domain

A. Problem Domain Complexity

The MSTG problem reduces to a minimum spanning tree (MST) problem if every node in the graph is selected as a goal node. However, in our scenario, the specific edges do not matter as long as the subgraph is connected, so in this case it is trivial. Also, if only two goal nodes are selected, we now have the shortest path problem (SPP). Both the MST and SPP are polynomial-time solvable ("Minimum Spanning Tree" 2018)("Shortest Path Problem" 2018). However, in every other scenario, the formulated MSTG is NP-Hard.

Since we are only considering whether each node is included in the Steiner Tree or not, we can represent the solution as a bit string of length n where n is the total number of nodes in the graph. Thus, we can see that there are 2^n total possible solutions. However, many of these bit strings are not be feasible solutions because they either don't include the goal nodes or they are not connected subgraphs. Nonetheless, the complexity of the problem domain search space is exponential (2^n).

According to (Klein and Ravi 1995) we cannot expect to obtain an approximation algorithm that achieves a performance ratio better than logarithmic. (Klein and Ravi 1995) proves that no polynomial-time approximation algorithm for set cover achieves an approximation factor smaller than ¼ unless deterministic time $n^{\text{polylog}\, n}$ contains NP.

B. Mathematical Formulation

Given an undirected Graph with weighted vertices $G = (V, E, W)$ and a set of Goal vertices S; we have $V = \{v_1, v_2, \dots, v_n\}$ is the set of vertices in G (v_0 is the start node and is in every S), $E \subseteq \left\{\{v_i, v_j\} | v_i \in V, v_j \in V, v_i \neq v_j\right\}$ is the set of edges in G, w is the set weights associated to each vertex in V such that $[\forall i_{0\ldots n} \mid v_i \rightarrow w_i]$, and $S \subseteq V$. Note that most Steiner tree problems have weighted edges not weighted vertices.

One constraint is that the solution must be a connected subgraph (tree). This means that for all vertices in the solution $\{u_1, u_2, \ldots, u_p\}$, for all values of k, $1 \leq k \leq p$, there is at least one edge $\{u_m . u_n\}$ between each vertex and another vertex in the solution and also there is some path from each vertex to any other vertex in the solution.

The solution tree *T* must also include all the goal vertices *S*. In mathematical terms: $S \subseteq T$.

The goal of our version of the Steiner Tree optimization problem is to find a tree T of G that spans S with minimal total cost/weight. If the vertices in the solution are $\{u_1, u_2, \ldots, u_p\}$ the objective is Min$\{\Sigma \mathrm{w}_k\}$ $(1 \leq k \leq p) \mid u_{1\ldots p} \rightarrow w_k$.

C. Algorithm Domain Selection & Specification

For this project we are required to employ at least one deterministic search technique which is guaranteed to give an optimal solution. Some examples of techniques in this category are global depth-first-search with back-tracking (global DFS_BT), global breadth-first-search (global BFS), Z*, A*, and dynamic programming. We must also choose and employ at least one biology-inspired stochastic search technique such as genetic algorithms (GA), ant colony optimization (ACO), particle swarms, etc. Finally, we must also employ a local search technique such as simulated annealing (SA) or Tabu search in the solution space for the bio-inspired stochastic search.

3. Deterministic Search

Before creating an efficient algorithm, we wanted to map the problem landscape and also find the global optimum values of each of the test graphs. This was accomplished with a brute force search that looked at every possible solution and found its respective fitness value (weight). The toy problem graph we used was a subgraph of a sample graph from an online graph database. Node weights were not present in the sample graph and were assigned based on what one might expect out of a graph on the CEH website. Nodes with lower weight are generally leaf nodes from larger topic area nodes. Our toy graph can be seen below in Figure 38.

There were five different sets of special nodes chosen and tested. The weights and edges remained constant. While we were mainly focused on testing the general efficacy of certain algorithm types, the algorithms we present should be tested on more graph variants in the future. The graph below has goal/special nodes shown in yellow. The numbers are in blue and the weights are in black next to each node. The special nodes for each of the five graphs are represented below.

```
SpecialNodes = "00101101101000100000" #A
SpecialNodes = "11100000000000000000" #B
SpecialNodes = "10010001010000000000" #C
SpecialNodes = "00000000101000000110" #D
SpecialNodes = "11100000000000000111" #E
```

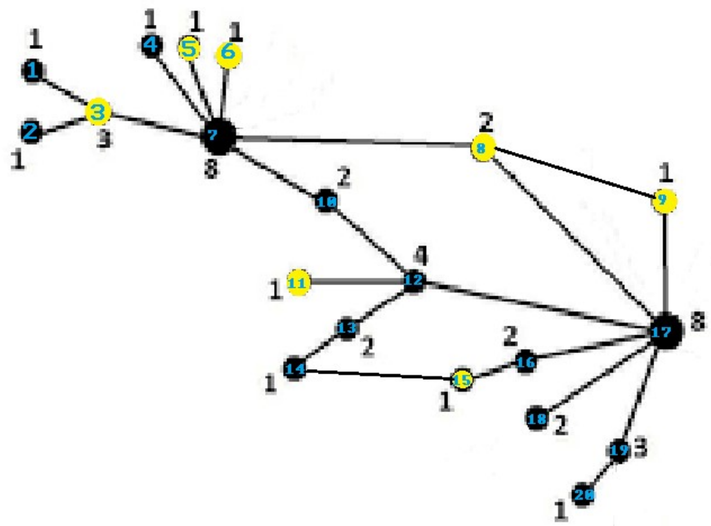


Figure 38. 1 of 5 sample Steiner Trees used for testing.

The following figures show the landscape of each of the five toy graph problems. The horizontal axis corresponds to the decimal number relating to the binary representation of each specific solution. The vertical axis represents the weight of each solution. Lower weights are better and solutions that are infeasible are given the maximum weight of 46.

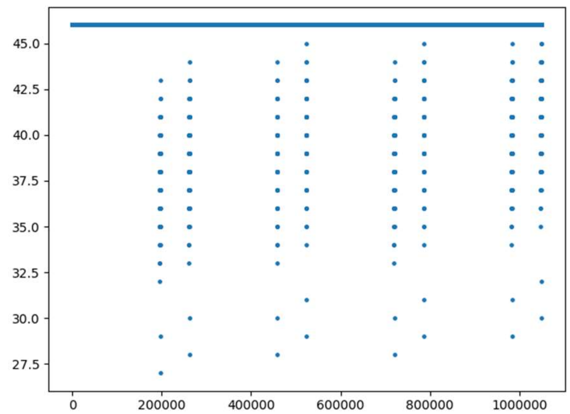


Figure 39. Landscape for Graph A

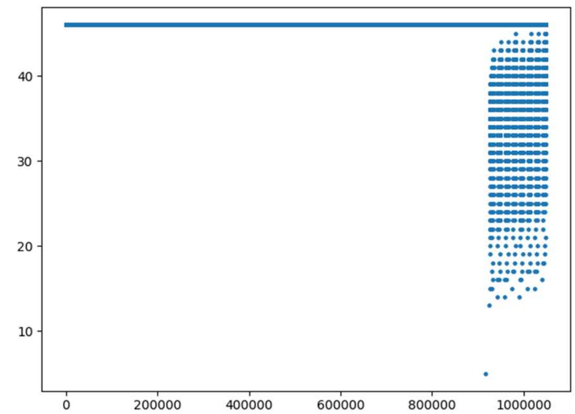


Figure 40. Landscape for Graph B

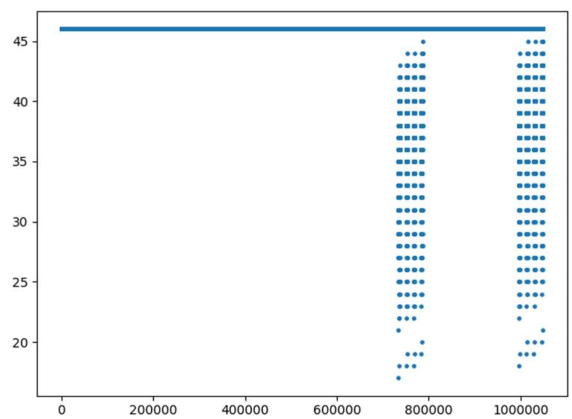


Figure 41. Landscape for Graph C

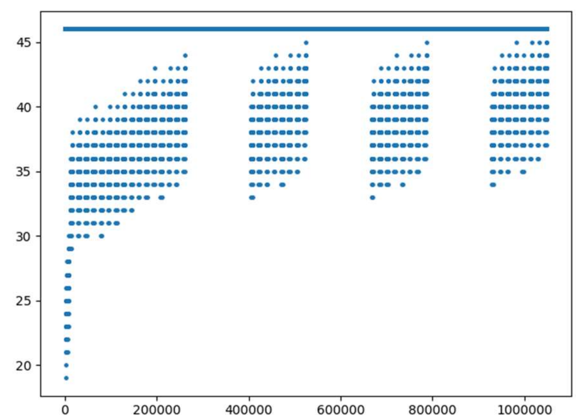


Figure 42. Landscape for Graph D

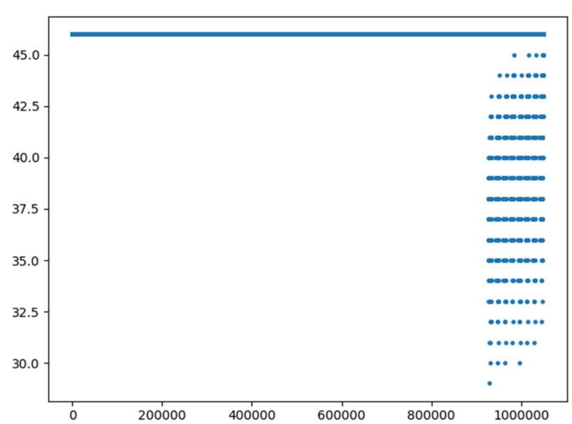


Figure 43. Landscape for Graph E

A. Top-Down Algorithm Design

Problem Domain Requirements Specification

- domains, D
  - input $D_i(G,S),\ graph\ G = (V,E,W)$, set of vertices V, set of edges E, set of weights associated with each vertex W, set of goal nodes S
  - output D_o – Steiner tree T with additive weight – a set of connected vertices including at least the goal nodes.
  - partial solution D_p- partial set of vertices with current partial solution weight z
- $I(x)$; input conditions on the domain satisfied
- $O(x,z)$; output conditions on output/input domain satisfied, i.e. a feasible/optimal solution with respect to the input domain
  - Minimize additive weight subject to the following:
  - Min{Σw_k} $(1 \le k \le p) \mid u_{1\ldots p} \to w_k$
  - Where $u_{1\ldots p}$ are the vertices of T.

PD/AD Integration Specification

Basic search constructs for gs-dfs/bt

- Initial set of candidates
  - Start with all nodes includes in solution $T = G$
  - Remove nodes / trim tree until we arrive at optimum solution
- Next-state-generator
  - $v \in T$
  - $I(x)$
- Selection
  - Remove some vertex
- Feasibility
  - $v \notin S$ – cannot remove a goal/special node
- Solution
  - T must remain connected after removal of node
  - $S \subseteq T$ – all goal/special nodes must remain
- Objective
  - Minimize additive weight subject to the following:
  - Min{Σw_k} $(1 \le k \le p) \mid u_{1\ldots p} \to w_k$
  - Where $u_{1\ldots p}$ are the vertices of T
- Delay Termination / Backtracking Loop
  - Find all minimal/minimum solutions within loop

B. Algorithm Design Specifications

Incorporating some heuristics:

1. Remove nodes with higher weight first
2. Don't allow removal of nodes if they disconnect T

- Initial set of candidates
  - Start with all nodes included in solution $T = G$
  - Remove nodes / trim tree until we arrive at optimum solution
- Next-state-generator
  - $v \in T$
  - $I(x)$
- Selection
  - Max$(w_i \to v_i)$
  - Remove candidate vertex with max weight
- Feasibility
  - $v \notin S$ – cannot remove a goal/special node
  - T must remain connected after removal of node
- Solution
  - If feasibility conditions are met, the solution is feasible
- Objective
  - Minimize additive weight subject to the following:
  - Min{Σw_k} $(1 \le k \le p) \mid u_{1\ldots p} \to w_k$
  - Where $u_{1\ldots p}$ are the vertices of T
- Delay Termination / Backtracking Loop
  - Find all minimal/minimum solutions within loop

C. Intermediate Algorithm Designs

Incorporating some more heuristics: (similar to those mentioned in (Koch and Martin 1970)).

Heuristic 1: Trim tree
Before starting search, remove all non-special nodes with a degree of one and adjust the adjacency matrix accordingly.

Heuristic 2: Supplement special nodes
If a special node has a degree of one, the vertex it is connected to is in the solution and essentially becomes a special node that cannot be removed.

Since these heuristics are applied before the search loop, below search constructs remain the same as previous subsection.

- Initial set of candidates
  - Start with all nodes includes in solution $T = G$
  - Remove nodes / trim tree until we arrive at optimum solution
- Next-state-generator
  - $v \in T$
  - $I(x)$
- Selection
  - $\text{Max}(w_i \rightarrow v_i)$
  - Remove candidate vertex with max weight
- Feasibility
  - $v \notin S$ – cannot remove a goal/special node
  - T must remain connected after removal of node
- Solution
  - If feasibility conditions are met, the solution is feasible
- Objective
  - Minimize additive weight subject to the following:
  - $\text{Min}\{\Sigma w_k\}\ (1 \leq k \leq p) \mid u_{1 \ldots p} \rightarrow w_k$
  - Where $u_{1 \ldots p}$ are the vertices of T
- Delay Termination / Backtracking Loop

Find all minimal/minimum solutions within loop

D. Algorithm Pseudo-Code

```
Initialization    \\ set up initial variables
Trim              \\ trim non-special nodes with degree of 1
Supplement        \\ add vertices connected to special nodes
                  w/ degree of 1 to list of special nodes
Loop
   Snip   \\ next-state, selection
Check     \\ feasibility, solution
   Weigh          \\ objective
   Backtrack \\ if necessary
```

E. Algorithm Implementation

We chose to use Python for the implementation of the DFS for several reasons. Foremost because we had already coded the GA and Tabu helper functions in Python, but also because of the ease of coding and understanding python as well as the fact that it is fast and has many available libraries. A shortened version of the main loop and the DFS( ) function is below to show the mapping between the search elements, pseudo code, and implemented code. Note that recursion and a duplicate list is used to accomplish the DFS loop with backtracking.

Main:

```
duplicateList = []                        Initialization
initial_solution=trim("11111111111111111111") Trim
newSpecialNodes = supplement(SpecialNodes)
                                          Supplement
DFS(initial_solution, newSpecialNodes)        Loop
```

```
def DFS(dna, special):
    fit = fitness(dna,special)  Weigh \\ objective

    if (dna not in duplicateList):   \\ selection
        duplicateList.append(dna)
        if (fit != TotalWeight) or (dna == "11111111111111111111"):
                                        Check \\ solution
            for i in range(DNA_SIZE): \\ next-state
                newSol = ""
                if (dna[i] == '1') and (special[i]!='1'):
                                        \\ feasibility
                    newSol = dna[:i]+'0'+dna[i+1:]
                                               Snip
                    DFS(newSol, special)  Backtrack
```

F. Evaluation Experimentation

After incorporating heuristics, the search was able to find the same optimum solutions much faster than the brute force solver. By drastically reducing the search space we were able to avoid wasting computation by unnecessarily checking certain solutions. Generally, it takes longer to get optimum solutions with fewer nodes as there are more possible solutions to check in the search tree. If the optimum solution has many nodes, we do not need to trim the tree as much and don't need to explore as deep. These results can be seen in Figure 44.

Graph	Min Weight	DFS (s)	Brute (s)
A	27	0.005	7.727
B	5	0.050	14.379
C	17	0.016	10.749
D	19	0.025	10.747
E	29	0.044	8.651

Figure 44. Results of the DFS deterministic search

4. Bio-inspired Stochastic Search

A. Top-Down Algorithm Design

In the 1980s the principles of evolution inspired computer scientists to design what are called evolutionary algorithms. Among the main sub-fields of evolutionary algorithms are evolution strategies, genetic programming, and genetic algorithms (Talbi 2009). These evolutionary algorithms are stochastic metaheuristics that have been applied to many problems. They are population-based algorithms that are based on the concept of competition, just like in Darwin's theory of evolution. Evolutionary algorithms are generally applied to problems with a large search space where it would take a very long time to find the exact optimal solution. Instead, evolutionary algorithms use competition and breeding with the goal of getting a good solution in a more feasible amount of time.

The evolutionary algorithm template from Talbi can be seen below in Figure 45. The initial population is generally randomly generated. Members of the population are solutions to the problem at hand; each has its own properties and associated fitness. The population size and number of generations are parameters to be set by the programmer to determine how long the search runs before returning the best-found solution.

```
Generate(P(0)) ; /* Initial population */
t = 0 ;
While not Termination_Criterion(P(t)) Do
    Evaluate(P(t)) ;
    P'(t)      = Selection(P(t)) ;
    P'(t)      = Reproduction(P'(t)); Evaluate(P'(t)) ;
    P(t + 1)   = Replace(P(t), P'(t)) ;
    t = t + 1 ;
End While
Output Best individual or best population found.
```

Figure 45. A template for evolutionary algorithms (Talbi 2009)

While the search is not terminated the following cycle repeats. Members of the population are selected to breed (or sometimes just to survive to the next generation). Selection methods vary, but one common way is the Roulette-strategy where each member has a chance to get selected that is proportionate to its fitness value. Like in nature, more fit members are more likely to survive and breed. Next, reproduction occurs. In this phase two or more members are combined in some manner to produce offspring. Most of the time, the offspring gains characteristics from each of the parents. The offspring are then assigned their own fitness value. Sometimes offspring are not feasible solutions, sometimes they have worse fitness than their parents, but sometimes they are better. Even a generally unfit parent can generate a fit offspring because they carried a good characteristic. This is why it is generally not a good idea to only select the most fit members for breeding. Doing this may result in getting stuck at a local optimal solution.

Another technique commonly used in genetic algorithms (and is also nature-inspired) is mutation. To help diversify the population and cover more area of the search space, members have a generally low chance of having a characteristic become mutated. This allows for some characteristics that were not in the initial population to be introduced to potentially help find better solutions.

After reproduction and mutation comes replacement. There are many different replacement strategies to determine the next generation. For example, only offspring can survive to the next generation, or the top 10% of the last generation can survive and the rest of the positions can be filled with new offspring. This loop continues until termination criteria is met. There are many parameters that can be tuned and several different strategies for each phase of the algorithm. Different parameters and strategies work better for different problems and different landscapes. There is no one-size-fits-all solution, just like there is no free lunch (NFL) ("No Free Lunch in

Search and Optimization" 2018). We decided to use a genetic algorithm approach for my stochastic search for the MSTG problem.

Problem Domain Requirements Specification

- domains, D
    - input $D_i(G,S),\ graph\ G = (V,E,W)$, set of vertices V, set of edges E, set of weights associated with each vertex W, set of goal nodes S
    - output D_o – Steiner tree T with additive weight – a set of connected vertices including at least the goal nodes.
- $I(x)$; input conditions on the domain satisfied
- $O(x,z)$; output conditions on output/input domain satisfied, i.e. a feasible/optimal solution with respect to the input domain.
    - Minimize additive weight subject to the following:
    - $\text{Min}\{\Sigma \text{w}_k\}\ (1 \leq k \leq p) \mid u_{1\ldots p} \rightarrow w_k$
    - Where $u_{1\ldots p}$ are the vertices of T.

PD/AD Integration Specification

- Next-state-generator
    - The next-state-generation for a Genetic Algorithm comes from crossover and generating new solutions to fill the population of the next generation.
    - Next set of candidates are binary strings which have genes/characteristics from parents of last generation.
- Selection
    - Selection in the GA comes in the form of choosing which solutions will be chosen to breed and make new solutions.
    - A Roulette selection strategy is used where the chance of a solution being selected to breed is proportional to its relative fitness in the population.
- Feasibility
    - All binary strings are feasible. If we restrict population members only to actual solutions, we may end up using a random number generator to solve the problem instead of allowing the GA to perform.
- Solution
    - If a binary string is an actual feasible solution, its associated weight will be lower than the maximum weight. A penalty will be imposed on binary strings that do not meet constraints.
- Objective

- The fitness of a solution is inverse to the total weight of all the nodes in T. The lower the weight, the better the solution. Binary strings that do not meet constraints will be assigned a maximum weight, equal to the total weight of all nodes in the graph.

B. Algorithm Design Specification

With the evolutionary algorithm template from Figure 45 and genetic algorithm techniques in mind we can begin to specify the specific techniques to be used in our application to solve the MSTG problem for this project.

Initial Population

As mentioned previously, genetic algorithms are generally used to find good solutions to complex problems. While initial population members are sometimes required to be feasible solutions, we decided to allow them to be infeasible solutions. To be more specific, we allowed the bit strings which represent which nodes are in the solution graph to not include the goal nodes and to also be disconnected subgraphs. Obviously, we want our final answer to meet both of these criteria, but if we required every initial member to be a feasible solution, it could take a long time considering the search space. So, to ensure that our final solution was feasible but also to save time we simply gave all infeasible solutions the worst possible fitness value. The population size in (Jones and Harris, n.d.) was 75 and in (Kapsalis, Rayward-Smith, and Smith 1993) was 10. We experiment with different population sizes, which will be discussed later.

Fitness

A majority of the time spent in GAs is determining the fitness value of a member. For my project the fitness was the total weight of all the nodes in the solution sub graph. If a solution was unfeasible it was assigned the same weight as if it included every single node in the graph. Checking if a member included all goal nodes was trivial but checking if the subgraphs was connected was a more expensive operation. This operation was performed by an iterative search through all nodes and connections using a master matrix of all nodes and connections.

Selection

The Roulette Strategy was used in this implementation. Although you run the risk of selecting the same solution every time (due to random chance), this strategy was found to be the safest by (Kapsalis, Rayward-Smith, and Smith 1993) although it is $O(n \log n)$ time versus $O(n)$ for Stochastic Uniform Selection (SUS) (Talbi 2009). In my opinion, the implementation in code of this strategy was easier than SUS and tournament selection. However, these techniques are also commonly used.

Reproduction / Crossover

After the candidates are selected to reproduce crossover must occur to create new offspring. Since we are dealing with bit strings, we are guaranteed to produce children that satisfy condition (1) based off of the Hamming distance if we use uniform or *n*-point crossover methods. Although we think uniform crossover is best, it takes some extra computing time because you must generate a random number for each bit every time we breed, which is very often. (Jones and Harris, n.d.)uses a 2-point crossover and (Kapsalis, Rayward-Smith, and Smith 1993) uses a uniform crossover. Booth of these papers found best results with a crossover probability between 0.9 and 0.95. The 1-point crossover was easy to implement and was very fast so we implemented this method.

(1) $$\text{Max}\{d(p1, o), d(p2, o)\} \leq d(p1, p2), \forall o \in O(p1, p2, Ox)$$

Mutation

Although we discussed how it takes a long time for uniform crossover, there is not a quick and effective alternative for mutation. Therefore, we use a uniform random mutation on each offspring to provide a source of diversity in the population.

Replacement

There are different replacement strategies but most fall into two categories: generational replacement (replace-all) or steady state replacement (replace-some). It seems intuitive that keeping some of the best of each generation around may lead to finding better solutions faster, but (Kapsalis, Rayward-Smith, and Smith 1993) found that the replace-all strategy worked best (for their Steiner Tree example, at least). So, we followed suit in my implementation.

C. Intermediate Algorithm Designs

Several changes were made to the genetic algorithm for this project. One of the major changes was revising and improving the algorithm which checked if a subgraph was complete to make it more accurate and efficient. Also, we made sure to remember the best-found solution instead of outputting the best solution from the last generation, because sometimes the mutation and crossover would cause all feasible solutions to be lost. Another major change was requiring all initial members to have the goal nodes. These changes (and some smaller ones) along with tuning parameters allowed the GA to be much more effective at solving the MSTG problem.

D. Algorithm Pseudo-Code

```
Generate(P(0)) ;  //Random Initial population
t = 0;
While not Termination Criterion(P(t)) Do
Evaluate(P(t));                  //objective minimize total weight
RememberBest(P(t));
P'(t) = Selection(P(t));         //selection use roulette strategy
P'(t) = Reproduction(P'(t)); Mutate(P'(t)); //feasibility
P(t + 1) = Replace(P(t), P'(t));          // replace-all
t = t + 1;
End While
Output Best individual found. // solution
```

E. Algorithm Implementation

We chose to use Python for the implementation of the GA for several reasons. Foremost because we found a good template on GitHub for GAs, but also because of the ease of coding and understanding python as well as the fact that it is fast enough and has many available libraries. A shortened version of the main loop is below to show the mapping between the search elements, pseudo code, and implemented code.

```
for POP_SIZE in [10,30,100,250]:
    for GENERATIONS in [10,100,500,1500,2500]:
        # Generate(P(0))
        population = random_population()
        for generation in range(GENERATIONS):
            weighted_population = []
                # Evaluate(P(t)); //objective, feasibility and solution
                fitness_val = fitness(individual)
                # RememberBest(P(t));
                if fitness_val <= MIN_FITNESS:
                    FITTEST_STRING = individual
                    MIN_FITNESS = fitness_val
                if fitness_val == 0:
                    pair = (individual, 1.0)
                else:
                    pair = (individual, 1.0 / fitness_val)
                weighted_population.append(pair)
            population = []
            # P'(t) = Selection(P(t));  //selection
            for _ in range(POP_SIZE // 2):
                ind1 = weighted_choice(weighted_population)
                ind2 = weighted_choice(weighted_population)
                # P'(t) = Reproduction(P'(t));
                ind1, ind2 = crossover(ind1, ind2)
                # Mutate(P'(t)); P(t + 1) = Replace(P(t), P'(t));
                population.append(mutate(ind1))
                population.append(mutate(ind2)
```

F. Evaluation Experiments between GA versions

As you can see in the code above, the GA was run with different population sizes and different generation limitations. The total time of each run as well as the best-found solution weight was recorded. The algorithm was implemented on the graph shown in Figure 38 with five different sets of special nodes chosen.

The charts in Figure 46 and Figure 47 show the results of the first and latest GA implementations, respectfully. The orange bars represent the generations, the blue line represents population size, and the grey area represents the weight for the best-found solution. Time was proportionate to the total members (pop size * generations). The latest GA implementation takes slightly longer, but it is much more effective than the first. We can see the difference in effectiveness in Figure 46 and Figure 47 and also in the differences in Figure 49 and Figure 49.

Figure 50 shows that the latest implementation can take a little longer (mostly because of the connectedness check). Notice in Figure 50 however, that the first implementation, GA0, often did not find a feasible solution (where grey dots are at a max weight of 46). The latest implementation is much more effective at finding feasible (and better) solutions than the first implementation because of the additional heuristics and refinement.

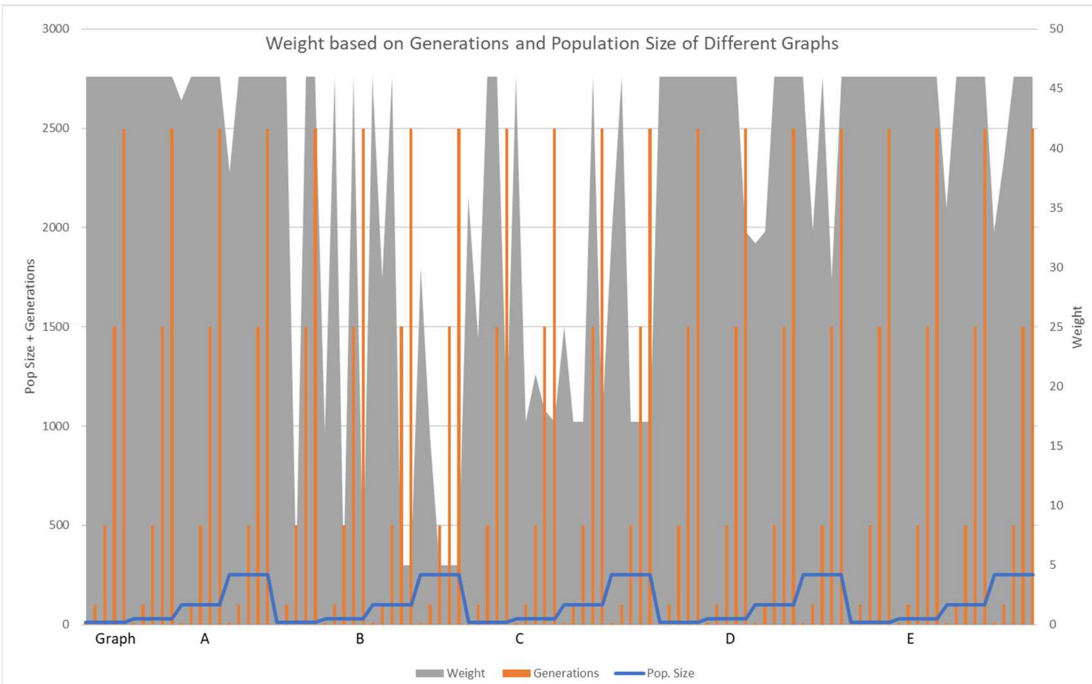


Figure 46. First GA implementation results

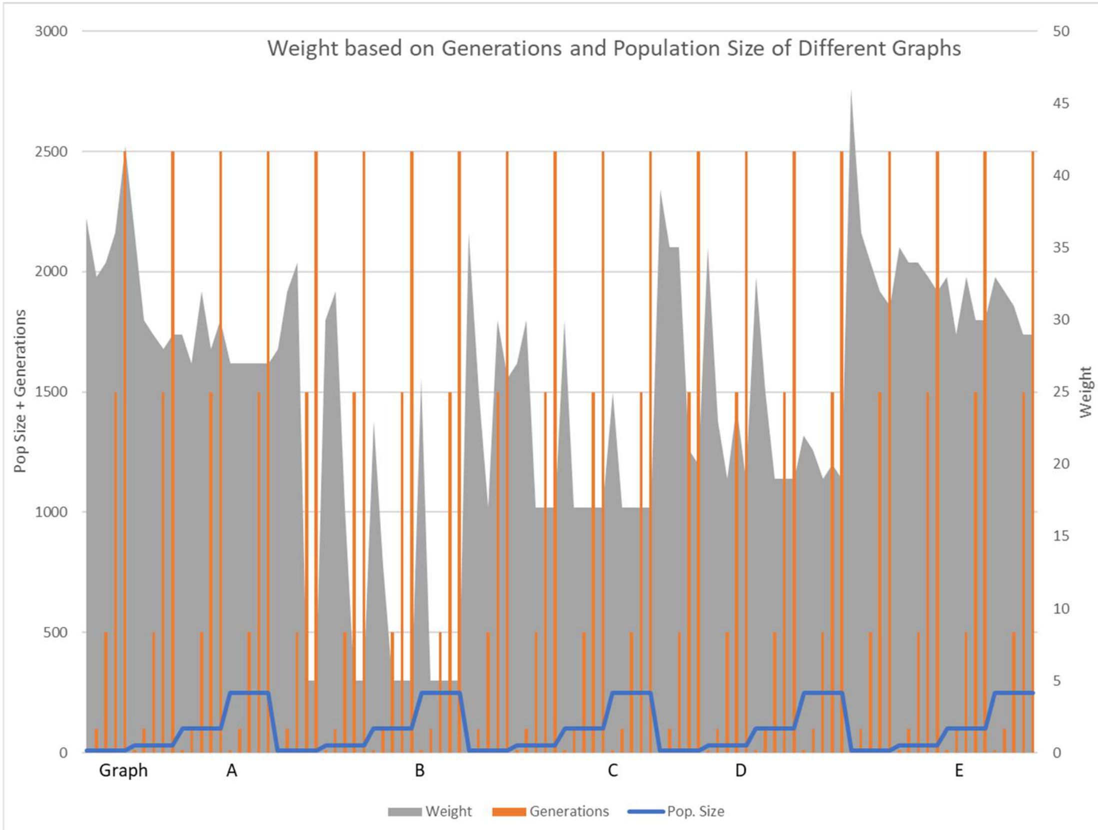


Figure 47. Latest GA implementation results

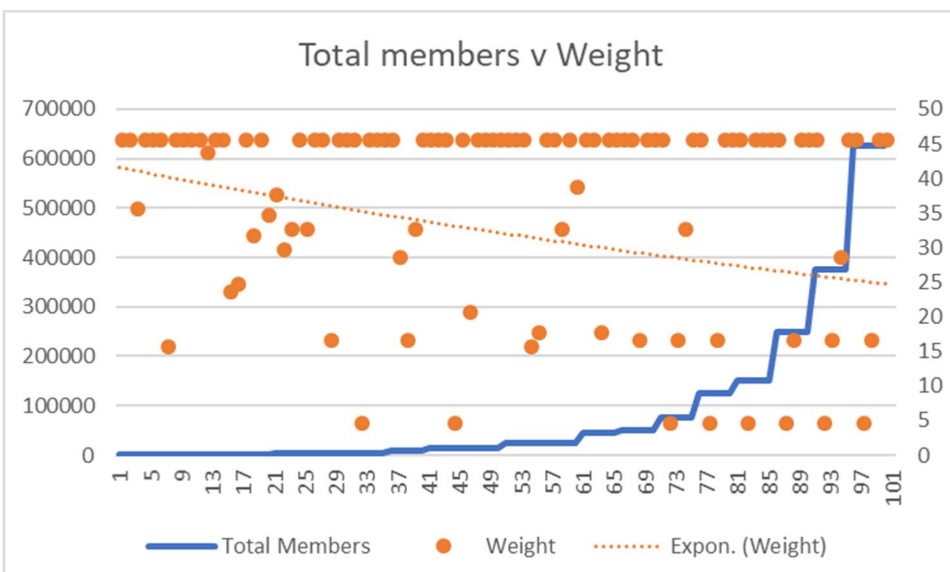


Figure 48. First GA implementation results

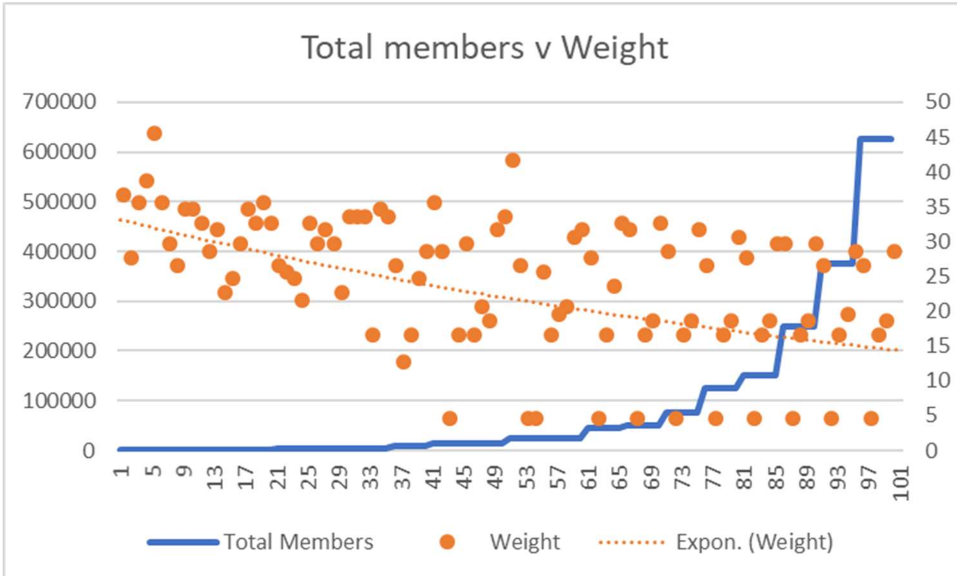


Figure 49. Latest GA implementation results

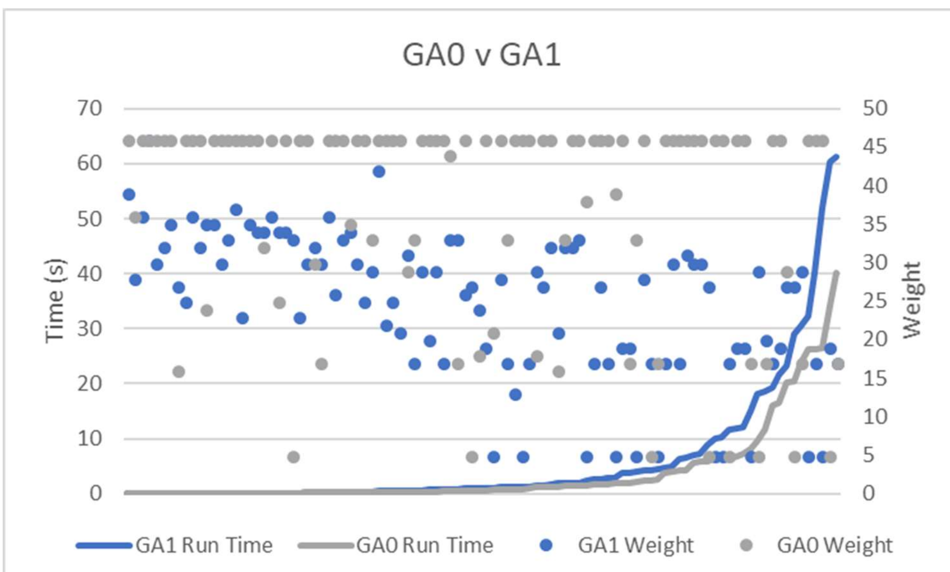


Figure 50. Time/Weight comparison for First GA0 and Latest GA1

5. Local Search

A. Top-Down Algorithm Design

Tabu search algorithm is an S-metaheuristic that has been a popular local search variant since the 1990s. It is very similar to a hill-climbing algorithm except it uses memory and accepts non-improving solutions to escape from local optima. Unlike the evolutionary algorithms, there is only one solution instead of many population members. The Tabu search (TS) uses short, medium, and long-term memories to diversify and reach good solutions. The short-term memory is the Tabu list, where recent moves are remembered and not repeated for a set number of moves, unless the solution surpasses some aspiration criteria. Medium-term memory is used to store the best solutions of the search and can be used to give priority to solutions that share characteristics with these solutions; this is known as intensification. Long-term memory can be used to remember the common moves and impose a penalty on any neighbor which includes these moves; this is diversification.

In the MSTG problem in this project the neighborhood solutions are defined in (Xu, Chiu, and Glover 1996) as any solution where a node is added, a node is removed, or nodes are swapped. We decided to use the first two of these three, which corresponds to a hamming distance of 1 because the solution is represented as a bit-string.

Problem Domain Requirements Specification

- domains, D
    - input $D_i(G,S),\ graph\ G=(V,E,W)$, set of vertices V, set of edges E, set of weights associated with each vertex W, set of goal nodes S
    - output D_o – Steiner tree T with additive weight – a set of connected vertices including at least the goal nodes.
- $I(x)$; input conditions on the domain satisfied
- $O(x,z)$; output conditions on output/input domain satisfied, i.e. a feasible/optimal solution with respect to the input domain

- Minimize additive weight subject to the following:
- $\text{Min}\{\Sigma w_k\}\ (1 \leq k \leq p) \mid u_{1\dots p} \rightarrow w_k$
- Where $u_{1\dots p}$ are the vertices of T.

PD/AD Integration Specification

- Next-state-generator
  - The next-state-generation for the Tabu Search comes from finding the best solution in the neighborhood of the current solution.
  - The neighborhood (next set of candidates) are binary strings which have a hamming distance of 1 from the current solution. This means that a node was either added to or removed from the current solution's set of vertices.
- Selection
  - In the TS the neighbor that is selected is the neighbor with the best fitness (even if it is worse than the current solution's fitness).
  - The neighbor must however, not be on the Tabu list, unless it meets the aspiration criteria.
  - The long-term memory is also used to impose a penalty on neighbors that are a result of moves that have been very common throughout the search. This encourages diversification.
- Feasibility
  - All binary strings are feasible. If we restrict the single population member only to actual solutions, we may end up using a random number generator to solve the problem instead of allowing the TS to perform.
- Solution
  - If a binary string is an actual feasible solution, its associated weight will be lower than the maximum weight. A penalty will be imposed on binary strings that do not meet constraints.
- Objective
  - The fitness of a solution is inverse to the total weight of all the nodes in T. The lower the weight, the better the solution. Binary strings that do not meet constraints will be assigned a maximum weight, equal to the total weight of all nodes in the graph.

B. Algorithm Design Specifications

Since we have already discussed that the solution is represented as a binary string and the neighborhood for a solution are those solutions with a hamming distance of 1, we must now discuss the other design aspects of the Tabu Search implementation.

First, we have the different memories, short (Tabu list), medium, and long-term. For this MSTG problem we have defined a move as adding or removing a node to a solution. For purposes of the Tabu list it is specifically adding or removing each specific node. For a certain preset number of moves (several numbers were tested, to be detailed later) a specific move is remembered in the Tabu list. Once in this list, we cannot select a neighbor that comes as a result of that specific move. For example, we do not want to continue adding and removing the same node. Having the Tabu list allows us to escape the local maxima and take a non-improving solution.

Next is medium-term memory, this is often used for intensification purposes. For example, we could remember the best 5 solutions and give priority to neighbors who have similar characteristics. We chose not to use medium-term memory because we wanted to get a wider scope of the landscape of the problem and avoid intensification.

Long-term memory was used for diversification purposes. Each time a specific node was added or removed a counter increased for that node. For every 5 counts a penalty of 1 weight was imposed on the neighboring solutions that would've resulted from that specific move. This encourages the search to look at adding and removing different nodes in the landscape, hence diversification.

Aspiration criteria for this problem was simple. If a Tabu solution was better than the current best solution, it met aspiration criteria and was permitted.

C. Intermediate Algorithm Designs

The first iteration of the TS worked quite well. Only a few changes were introduced along the way. One change included implementing another diversification measure. If no improving solution was found (or all moves were Tabu) for 10 turns in a row, a new random solution replaced the current member. This allows us to better escape local loops.

Second, we required the first solution to include all special nodes. This helped on some graphs and hurt on others (No Free Lunch), as we will see later. We considered making the initial solution include all nodes so that we could prune nodes until we got to a good solution, but this would basically turn into something very close to a greedy deterministic search.

D. Algorithm Pseudo-Code

```
s = s0; /* Initial solution */
Initialize the tabu list and long-term memories;
Repeat
Find best admissible neighbor s'; //objective, feasibility, and solution
/* non-tabu or aspiration criterion holds */
s = s';
Update tabu list, aspiration conditions and long-term memories;
Until Stopping criteria satisfied
Output: Best solution found. // solution
```

E. Algorithm Implementation

For many of the same reasons listed in the GA section above, we chose to implement my TS algorithm in Python. A shortened version of the main loop is below to show the mapping between the search elements, pseudo code, and implemented code.

```
TABU = mem()            # Init memories
LONGTERM = mem()
for ITERATIONS in [10,50,200,500]:
   for TABUNUMBER in [1,2,3,4]:
      solution = random_solution()
      iter = 0
      while iter < ITERATIONS:
         solution = findBestNeighbor(solution) # s = s';
        //objective, feasibility, and solution
         if fitness(solution) <= MIN_FITNESS:
            MIN_FITNESS = fitness(solution)
            FITTEST_STRING = solution   // solution
         iter +=1
```

F. Evaluation Experiments between TS versions

Note that in the code above the TS was tested with different iteration limits and Tabu numbers (the number of turns a move remains Tabu) on the five different versions of the graph in Figure 38. Figure 51 and Figure 52 show how the number of iterations and Tabu number affects the weight of the best-found solution. Note that in none of these tests, there was an improvement from Tabu number of 1 to 2 and 2 to 3, but no benefit was realized from 3 to 4. Figure 53 shows that Tabu 1 takes slightly more time but is much better at finding feasible (and better) solutions in general.

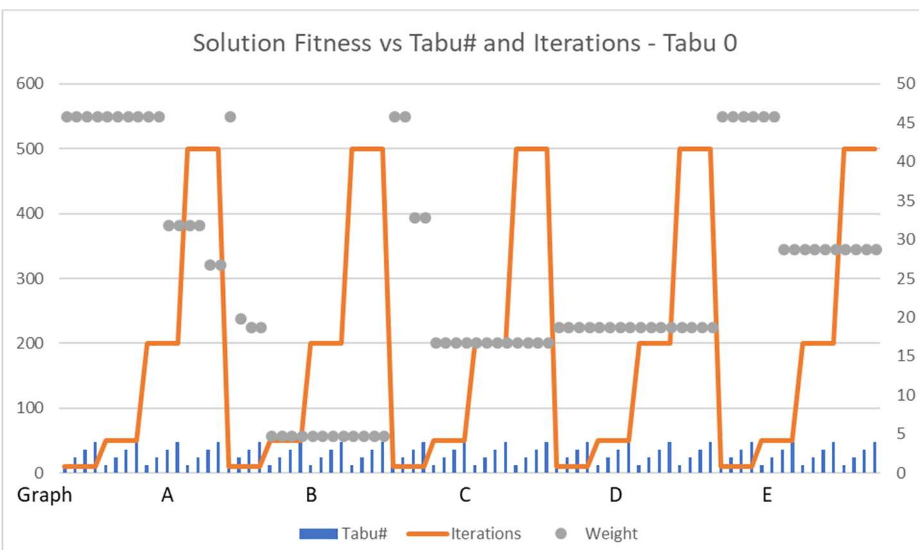


Figure 51. Initial TS Results

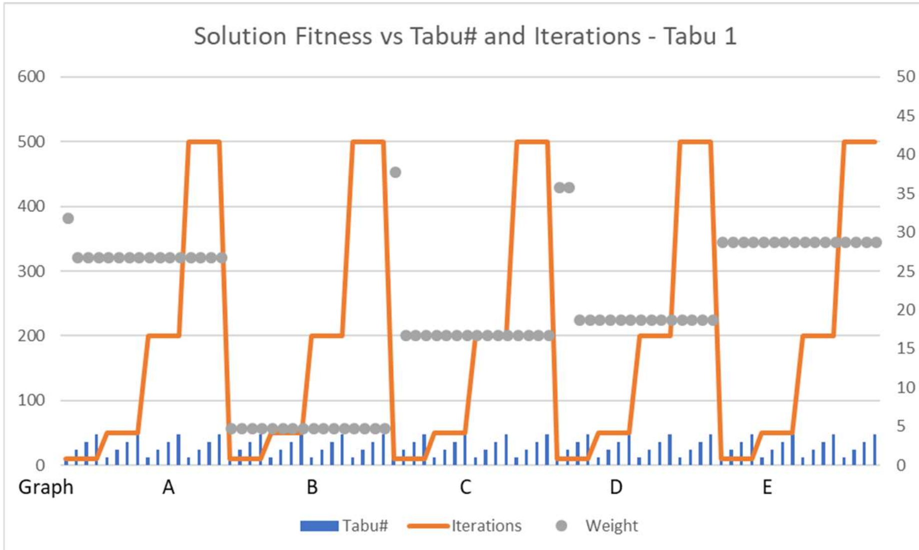


Figure 52. Latest TS Results

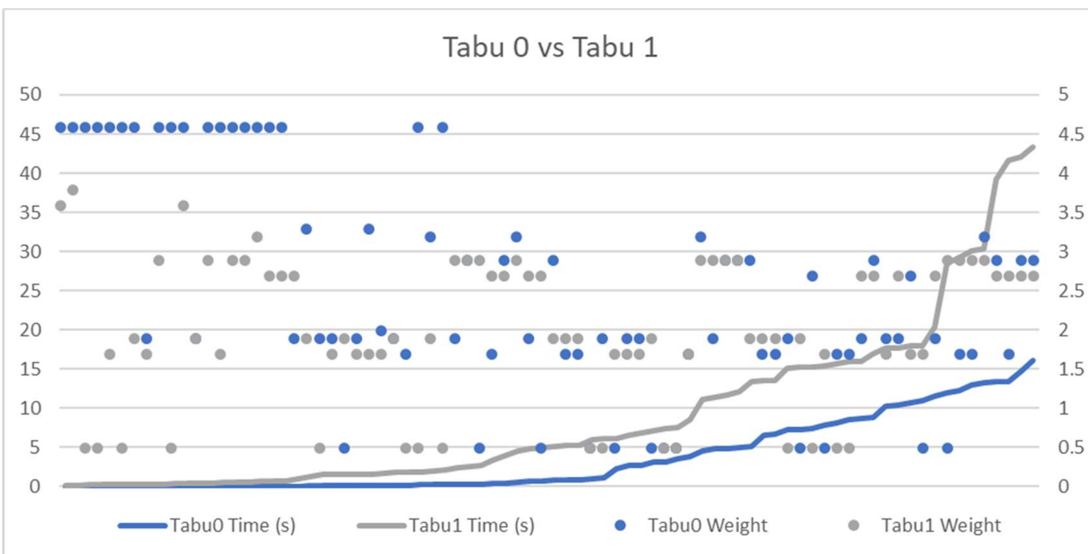


Figure 53. TS Weight/Time Comparison

6. Design and Evaluation of Experiments

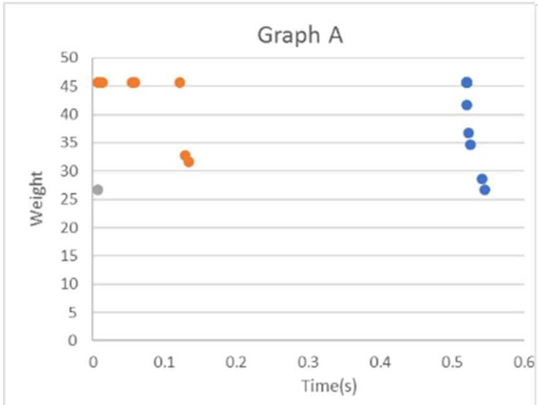
After implementing all three algorithms they were tuned once more and compared against each-other for the same five example problems used throughout this paper. The comparison was performed to see the relative computing performance of the three different types of searches. A summary of the results can be seen in the table below. More comprehensive charts that detail each graph problem can be seen later.

The primary adjustments to the algorithms were selecting specific values for population size (100) and generations (1500) for the GA and iterations (80) and Tabu number (3) for the TS. We tried to incorporate some of the heuristics of the DFS into these other searches but this actually had a negative effect on the computing performance. Since we trimmed the tree and made less solutions feasible, the stochastic search (GA) and the stochastic elements of the TS had a harder time judging fitness because solutions were now either very good or (in most cases) unfeasible. So, although the heuristics worked well for the DFS on these examples, they did not help us in our other searches. Again, refer to the NFL theorem ("No Free Lunch in Search and Optimization" 2018).

The programs were tested on a personal machine with an Intel i5 3.8GHz 4-core CPU and 16 GB 2400 MHZ DDR4 RAM.

Graph	Min Weight	DFS (s)	Brute (s)	GA (s)	Tabu (s)	Non-Optimum Weight
A	27	0.005	7.727	0.543	x	32
B	5	0.050	14.379	0.185	0.044	N/A
C	17	0.016	10.749	0.571	0.064	N/A
D	19	0.025	10.747	0.198	0.063	N/A
E	29	0.044	8.651	x	0.098	32

Figure 54. Comparison of all project algorithms



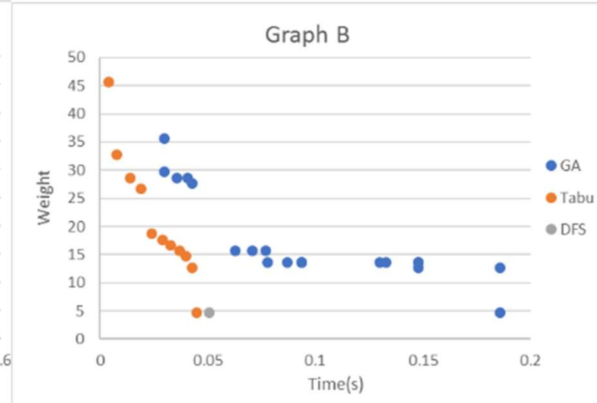
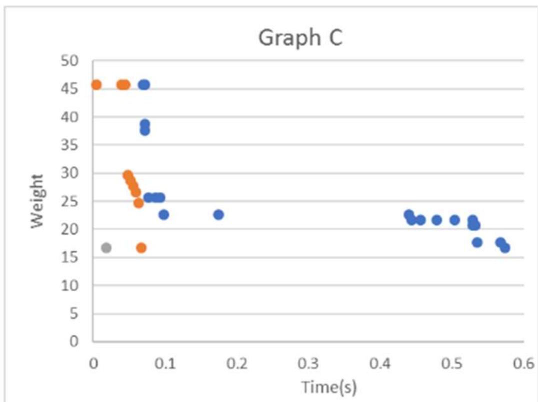


Figure 55. Comparison of algorithms for Graph A

Figure 56. Comparison of algorithms for Graph B



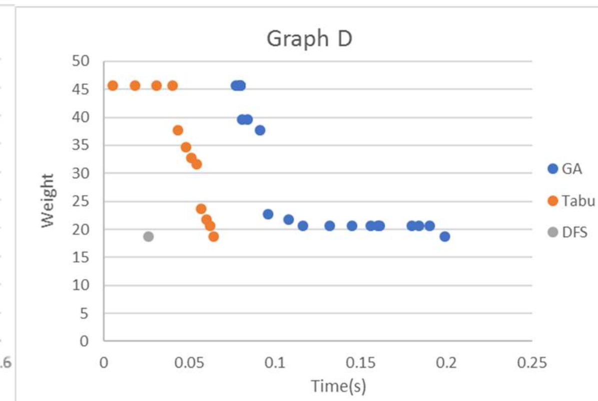


Figure 57. Comparison of algorithms for Graph C

Figure 58. Comparison of algorithms for Graph D

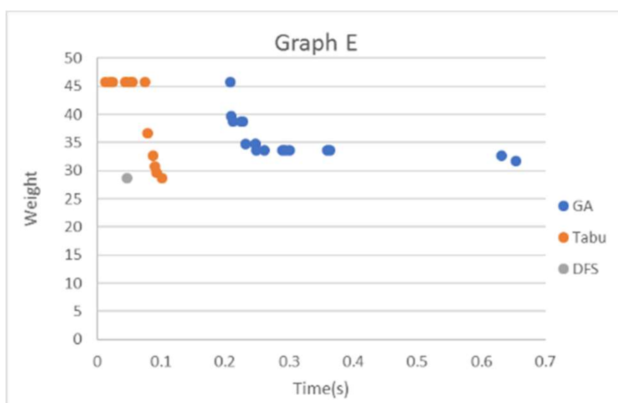


Figure 59. Comparison of project algorithms for Graph E

7. Conclusions

An interesting observation from the Figures above is that the DFS outperformed the others on all but one of the five examples. We thought that it was because a DFS may work better on smaller problems but it was actually outperformed on Graph B, which had the smallest set of special nodes. Since the TS was able to perform a good local search it was able to find a solution faster while the DFS had to span many more branches because of the many node removal permutations.

The TS outperformed the GA on all three graphs where they both found the optimum in less than one second. The GA has a high start time to find the first solution because it has the requirement that every member in the initial population has to include the goal nodes, which chews a lot of clock. Each search failed to find the optimum on one of the five graphs. Again, note that these are only examples and the GA could outperform the TS or even both other searches given a different problem landscape. More research needs to be done to determine what size of a problem, if any, is needed to see this performance change.

In conclusion, each search technique we developed in the project is valuable depending on the problem instance and landscape. Thanks to the reality of the NFL theorem ("No Free Lunch in Search and Optimization" 2018) it is likely in our best interest to keep all of these algorithms in our back pocket for use in different scenarios. These algorithms should be tested more thoroughly and improved in the future. There are many useful techniques in designing an algorithm but there is no parallel to spending time understanding the problem domain and using your knowledge to incorporate appropriate heuristics into your search algorithms.

Figures, Tables, and Documents

A. Base Survey

Determining the Motivation Model of Airmen for Use in Developing Effective Education and Training Platforms

Career Information and Demographics

1. **Age**

   ____

2. **Years of Service**

   Military and/or government service

   ____

3. **Status**

   Mark only one oval.

   ◯ Officer
   ◯ Enlisted
   ◯ Govt. Civilian

4. **Career Field (AFSC, etc.)**

   ____

5. **What is your highest education level?**

   Mark only one oval.

   ◯ No diploma
   ◯ High school graduate or GED
   ◯ Some college
   ◯ Associate's degree
   ◯ Bachelor's degree
   ◯ Master's degree
   ◯ Doctoral or professional school degree (PhD, MD, JD, DVM, EdD)

6. **Do you have any cyber security work experience?**

   Mark only one oval.

   ◯ Yes
   ◯ No

7. **Have you earned any cyber security certifications? (education / training)**

   Check all that apply.

   ☐ CompTIA Security+
   ☐ CISSP: Certified Information Systems Security Professional
   ☐ CISM: Certified Information Security Manager
   ☐ CEH: Certified Ethical Hacker
   ☐ GSEC: SANS GIAC Security Essentials
   ☐ None
   ☐ Other: ______________________

Motivation

8. *Mark only one oval per row.*

		Strongly disagree	Disagree	Slightly disagree	Nether agree or disagree	Slightly agree	Agree	Strongly Agree
(CD1.1)	I choose to do things that make me feel like I am serving a higher purpose than myself							
(CD1.2)	I choose to do things that make me feel like I am fighting for a greater good							
(CD2.1)	I choose to do things that make me feel like I am developing myself							
(CD2.2)	I choose to do things that allow me to track my accomplishments and show progress							
(CD3.1)	I choose to do things that make me feel empowered to use my creativity							
(CD3.2)	I choose to do things that give me feedback on how I am doing							
(CD4.1)	I choose to do things that allow me to own things							
(CD4.2)	I choose to do things that allow me to customize things							
(CD5.1)	I choose to do things that relate to me or my past							
(CD5.2)	I choose to do things that allow me to interact with friends and/or family							
(CD6.1)	I choose to do things that are limited events or have exclusive prizes							
(CD6.2)	I choose to do things that are difficult challenges that are rare to achieve							
(CD7.1)	I choose to do things that have unpredictable elements							
(CD7.2)	I choose to do things that pique my curiosity							
(CD8.1)	I choose to do things to avoid punishment or loss							
(CD8.2)	I choose to do things to avoid missing out on opportunities							

Gaming

On open ended questions: "Do NOT discuss or comment on classified or operationally sensitive information. We cannot provide confidentiality to a participant regarding comments involving criminal activity/behavior, or statements that pose a threat to yourself or others."

9. **How often do you participate in gaming using each of the following?**

   Mark only one oval per row.

	Daily	Weekly	Monthly	Rarely	Never
Desktop/Laptop	○	○	○	○	○
TV/Game Console	○	○	○	○	○
Smartphone/Tablet	○	○	○	○	○
Board Games	○	○	○	○	○

10. **What kind of games do you play?**

    Check all that apply.

- ☐ MOBA (Massive Online Battle Arena - ex. League of Legends, DOTA, Heroes of the Storm)
- ☐ RPG (Role-Playing Game - ex. Skyrim, Pokemon, Final Fantasy)
- ☐ MMORPG (Massive Multiplayer Online Role-Playing Game - ex. World of Warcraft)
- ☐ RTS (Real Time Strategy - ex. Starcraft, Civilization, Age of Empires)
- ☐ Shooters (Call of Duty, Halo, Medal of Honor, Doom)
- ☐ Fighters (Mortal Kombat, Super Smash Brothers, Tekken)
- ☐ Puzzle (Candy Crush, Tetris, Portal)
- ☐ Racing (Need 4 Speed, Forza, Dirt)
- ☐ Survival (Fortnite, Zombies, PUBG)
- ☐ Sandbox/Open World (Grand Theft Auto, Minecraft, Kerbal Space Program)
- ☐ Sports Games (NBA2K, Madden, Rocket League)
- ☐ Other: ______________________

11. **What is your favorite thing about games? Why do you play games?**

____

____

____

____

____

Enjoyment

For the following questions: If you don't play/enjoy games, replace the word "games" with "activities" and answer the questions accordingly. Thanks!

12. *Mark only one oval per row.*

		Strongly disagree	Disagree	Slightly disagree	Nether agree or disagree	Slightly agree	Agree	Strongly Agree
(CD1.1)	I enjoy games that make me feel like I am serving a higher purpose	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD1.2)	I enjoy games that make me feel like I am fighting for a greater good	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD2.1)	I enjoy games that make me feel like I am developing my character	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD2.2)	I enjoy games that allow me to track my accomplishments.	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD3.1)	I enjoy games that make me feel empowered to use my creativity	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD3.2)	I enjoy games that give me feedback	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD4.1)	I enjoy games that allow me own things	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD4.2)	I enjoy games that allow me to customize things	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD5.1)	I enjoy games that relate to me or my past	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD5.2)	I enjoy playing games with friends/family	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD6.1)	I enjoy games that have limited events and exclusive prizes	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD6.2)	I enjoy games that have difficult challenges that are rare to achieve	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD7.1)	I enjoy games that have unpredictable elements	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD7.2)	I enjoy games that piquemy curiosity	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD8.1)	I enjoy games that I have to play in order to avoid losing my progress	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD8.2)	I enjoy games that I have to play in order to avoid missing opportunities	⬭	⬭	⬭	⬭	⬭	⬭	⬭

13. **Would you consider yourself a "gamer?"**
Mark only one oval.

- Yes
- No
- Not sure

Air Force Education Platforms

14. **How often do you take the DoD Cyber Awareness Challenge?**

Mark only one oval.

- Daily
- Weekly
- Monthly
- Only When Required
- Never

15. **How often do you use the Air Force's Advanced Distributed Learning Service (ADLS) ?**
Mark only one oval.

- Daily
- Weekly
- Monthly
- Only when required
- Rarely
- Never

16. **How often do you use military education/training platforms such as milSuite, milTube, etc.?**
Mark only one oval.

- Daily
- Weekly
- Monthly
- Only when required
- Rarely
- Never

User Identification

17. **Participant EID Number**

This survey is confidential and the assigned Experiment ID number (found on the Cyber Education Hub) is obscured to ensure that your identity is protected. Personal data will not be shared or released.

B. Post-Experiment Survey

Determining the Motivation Model of Airmen for Use in Developing Effective Education and Training Platforms Post-Survey

On open ended questions: "Do NOT discuss or comment on classified or operationally sensitive information. We cannot provide confidentiality to a participant regarding comments involving criminal activity/behavior, or statements that pose a threat to yourself or others."

1. **Cyber Education Hub (CEH)**

Mark only one oval per row.

	Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
I enjoy using the Cyber Education Hub	⬭	⬭	⬭	⬭	⬭	⬭	⬭
Using the Cyber Education Hub motivated me to consume more outside educational/training content	⬭	⬭	⬭	⬭	⬭	⬭	⬭
I am motivated to continue using the Cyber Education Hub in the future	⬭	⬭	⬭	⬭	⬭	⬭	⬭

2. **Why? Why not?**

Referring to the above question "I am motivated to continue using the Cyber Education Hub in the future"

____

____

____

____

____

3. **Over the past three weeks I consumed more educational content on:**

Mark only one oval.

- ⬭ The Cyber Education Hub
- ⬭ Sources other than the Cyber Education Hub
- ⬭ Each Equally

4. **Over the past three weeks, how much time overall did you spend on the CEH?**
Mark only one oval.

- < 2 hours
- 2-4 hours
- 5-8 hours
- 9-12 hours
- 13-16 hours
- 17-24 hours
- 25-48 hours
- 48+ hours

5. **How frequently did you access the Cyber Education Hub in each of the following locations?**
Mark only one oval per row.

	Never	Rarely	Sometimes	Often	Very Often
From Home					
At Work / In the Office					
During Class					
Other					

6. **Did you have a CAC reader for home access?**
Mark only one oval.

- Yes
- No

7. **If not, would having a CAC reader at home have increased your time spent on the Cyber Education Hub?**
Mark only one oval.

- Yes
- No
- Maybe / Not sure

8. **Over the past three weeks, how much time overall did you spend consuming outside cyber educational content?**
Mark only one oval.

- < 2 hours
- 2-4 hours
- 5-8 hours
- 9-12 hours
- 13-16 hours
- 17-24 hours
- 25-48 hours
- 48+ hours

9. **Reasons why I used CEH:**
Check all that apply.

- [ ] Work/Job-related reasons
- [ ] Learn about cyber
- [ ] Curious about a new platform
- [ ] Give feedback and help shape the platform
- [ ] Referred by a friend/peer
- [ ] Referred by a superior
- [ ] See recent content from the community
- [ ] Other: ____________________

10. **Reasons why I did not use CEH:**
Check all that apply.

- [ ] Not enough content for me
- [ ] Too busy
- [ ] Frustrating internet connection
- [ ] Frustrating interface
- [ ] Not interested in cyber
- [ ] Other: ____________________

11. *Mark only one oval per row.*

		Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
(CD1.1)	Using the CEH makes me feel like I am serving a higher purpose	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD1.2)	Using the CEH makes me feel like I am acting for a greater good	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD2.1)	Using the CEH makes me feel like I am developing myself	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD2.2)	The CEH allows me to track my accomplishments	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD3.1)	Using the CEH makes me feel empowered to use my creativity	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD3.2)	The CEH gives me feedback	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD4.1)	Using the CEH allows me to own content on the site	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD4.2)	The CEH allows me to customize the site	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD5.1)	The CEH relates to me and my career	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD5.2)	The CEH allows me to interact with peers/friends	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD6.1)	The CEH has limited/exclusive elements	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD6.2)	The CEH has difficult challenges to strive towards	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD7.1)	The CEH has unpredictable elements when it comes to content	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD7.2)	The CEH piques my curiosity	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD8.1)	The CEH is something that I feel I have to use in order to avoid losing my educational/training progress	⬭	⬭	⬭	⬭	⬭	⬭	⬭
(CD8.2)	The CEH is something that I feel I have to use in order to avoid missing out on opportunities	⬭	⬭	⬭	⬭	⬭	⬭	⬭

Topic Map / Knowledge Skill Ability (KSA) Tree

On open ended questions: "Do NOT discuss or comment on classified or operationally sensitive information. We cannot provide confidentiality to a participant regarding comments involving criminal activity/behavior, or statements that pose a threat to yourself or others."

12. *Mark only one oval per row.*

	Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
The Topic Map helped me to consume more content on the Cyber Education Hub	⬭	⬭	⬭	⬭	⬭	⬭	⬭
The Topic Map helped me to consume more content OUTSIDE of the Cyber Education Hub	⬭	⬭	⬭	⬭	⬭	⬭	⬭
The layout of the Topic Map was intuitive	⬭	⬭	⬭	⬭	⬭	⬭	⬭
The use of the Topic Map was intuitive	⬭	⬭	⬭	⬭	⬭	⬭	⬭

13. **Suggestions for improvement of the Topic Map:**

____
____
____
____
____

14. **Reasons why I used the Topic Map:**
Check all that apply.

- [ ] Dicover content on the platform
- [ ] To navigate to/from familiar topics
- [ ] Explore the 'cyber universe'
- [ ] Interesting new/unique feature
- [ ] Contributing content and associating topics
- [ ] Other: _______________

15. **Reasons why I did not use the Topic Map:**
Check all that apply.

- [ ] Latency
- [ ] Confusing
- [ ] Not interesting
- [ ] No way to track what I've already seen or where I have been
- [ ] Lack of orientation
- [ ] Could not find what I was looking for
- [ ] Not personalizable
- [ ] Zoom In/Out was frustrating
- [ ] Other: _______________

16. *Mark only one oval per row.*

	Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
The KSA Tree helped me to consume more content on the Cyber Education Hub	⬭	⬭	⬭	⬭	⬭	⬭	⬭
The KSA Tree helped me to consume more content OUTSIDE of the Cyber Education Hub	⬭	⬭	⬭	⬭	⬭	⬭	⬭
The layout of the KSA Tree was intuitive	⬭	⬭	⬭	⬭	⬭	⬭	⬭
The use of the KSA Tree was intuitive	⬭	⬭	⬭	⬭	⬭	⬭	⬭

17. **Suggestions for improvement of the KSA Tree:**

____
____
____
____
____

18. **Reasons why I used the KSA Tree:**
Check all that apply.

- ☐ Interesting new/unique feature
- ☐ Track my accomplishments
- ☐ Follow a guided learning path
- ☐ Good balance of choice and structure
- ☐ interesting topics
- ☐ Other: ____________________

19. **Reasons why I did not use the KSA Tree:**
Check all that apply.

- ☐ Latency
- ☐ Confusing
- ☐ Not interesting
- ☐ No way to track what I've already seen or where I have been
- ☐ Lack of orientation
- ☐ Could not find what I was looking for
- ☐ Not personalizable
- ☐ Zoom In/Out was frustrating
- ☐ Too restrictive
- ☐ Not restrictive enough
- ☐ Other: ____________________

20. *Mark only one oval per row.*

	Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
I am familiar with YouTube's layout for content	⬭	⬭	⬭	⬭	⬭	⬭	⬭
I am familiar with prerequisite-style learning (content that you must watch in order)	⬭	⬭	⬭	⬭	⬭	⬭	⬭

21. **Which layout of content do you prefer?**
Mark only one oval.

	1	2	3	4	5	6	7	
Topic Map layout	⬭	⬭	⬭	⬭	⬭	⬭	⬭	YouTube-style layout

22. **Which layout of content do you prefer?**
Mark only one oval.

	1	2	3	4	5	6	7	
Topic Map layout	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Prerequisite-style layout

23. **Which layout of content do you prefer?**
Mark only one oval.

	1	2	3	4	5	6	7	
YouTube-style layout	⬭	⬭	⬭	⬭	⬭	⬭	⬭	KSA Tree layout

24. **Which layout of content do you prefer?**
Mark only one oval.

	1	2	3	4	5	6	7	
Prerequisite-style layout	⬭	⬭	⬭	⬭	⬭	⬭	⬭	KSA Tree layout

25. **Which layout of content do you prefer?**
Mark only one oval.

	1	2	3	4	5	6	7	
Topic Map layout	⬭	⬭	⬭	⬭	⬭	⬭	⬭	KSA Tree layout

26. *Mark only one oval per row.*

	Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
I enjoyed the CSCE 525 KSA Tree challenges and tasks	⬭	⬭	⬭	⬭	⬭	⬭	⬭
I enjoyed the Mobile Technology KSA Tree challenges and tasks	⬭	⬭	⬭	⬭	⬭	⬭	⬭
I enjoyed the KSA Tree topic of CSCE 525	⬭	⬭	⬭	⬭	⬭	⬭	⬭
I was interested in CSCE 525 before using the KSA Tree	⬭	⬭	⬭	⬭	⬭	⬭	⬭
The skill tree motivated me to learn about CSCE 525	⬭	⬭	⬭	⬭	⬭	⬭	⬭
I enjoyed the KSA Tree topic of Mobile Technology	⬭	⬭	⬭	⬭	⬭	⬭	⬭
I was interested in Mobile Technology before using the KSA Tree	⬭	⬭	⬭	⬭	⬭	⬭	⬭
The KSA Tree motivated me to learn about Mobile Technology	⬭	⬭	⬭	⬭	⬭	⬭	⬭

27. **Would you like to see a Topic Map or KSA Tree implemented in other applications?**
Mark only one oval.

- ⬭ Yes
- ⬭ No
- ⬭ Maybe

28. **If so, where?**
Check all that apply.

- ☐ Air Force Training
- ☐ Air Force Career Paths
- ☐ AFIT Degree Programs
- ☐ Other: ____________________

29. **Who should create new Topic Maps / KSA Trees?**
Check all that apply.

- ☐ The Community of Airmen
- ☐ Career Field Managers
- ☐ AFIT Academic Program Managers
- ☐ Individuals using these elements
- ☐ Supervisors of Airmen using these elements
- ☐ Commanders of Airmen using these elements
- ☐ Other: ____________________

User Identification

30. **Participant EID Number**

This survey is confidential and the assigned Experiment ID number (found on the Cyber Education Hub) is obscured to ensure that your identity is protected. Personal data will not be shared or released.

____

Other Feedback

There are no more questions on this survey. If you have comments or concerns that you were not able to express in answering this survey, please enter them in the space provided. Your feedback is useful and appreciated!

On open ended questions: "Do NOT discuss or comment on classified or operationally sensitive information. We cannot provide confidentiality to a participant regarding comments involving criminal activity/behavior, or statements that pose a threat to yourself or others."

31.

____

____

____

____

____

C. Article VII. Survey Results

In this section we present survey data that may be of interest to other researchers, but would potentially distract from the flow and scope of the main document. Various tables, figures and text excerpts are presented.

	CD 1.1	CD 1.2	CD 2.1	CD 2.2	CD 3.1	CD 3.2	CD 4.1	CD 4.2	CD 5.1	CD 5.2	CD 6.1	CD 6.2	CD 7.1	CD 7.2	CD 8.1	CD 8.2
Avg.	6.25	6.00	6.33	5.92	5.00	5.67	5.25	5.42	5.58	6.00	4.08	5.17	5.00	6.17	5.17	5.58
δ	0.25		0.42		0.67		0.17		0.42		1.08		1.17		0.42	
σ	0.60	0.82	0.94	0.76	1.41	1.11	1.48	0.95	1.61	1.58	1.85	1.62	1.29	0.90	1.40	1.11
Max	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
Min	5	4	4	5	2	4	2	3	1	1	1	2	2	4	2	3

Table 25 – Sub-question statistics of CD: Act questions for the classroom group.

	CD 1.1	CD 1.2	CD 2.1	CD 2.2	CD 3.1	CD 3.2	CD 4.1	CD 4.2	CD 5.1	CD 5.2	CD 6.1	CD 6.2	CD 7.1	CD 7.2	CD 8.1	CD 8.2
Avg.	4.42	4.58	6.08	5.92	4.92	5.33	5.25	5.50	5.17	6.58	4.92	6.08	6.25	6.25	3.83	3.42
d	0.17		0.17		0.42		0.25		1.42		1.17		0.00		0.42	
sd	1.04	1.32	1.26	1.11	1.85	1.03	1.79	1.71	1.28	0.64	1.66	1.44	0.92	0.72	1.99	1.80
Max	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	6
Min	3	2	3	4	1	4	2	2	3	5	2	2	4	5	1	1

Table 26 – Sub-question statistics of CD: Game questions for the classroom group.

	CD 1.1	CD 1.2	CD 2.1	CD 2.2	CD 3.1	CD 3.2	CD 4.1	CD 4.2	CD 5.1	CD 5.2	CD 6.1	CD 6.2	CD 7.1	CD 7.2	CD 8.1	CD 8.2
Avg.	3.86	4.14	5.29	5.86	3.71	3.43	5.71	3.86	5.57	4.14	4.29	4.00	4.00	5.43	3.86	3.86
δ	0.29		0.57		0.29		1.86		1.43		0.29		1.43		0.00	
σ	1.36	1.55	1.48	1.36	1.03	1.59	1.03	1.36	1.76	1.55	1.03	1.60	1.41	0.73	1.64	1.64
Max	6	6	7	7	5	6	7	6	7	6	5	6	6	6	6	6
Min	1	1	2	3	2	1	4	2	2	2	2	2	2	4	2	2

Table 27 – Sub-question statistics of CD: CEH questions for the classroom group.

Average for:	CD: Act	CD: Game	CD: CEH
Avg. (1-7)	5.54	5.28	4.44
δ	0.57	0.50	0.77
σ	1.21	1.35	1.38
Max	7.00	6.94	6.13
Min	2.81	2.75	2.13

Table 28 – Comparison of averages for each metric used in Table 25, Table 26, and Table 27.

Figure 60 – Classroom participant CEH access location responses.

	CD 1.1	CD 1.2	CD 2.1	CD 2.2	CD 3.1	CD 3.2	CD 4.1	CD 4.2	CD 5.1	CD 5.2	CD 6.1	CD 6.2	CD 7.1	CD 7.2	CD 8.1	CD 8.2
Avg.	5.00	5.44	6.00	5.56	5.67	5.67	6.11	6.22	5.67	5.44	4.67	5.56	4.89	6.11	5.11	5.22
δ	0.44		0.44		0.00		0.11		0.22		0.89		1.22		0.11	
σ	1.33	1.34	0.47	1.07	1.15	0.82	0.99	1.03	1.05	1.42	1.33	0.68	1.29	0.57	1.85	1.40
Max	7	7	7	7	7	7	7	7	7	7	6	7	7	7	7	7
Min	2	2	5	3	3	4	4	4	4	2	3	5	2	5	1	2

Table 29 – Sub-question statistics of CD: Act questions for the operational group.

	CD 1.1	CD 1.2	CD 2.1	CD 2.2	CD 3.1	CD 3.2	CD 4.1	CD 4.2	CD 5.1	CD 5.2	CD 6.1	CD 6.2	CD 7.1	CD 7.2	CD 8.1	CD 8.2
Avg.	3.78	4.11	5.78	6.22	5.67	5.67	5.67	5.78	4.44	5.89	5.22	6.00	5.67	5.56	2.44	3.22
δ	0.33		0.44		0.00		0.11		1.44		0.78		0.11		0.78	
σ	1.31	1.45	1.62	0.92	1.33	0.94	1.05	1.13	1.07	1.10	0.92	1.05	0.94	1.26	1.34	1.23
Max	6	7	7	7	7	7	7	7	6	7	7	7	7	7	5	5
Min	2	2	2	4	3	4	4	4	2	4	4	4	4	3	1	1

Table 30 – Sub-question statistics of CD: Game questions for the operational group.

	CD 1.1	CD 1.2	CD 2.1	CD 2.2	CD 3.1	CD 3.2	CD 4.1	CD 4.2	CD 5.1	CD 5.2	CD 6.1	CD 6.2	CD 7.1	CD 7.2	CD 8.1	CD 8.2
Avg.	5.71	5.76	6.19	5.76	5.29	5.67	5.62	5.76	5.62	5.76	4.33	5.33	4.95	6.14	5.14	5.43
δ	0.05		0.43		0.38		0.14		0.14		1.00		1.19		0.29	
σ	1.16	1.11	0.79	0.92	1.35	0.99	1.36	1.06	1.40	1.54	1.67	1.32	1.29	0.77	1.61	1.26
Max	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
Min	2	2	4	3	2	4	2	3	1	1	1	2	2	4	1	2

Table 31 – Sub-question statistics of CD: Act questions for all participants.

	CD 1.1	CD 1.2	CD 2.1	CD 2.2	CD 3.1	CD 3.2	CD 4.1	CD 4.2	CD 5.1	CD 5.2	CD 6.1	CD 6.2	CD 7.1	CD 7.2	CD 8.1	CD 8.2
Avg.	4.14	4.38	5.95	6.05	5.24	5.48	5.43	5.62	4.86	6.29	5.05	6.05	6.00	5.95	3.24	3.33
δ	0.24		0.10		0.24		0.19		1.43		1.00		0.05		0.10	
σ	1.21	1.40	1.43	1.05	1.69	1.01	1.53	1.50	1.25	0.93	1.40	1.29	0.98	1.05	1.87	1.58
Max	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	6
Min	2	2	2	4	1	4	2	2	2	4	2	2	4	3	1	1

Table 32 – Sub-question statistics of CD: Game questions for all participants.

Operational User Post-Experiment Survey (Single Response)

For context please refer to VII. §2C. The participant agreed that they enjoy the platform and are motivated to use it in the future. The participant said that they were 'always interested in learning more about cyber' and they consumed more cyber educational/training content from sources other than the CEH. The participant said that they rarely accessed the platform and when they did, it was at work and also that having a CAC reader at home would not have increased their time on the website. The reasons listed for why they did not use the platform were 'too busy,' and that there was 'not enough content for me.' Perhaps if there were more users, there would be enough content for everyone, and the platform would pass the tipping point (Gladwell 2002).

CD: CEH – CD5 had the only two statements relating to Octalysis that the participant strongly agreed with. The participant responded that they used the Topic Map and it helped them consume more content on the CEH. They also responded that they like the Topic Map more than both the You-tube style and prerequisite-style layouts. The participant was less pleased with the KSAT, citing that 'no content appears for me.' This is either a design issue where the user does not understand how to use the KSAT, and enhancements need to be made, or the participant simply meant that the two available KSAT topics of 'Mobile Technology' and 'CSCE 525' were not relevant to them. Finally, the participant stated "I haven't used the site yet as much as I'd like. Would like to see more content and maybe some type of training/walkthrough of the site." A KSAT that works as a CEH website tutorial vehicle should be considered for future development.

Officers v Enlisted Octalysis Models

Based on the demographics reported above we can see that 62% of participants were officers (n = 13) and 33% were Enlisted members (n = 7). Civilian Personnel were not included in either group for this comparison. Similar results are given for Younger and Older participants. These groups were selected based on age with a cap placed at age 27 for Younger participants. With this cap, Younger members included 57% (n=12) of participants and the remaining Older members made up 43% (n=9) of participants. The average age of Officers in this survey was 23.9 and the average age of Enlisted members was 33.2. After closer inspection of the resulting groups, 3 Officers were in the Older group and 1 Civilian and 1 Enlisted member were in the younger group.

Figure 61 shows the resulting Octalysis Models for these groups and Table 33 shows the spreads (calculated via Equation (4)) between them. Table 33 also shows the spreads for the Younger v Older groups, which was quite similar to the Officer v Enlisted spreads. However, this only shows differences and not the average values. To further demonstrate the slight differences that resulted from the group changes of a few members, we also show the spreads of Younger v Officer and Older v Enlisted groups in Table 33. Since there were no spreads greater than 6% in these comparisons, Octalysis Models are not shown for Older and Younger groups, as they are very much like those seen in Figure 61. While potentially redundant for this survey, differences in these groups may appear in other iterations of this study with a different or larger population.

The largest spreads between both comparisons in Figure 61 were revealed to be in CD5 for both CD: Act and CD: Game questions. From Officers to Enlisted and Younger to Older CD5 decreases by approximately 2 categories on the Likert scale (which would be 33% spread) for CD: Act and 1 category (which would be 17% spread) for CD: Game responses. This suggests that Enlisted and Older members are less motivated to act and also garner less enjoyment through CD5:

Social Influence and Relatedness game elements. No other spreads were over 10%. For this study, prior-enlisted officers were not separated from the officer group for analysis, but this should be explored in the future.

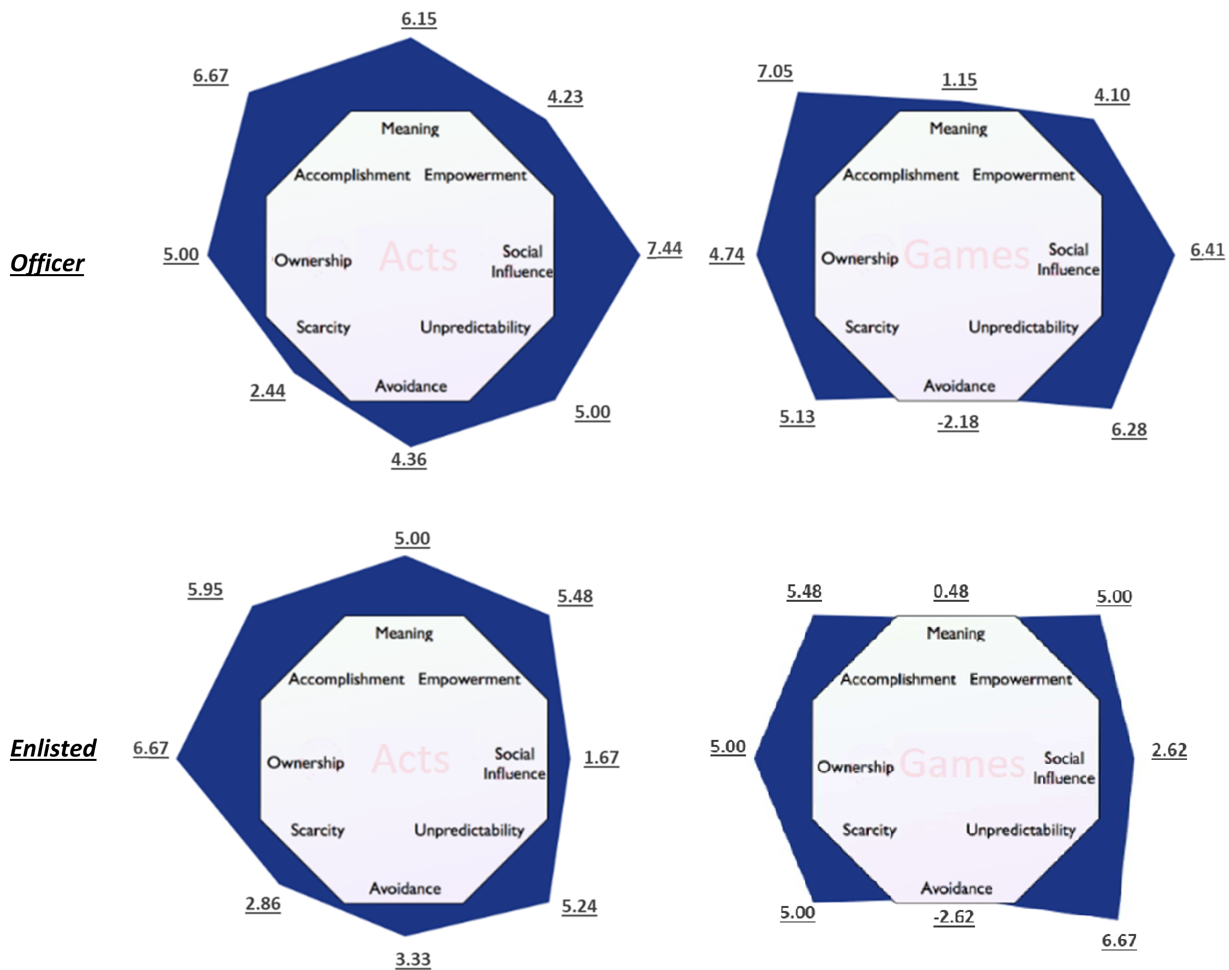


Figure 61 – Octalysis Models for Officers and Enlisted participants.

Officers v Enlisted	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	6%	4%	6%	8%	29%	2%	1%	5%	8%
Games	3%	8%	4%	1%	19%	1%	2%	2%	5%

Younger v Older	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	2%	4%	6%	1%	25%	10%	8%	10%	8%
Games	2%	8%	1%	9%	17%	9%	7%	0%	7%

Younger v Officer	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	1%	1%	6%	3%	2%	6%	4%	4%	3%
Games	2%	2%	2%	5%	1%	4%	2%	1%	2%

Older v Enlisted	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	3%	1%	6%	6%	6%	6%	5%	1%	4%
Games	3%	1%	3%	5%	4%	5%	4%	1%	3%

Table 33 – Spreads for CDs of different subgroups.

Junior v Senior Octalysis Models

Next, a Junior v Senior comparison based on years of service (YoS) was performed. Capping the YoS at 5 years would result in a nearly the same grouping as Younger v Older save one member. Capping at 6 years would also only make a 1-member difference. Thus, the cap was placed at 7 years, resulting in 71% Junior members (n=15) and 29% Senior members (n=6).

Figure 62 shows the resulting Octalysis Models for the Junior and Senior groups. As you may notice, the differences are more noticeable than those in the Officer v Enlisted subgroups. Table 34 also shows these results in terms of percent spread between each group for each CD. The largest differences between Junior and Senior participants are apparent in CD5, CD8, and CD6 for both CD: Act and CD: Game responses and also in CD: Game - CD2. In all but 2 instances, the responses are higher for Junior members than for those Senior. The exceptions are CD: Act – CD4 and CD: Game – CD7. While this may account for some of the spread, the shapes of the Octalysis Models between Junior and Senior participants still vary, especially in CD5. These results suggest that it may be more difficult to motivate senior members with game elements relating to CD5, CD6, and CD8 compared to Junior members.

Since years of service and age are related, we also performed Junior v Younger and Senior v Older comparisons. Again, the Octalysis models for Younger and Older groups are not presented due to the low variance between those groups and what we have already presented. One finding is that compared to the Senior participants, the Older participants responded more positively to CD8 questions by nearly 1 Likert category. Since all the Younger participants are also Junior, we can see that the Older Junior participants account for this spread.

Junior

Senior

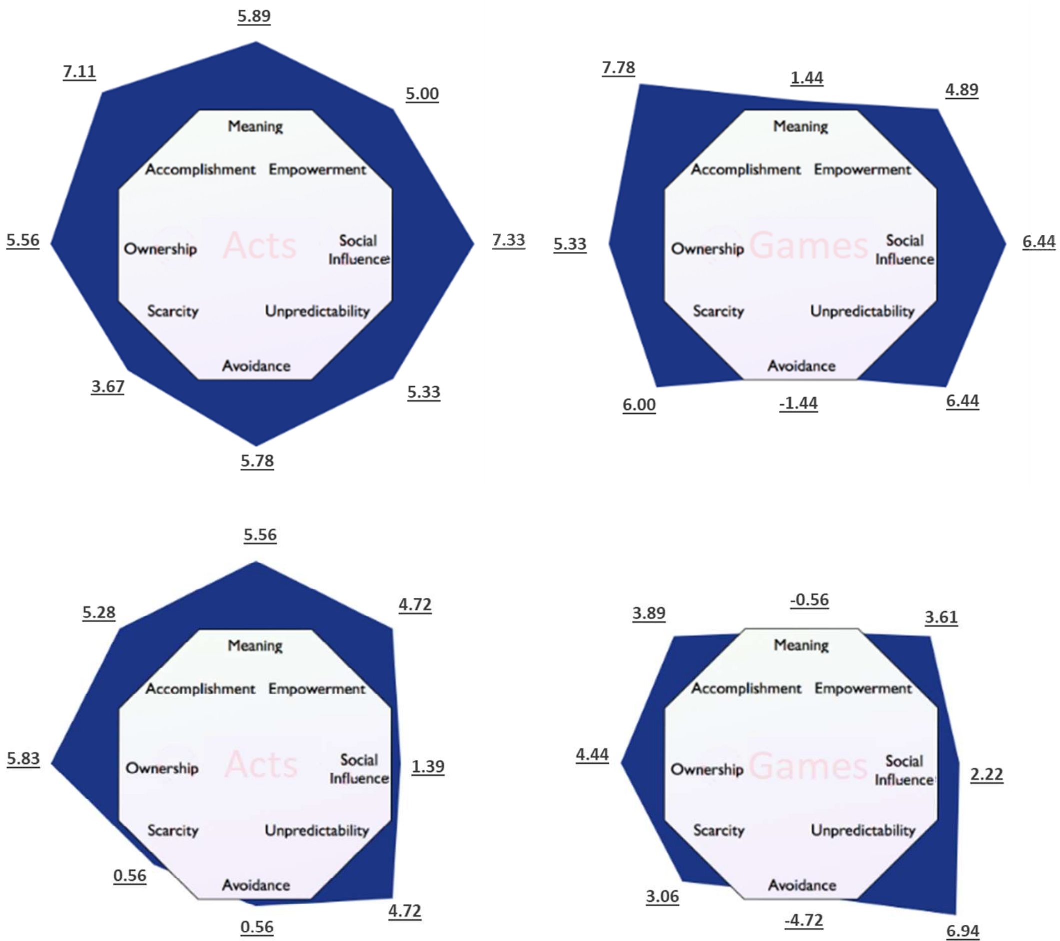


Figure 62 – Octalysis Models for Junior and Senior participants.

Junior v Senior	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	2%	9%	1%	1%	30%	16%	3%	26%	11%
Games	10%	19%	6%	4%	21%	15%	2%	16%	12%

Younger v Older	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	2%	4%	6%	1%	25%	10%	8%	10%	8%
Games	2%	8%	1%	9%	17%	9%	7%	0%	7%

Junior v Younger	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	0%	1%	2%	1%	2%	0%	3%	3%	2%
Games	4%	2%	2%	3%	1%	0%	2%	5%	2%

Senior v Older	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	Average
Acts	0%	4%	2%	1%	7%	6%	2%	13%	4%
Games	8%	9%	4%	2%	6%	5%	2%	12%	6%

Table 34 – Spreads for CDs of different subgroups.

Gamer and Non-Gamer Additional Insights

Looking more into the Gamers versus Non-Gamers and game-related questions, we found that from a list of 11 game types, seen in Table 35, gamers claimed to play 5.3 different game types on average, whereas non-gamers play 2.7 different game types on average. Some observations include that non-gamers play more RTS and Puzzle games than non-gamers based on percentage. Also, 87% of gamers play 'Shooter' games and 0% of non-gamers play MOBA, MMORPG, Survival, or Sandbox type games. While these extreme percentages are likely a result of the sample-size, it may be interesting to see if these results are similar on a larger scale. Perhaps there are certain game-elements in different game types that specifically appeal or repulse Gamers or Non-Gamers. The effectiveness of different game genres at capturing or diverting the attention of the target audience of the CEH may provide insight into why certain types of users interacted with the platform more than others, and may also inform design decisions of the website moving forward.

Game Type	# of Players (21 Total)
Multiplayer Online Battle Arena (MOBA)	3
Racing	4
Massive Multiplayer Online Role-Playing Game (MMORPG)	7
Fighters	8
Puzzle	8
Real-Time Strategy	9
Survival	9
Sandbox / Open World	9
Sports	9
Role-Playing Game (RPG)	11
Shooters	15

Table 35 – Number of players for each game type

Comparison of Gaming Octalysis Models

An open-ended question asked why participants play games and what about them they enjoy. Incidences of statements within responses relating to one of the eight CDs were tallied and presented as percentages of all participants who answered the question. For example, one response was "they provide a challenge, milestones, opportunities to get better at something to accomplish something harder, entertaining stories, competitiveness." This response includes allusions to CDs 2: Development and Accomplishment, 7: Unpredictability and Curiosity, and 5: Social Influence and Relatedness.

Since we have already generated the Octalysis Model for 'what participants enjoy about games/activities' for all participants based on the Likert-type responses, we can compare these two models. Because the models were generated in two distinct manners, we will solely compare the general shapes of the models. In Figure 63, the left Octalysis corresponds to the percentage of incidences of each CD in the open-ended question and the right Octalysis is the previously-seen (in Figure 33) model based on Likert responses. The percentages were rounded to their closest 0 to 10 relative and used for the scale of the blue area of the online Octalysis tool ("Octalysis / Gamification Building Developing Online Tool - by Yukai Chou" n.d.). For example, 6 was used for 56%, 3 was used for 33%, and so on.

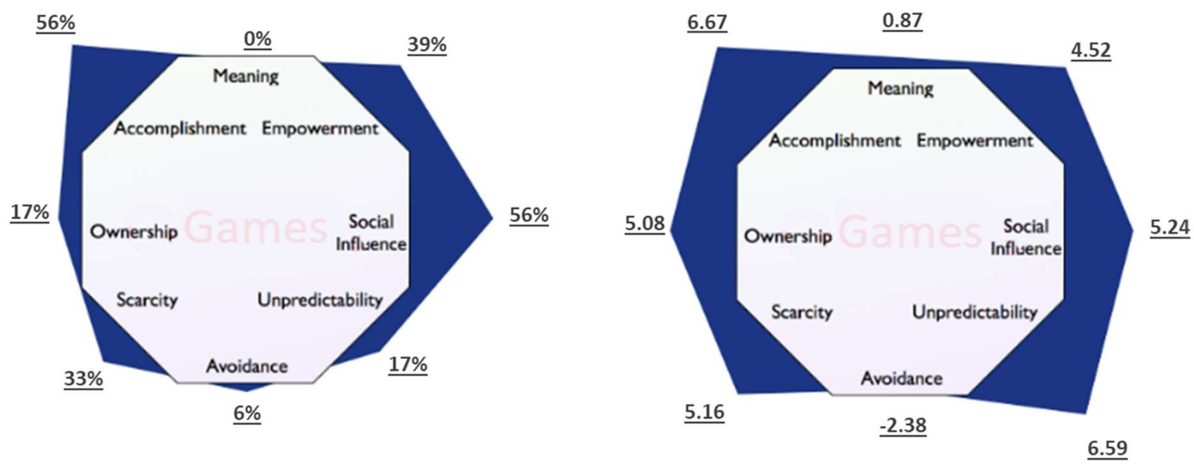


Figure 63 – Open-ended responses (left) and Likert-scale responses (right) about what participants enjoy in games.

The major difference in what participants said they enjoyed in games in the open-ended format (left) and the Likert statement format (right) can primarily be seen in CD7, CD4, and CD6. Participants responded very similarly about CD5, CD1, CD2, and CD3. Again, these responses were scaled differently and are from different question types, but something that we may take away is that these participants are familiar with the game elements in CD5, CD2, and CD5, but they also enjoy other game elements that they do not initially assert. This should be considered when looking at feedback taken from subjects, especially when delivered in open-ended responses. At some point, when users are asked to give feedback about the CEH, they may not have criticisms or not offer any suggestions of additions, but this piece of data may suggest that there are still areas that can be improved to increase enjoyment with the CEH platform and subsequently motivation and engagement.

D. Article VII. Engagement Results

In this section we present engagement tracking data that may be of interest to other researchers, but would potentially distract from the flow and scope of the main document. Various tables, figures, and text excerpts are presented.

Views

When looking at the data for content item views by participants, in Table 36 and Table 37, it is important to note that video and YouTube views may appear inflated. For an external webpage or file type content item, the view count is incremented each time as user opens up the webpage or downloads the file. In contrast, internal and external video view counts increase each time the 'play' button is clicked or a user skips to a new time in the video. There may also be other factors for this disparity in view counts by subtype: a user could bookmark an external webpage and access it without going through the CEH, or a user could reopen a file without redownloading it from the CEH, these views are not accounted for on the platform, and likely cannot be.

	All	CSCE 525 Class	CSCE 525 Post-Class	All Operational	88CS	33NWS
Video	410	388	3	19	14	5
Web Page	134	114	1	19	15	4
PDF	120	116	0	4	4	0
YouTube	85	85	0	0	0	0
Other	5	3	0	2	2	0

Table 36 – Content item view count by subtype.

When looking at the view counts broken down by source in Table 37, we see that the majority of views from the classroom participants came via 'KSAT Activities.' However, even more views came from 'Recently Added' content when combining the two ways to access this feature. 'Suggested Content-Recently Added' is the list of content items that appears as suggestions when accessing another content item; 'Recently Added' is currently the default suggestion tool. Once the 'Popular' and 'Trending' algorithms are implemented, perhaps they will draw more views. During the experiment, users likely saw the same content suggested by these tools during each session because they used placeholder values to determine the suggestion order.

	All	CSCE 525 Class	CSCE 525 Post-Class	All Operational	88CS	33NWS
Total Views	754	706	4	44	35	9
KSAT Activity	251	250	0	1	0	1
Home-Recently Added	182	166	1	15	15	0
Sugested Content-Recently Added	123	113	1	9	3	6
EditWrap-	106	106	0	0	0	0
?	40	37	0	3	3	0
Search	35	32	3	0	0	0
Home-NICE Academic Category	10	0	10	0	0	0
Home - Operational Category	7	0	3	4	2	2
Home-Trending	2	2	0	0	0	0

Table 37 – Content item view count by source.

Likes and Comments

Among all 40 participants there were 7 'likes' from 3 users. Comments were primarily for the classroom CTCs, which were almost certainly all 'Long' and 'Very Long' comments as discussed in VII. §2F. 106 content items were contributed by classroom participants, with only 72 likely CTC comments. This suggests that 34 content items were contributed and 14 comments were provided for other-than-classroom purposes.

	All	CSCE 525 Class	CSCE 525 Post-Class	All Operational	88CS	33NWS
Total Comments	88	86	0	2	2	0
Very Short Comments	3	1	0	2	2	0
Short Comments	4	4	0	0	0	0
Medium Comments	9	9	0	0	0	0
Long Comments	60	60	0	0	0	0
Very Long Comments	12	12	0	0	0	0
Longest Comment (char)	3919	3919	0	43	43	0

Table 38 – Comment counts.

Topic Map

Topic Map use primarily consisted of participants associating their contributed content with the nodes of the Topic Map. Participants utilized the Topic Map as a way to potentially get more views on their content by making it easier to find in more places, but ironically, very few users explored the Topic Map nodes for content, and as mentioned previously, the Topic Map was not a direct source for any content item views. The most popular nodes in terms of contributions to and clicks are shown in Table 39. A visual of the Topic Map and the counts for each node can be seen in Figure 64. In total, there were 115 Topic Map associations/contributions (of 111 contributed content items) and 40 Topic Map clicks.

Topic Map Node	Contributions	Clicks
Cyber Warfare	11	2
Strategy	9	2
Vulnerabilities	8	2
Human Element	7	0
Social Engineering	5	0
Risk Management	5	1
Vulnerability Assessment	4	1
China	4	0
Education	3	0
Networks	3	2
Access Control	3	1

Topic Map Node	Contributions	Clicks
Incident Response Team	0	9
Databases	0	4
Hardware	2	3
Software	2	3
Cyber Warfare	11	2
Strategy	9	2
Vulnerabilities	8	2
Networks	3	2

Table 39 – Popular Topic Map nodes by contributions (left) and clicks (right).

Inspection of Figure 64 shows us that much of the Topic Map was contributed to and most of the contributions were in the area near ‘Cyber Warfare.’ This makes sense because most of the contributions were from student participants in the ‘Introduction to Cyber Warfare’ class. The exploration (or clicks) of the Topic Map were spread out. Most of the larger nodes were clicked, but most of the clicks were not necessarily on the largest nodes. About 16 of 40 clicks were on the larger nodes. ‘Malware’ and ‘Social Engineering’ were the largest nodes to not be clicked. The ‘Cryptography,’ ‘Digital Forensics,’ and ‘Networks’ sections of the Topic Map were largely untouched. Asterisks (*) in Figure 64 represent 1 click or 1 contribution

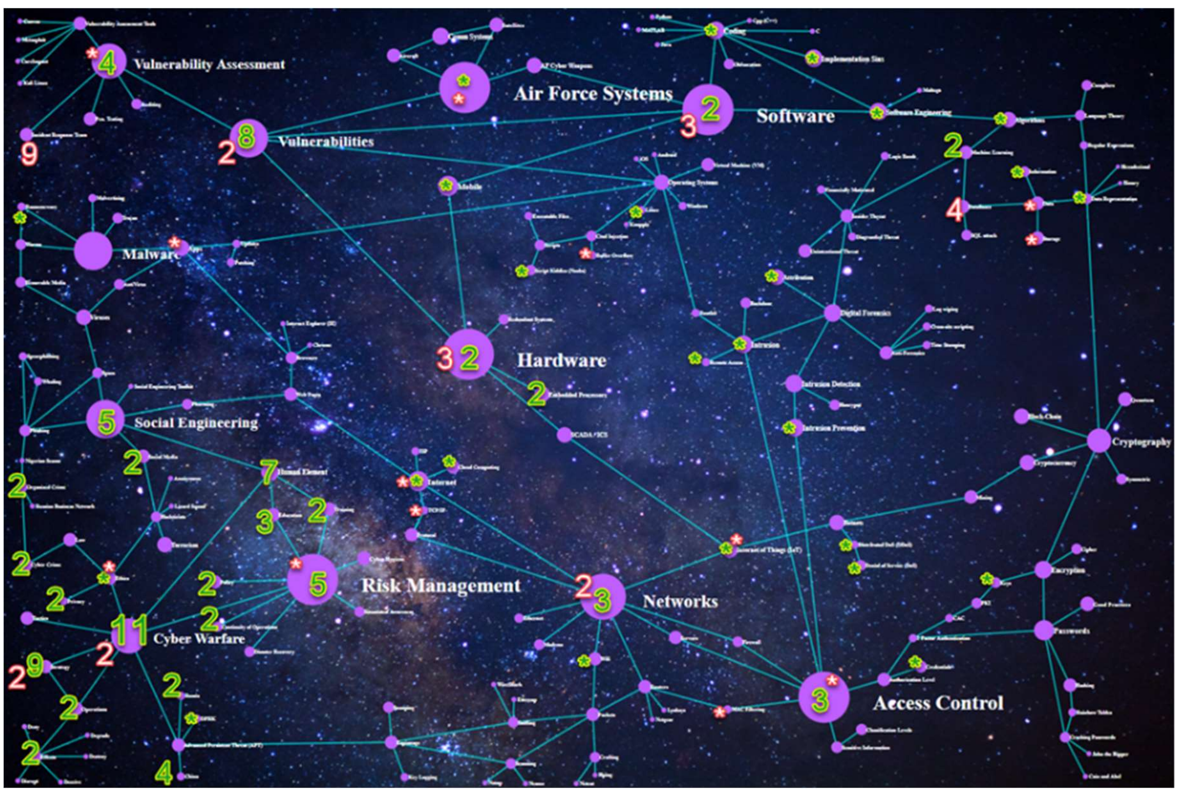


Figure 64 – Topic Map contribution counts per node (green text, yellow outline) and clicks per node (white text, red outline) (“Cyber Education Hub” 2019).

KSA Tree

Table 40 shows the total KSAT Activity for the CSCE 525 Tree. Classroom and Operational group activity are included in the table, but Operational users only accounted for 2 'navs' (navigations) and 1 net completed activity. The 'unique completed activities' and 'available unique activities' within each node are displayed in Figure 65 within parentheses. Pink numbers in the figure represent the net number of completed activities per node by all participants.

KSAT Node	Navs	Links Clicked	Completed Activities	Unique Activities Completed
Cyber Actors	45	18	59	7
Cyber Battlefield	39	33	46	6
Introduction to Cyber Warfare	35	17	85	8
Security Design Principles	30	41	21	5
Cyber Doctrine	27	26	52	7
Embedded Systems	26	4	17	4
Fundamental Design	26	21	14	4
Cyber Threatscape	24	22	60	6
Social Engineering	20	18	27	6
Network Defenses	17	5	23	6
Network Exploitation	16	14	15	4
Cyber Warriors	15	13	31	6
Cyber Weapons	15	19	28	5
Cyber Deterrence	14	13	42	8
Legal/Ethics	13	7	25	8
Network Attacks	12	23	27	9
Cyber Risk	8	11	10	6
Logical Defenses	6	0	9	8
Physical Attacks	5	0	8	8
Logical Attacks	4	1	8	8
Regular Expressions	4	2	0	0
Challenges and the Future	3	0	6	6
APT	3	2	2	4
Hacktivists	2	0	0	0
IoT	2	1	2	2
Physical Defenses	1	0	8	8
Quantum	1	0	3	3
Hashing	1	0	3	3
Multi-Factor Authentication	1	2	4	4

Table 40 – Total KSAT Activity.

Most completions were near the start node (bottom left, 'Introduction to Cyber Warfare'). Nodes that corresponded to class lessons (square) were completed the most, but some optional nodes (circle) were still explored and completed. All activities from course-related nodes were completed at least once. Table 41 shows keywords within activity names, the amount of unique activities containing the keyword, and the number of associated completions for each keyword.

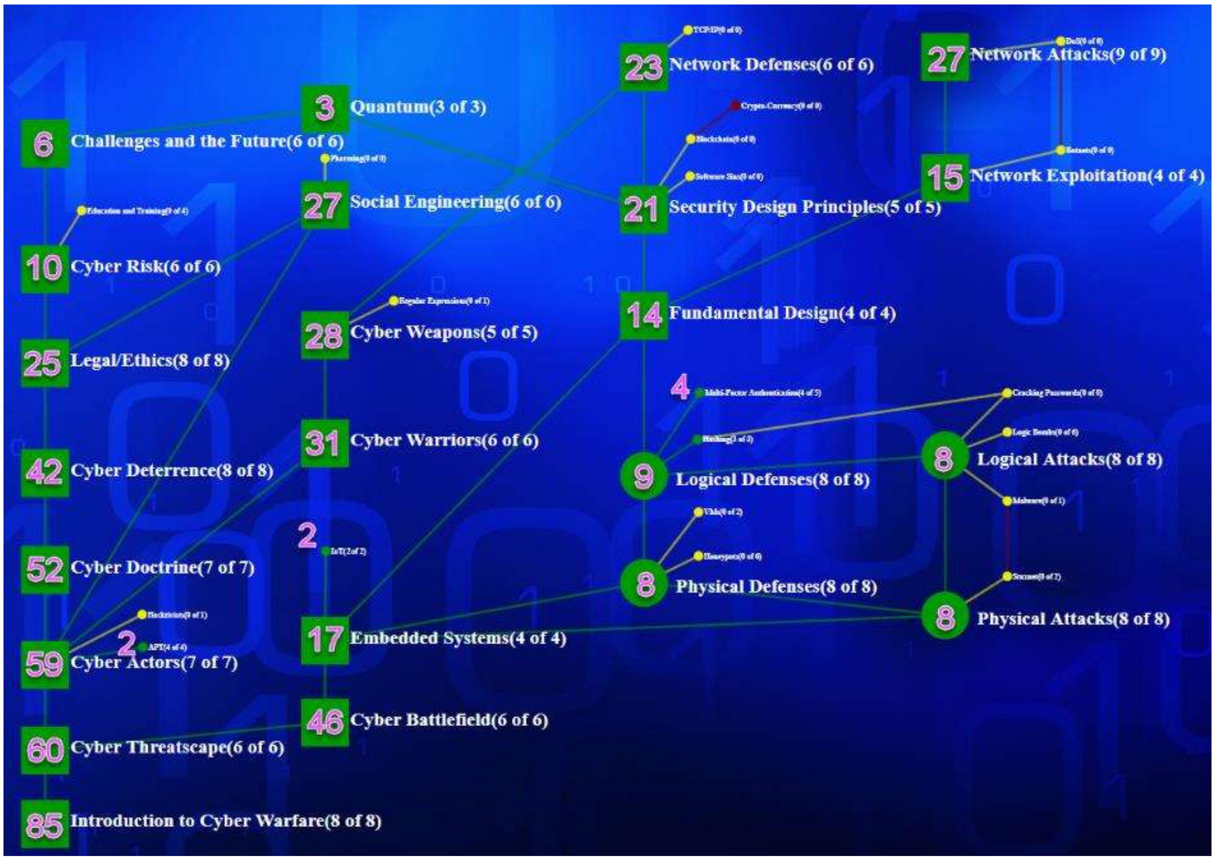


Figure 65 – CSCE 525 KSAT. Total net activities completed per node (pink numbers) and the number of unique activities completed and available (x of y, in parentheses) are shown (“Cyber Education Hub” 2019).

‘Review’ and ‘Easy’ were the two most prevalent keywords within activities, but their ratio of net completions over unique activities completed are much different. Also, activities with the keyword ‘Optional’ were completed an average of almost 4 times each out of 14 participants. ‘Optional’ only appeared within class-related nodes, activities within circle nodes were already optional and more frequently contain the ‘Easy,’ ‘Medium,’ ‘Hard,’ etc. keywords.

Keyword	Completed Activities	Unique Acts	Ratio
Total Completed Activities	635	x	x
Review	227	31	7.3226
Critical Thinking Check	102	18	5.6667
Skim	58	14	4.1429
Easy	35	33	1.0606
Optional	15	4	3.75
Explore	12	4	3
View	9	5	1.8
Medium	6	6	1
Hard	6	6	1
Expert	5	5	1
Game	3	1	3
Watch	1	1	1
Master	0	0	0

Table 41 – Keywords and counts of completed activities.

X. Bibliography

Adner, Ron, and Rahul Kapoor. 2016. “Innovation Ecosystems and the Pace of Substitution: Re-Examining Technology S-Curves.” *Strategic Management Journal* 37 (4): 625–48. https://doi.org/10.1002/smj.2363.

“AFSC 17X Cyberspace Operations Officer Career Field Education and Training Plan.” 2015. https://static.e-publishing.af.mil/production/1/saf_cio_a6/publication/cfetp17x/cfetp17x.pdf.

Alm, Antonie. 2006. “CALL for Autonomy,Competence and Relatedness: Motivating Language Learning Environments in Web 2.0.” *The JALT CALL Journal* 2 (3): 29–38. http://jaltcall.org/journal/articles/2_3_Alm.pdf.

Andress, Jason, and Steve Winterfield. 2014. *Cyber Warfare Techniques, Tactics and Tools for Security Practitioners*. 2nd ed. Waltham, MA: Elselvier.

Antiel, Ryan M., Scott M. Thompson, Christopher L. Camp, Geoffrey B. Thompson, and David R. Farley. 2012. “Attracting Students to Surgical Careers: Preclinical Surgical Experience.” *Journal of Surgical Education* 69 (3). Elsevier: 301–5. https://doi.org/10.1016/J.JSURG.2011.10.001.

Batagelj, Vladimir, and Andrej Mrvar. 2011. “Pajek.” http://vlado.fmf.uni-lj.si/pub/networks/pajek/.

“Borderlands 2.” 2012. Gearbox Software. 2012. https://borderlands.com/en-US/.

Burke, Brian. 2014. *Gamify: How Gamification Motivates People to Do Extraordinary Things*. Brookline, MA: Bibliomotion.

Carifio, James, and Rocco Perla. 2008. “Resolving the 50-Year Debate around Using and Misusing Likert Scales.” *Medical Education* 42 (12). John Wiley & Sons, Ltd (10.1111): 1150–52. https://doi.org/10.1111/j.1365-2923.2008.03172.x.

Chlebík, Miroslav, and Janka Chlebíková. 2002. “Approximation Hardness of the Steiner Tree Problem on Graphs.” In , 170–79. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45471-3_18.

Chou, Yu-Kai. 2015. *Actionable Gamification*. California: Octalysis Media.

Cisco Advisory Services. 2015. “Mitigating the Cybersecurity Skills Shortage,” 6. http://www.cisco.com/c/dam/en/us/products/collateral/security/cybersecurity-talent.pdf.

"Civilization V." 2010. 2K Games. 2010. https://civilization.com/civilization-5/.

"Computer Engineering Major." 2018. Air Force Academy. 2018. https://www.academyadmissions.com/the-experience/academics/majors/computer-engineering-major/.

Corson, Trevor. 2008. *The Story of Sushi : An Unlikely Saga of Raw Fish and Rice*. Harper Perennial.

Csikszentmihalyi, M. 1990. "Flow. The Psychology of Optimal Experience." *Harper-Perrenial New York*. https://opus4.kobv.de/opus4-Fromm/frontdoor/index/index/docId/27641.

"Cyber Education Hub." 2019. U.S. Air Force. 2019. http://www.afit.edu/ceh.

Dacus, Chad. 2018. "Cyber Education & Training Presentation." Air Force Cyber College.

Danforth, Melissa, and Charles Lam. 2017. "Effects of a Four-Week Cyber Security Summer Program on the Attitudes and College Interests of High School Students." *Journal of The Colloquium for Information System Security Education (CISSE)* 4 (2): 75–93. https://cisse.info/resources/archives/category/43-edition-4-issue-2.

"Definition of Gamification." 2018. Merriam-Webster. 2018. https://www.merriam-webster.com/dictionary/gamification.

Desta, Yohana. 2018. "Netflix Pulls the Plug on Feature Designed to Get Kids Addicted to Netflix." Vanity Fair. 2018. https://www.vanityfair.com/hollywood/2018/03/netflix-patch-testing-kids-binge-watching.

DISA. 2018. "Cyber Awareness Challenge." 2018. https://iatraining.disa.mil/eta/disa_cac2018/launchPage.htm.

Dunning, David. 2011. *The Dunning – Kruger Effect : On Being Ignorant of One ' s Own Ignorance*. *Advances in Experimental Social Psychology*. 1st ed. Vol. 44. Elsevier Inc. https://doi.org/10.1016/B978-0-12-385522-0.00005-6.

Eddins, J. M. Jr. 2018. "Byte-Size Learning." Airman Magazine. 2018. http://airman.dodlive.mil/2018/05/08/bytesizelearning/.

Endicott-Popovsky, Barbara, and Viatcheslav Popovsky. 2017. "Searching and Developing Cybersecurity Talent." *Journal of The Colloquium for Information System Security Education* June.

Eyal, Nir. 2018. “People Don’t Want Something Truly New, They Want the Familiar Done Differently.” Nir and Far. 2018. https://medium.com/thrive-global/people-dont-want-something-truly-new-they-want-the-familiar-done-differently-979302796cc5.

Fruchterman, Thomas M. J., and Edward M. Reingold. 1991. “Graph Drawing by Force-directed Placement.” *Journal of Software: Practice and Experience* 21 (11): 1129–64.

Fui-Hoon Nah, Fiona, Qing Zeng, Venkata Rajasekhar Telaprolu, Abhishek Padmanabhuni Ayyappa, and Brenda Eschenbrenner. 2014. “Gamification of Education: A Review of Literature.” *Lncs* 8527: 401–9. https://doi.org/10.1007/978-3-319-07293-7.

Gladwell, Malcolm. 2002. *The Tipping Point: How Little Things Can Make a Big Difference*. Back Bay Books.

Hamari, Juho, Jonna Koivisto, and Harri Sarsa. 2014. “Does Gamification Work? - A Literature Review of Empirical Studies on Gamification.” *Proceedings of the Annual Hawaii International Conference on System Sciences*, 3025–34. https://doi.org/10.1109/HICSS.2014.377.

Howard, Michael, David LeBlanc, and John Viega. 2010. *24 Deadly Sind of Software Security: Programming Flaws and How to Fix Them*. McGraw-Hill.

Jamieson, Susan. 2004. “Likert Scales : How to (Ab) Use Them,” 1217–18. https://doi.org/10.1111/j.1365-2929.2004.02012.x.

Jones, Joseph, and Frederick C. Jr. Harris. n.d. “A Genetic Algorithm of the Steiner Minimal Tree Problem.”

Kamada, Tomihisa, and Satoru Kawai. 1989. “An Algorithm for Drawing General Undirected Graphs.” *Information Processing Letters* 31 (1). Elsevier: 7–15. https://doi.org/10.1016/0020-0190(89)90102-6.

Kapsalis, A, Victor Rayward-Smith, and George Smith. 1993. *Solving the Graphical Steiner Tree Problem Using Genetic Algorithms*. *Journal of the Operational Research Society*. Vol. 44. https://doi.org/10.1038/sj/jors/0440408.

Kearney, A.T. 2016. “Tough Choices: The Real Reasons A-Level Students Are Steering Clear of Science and Maths.” *Your Life*, 1–26. https://1cwcpp2qp6hi1exr69469mgu-wpengine.netdna-ssl.com/wp-content/uploads/2016/06/YL-ATK-Tough-Choices-Research-Report-FINAL-3-02-16.pdf%0Ahttp://yourlife.org.uk/tough-choices-the-real-reasons-a-level-students-are-steering-clear-of-science-and-maths/.

Kelley, Paul, and Terry Whatson. 2013. "Making Long-Term Memories in Minutes : A Spaced Learning Pattern from Memory Research in Education." *Human Neuroscience* 7 (September): 1–9. https://doi.org/10.3389/fnhum.2013.00589.

Klein, P., and R. Ravi. 1995. "A Nearly Best-Possible Approximation Algorithm for Node-Weighted Steiner Trees." *Journal of Algorithms* 19 (1). Academic Press: 104–15. https://doi.org/10.1006/JAGM.1995.1029.

Koch, Thorsten, and Alexander Martin. 1970. *Solving Steiner Tree Problems in Graphs to Optimality*. *Networks*. Vol. 32. https://doi.org/10.1002/(SICI)1097-0037(199810)32:33.0.CO;2-O.

Kusurkar, R. A., G. Croiset, and Olle Th. J. Ten Cate. 2011. "Twelve Tips to Stimulate Intrinsic Motivation in Students through Autonomy-Supportive Classroom Teaching Derived from Self-Determination Theory." *Medical Teacher* 33 (12): 978–82. https://doi.org/10.3109/0142159X.2011.599896.

Kyewski, Elias, and Nicole C. Krämer. 2018. "To Gamify or Not to Gamify? An Experimental Field Study of the Influence of Badges on Motivation, Activity, and Performance in an Online Learning Course." *Computers & Education* 118 (March). Pergamon: 25–37. https://doi.org/10.1016/J.COMPEDU.2017.11.006.

Lamb, Daniel. 2017. "Building on a Firm Foundation: Redesigning the Air Force Assignment System to Improve Readiness." OTH. 2017. https://othjournal.com/2017/08/24/building-on-a-firm-foundation-redesigning-the-air-force-assignment-system-to-improve-readiness/.

Leh, John, and Karl Kapp. 2018. "Gamification and Learning - Episode 3." The Talented Learning Show. https://talentedlearning.com/learning-gamification-podcast-karl-kapp/.

Lou, Jing-Kai, Kuan-Ta Chen, Hwai-Jung Hsu, and Chin-Laung Lei. 2012. "Forecasting Online Game Addictiveness." https://dl.acm.org/citation.cfm?id=2501568.

Maybury, Mark. 2015. "Toward the Assured Cyberspace Advantage: Air Force Cyber Vision 2025." *IEEE Security & Privacy* 13 (1). IEEE Computer Society: 49–56. https://doi.org/10.1109/MSP.2013.135.

Mcfarlane, Jameson, and Thorne Jameson Mcfarlane. 2017. "Academic Influence of Social Network Sites on the Collegiate Performance of Technical." *Journal of The Colloquium for Information System Security Education (CISSE)*, no. June.

Mehlhorn, Kurt. 1988. "A Faster Approximation Algorithm for the Steiner Problem in Graphs." *Information Processing Letters* 27 (3). Elsevier: 125–28. https://doi.org/10.1016/0020-

0190(88)90066-X.

"Minimum Spanning Tree." 2018. Wikimedia Foundation. 2018. en.wikipedia.org/wiki/Minimum_spanning_tree.

Neill, Thomas A O. 2017. "An Overview of Interrater Agreement on Likert Scales for Researchers and Practitioners" 8 (May). https://doi.org/10.3389/fpsyg.2017.00777.

Newhouse, William, Stephanie Keith, Benjamin Scribner, and Greg Witte. 2017. "National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework." https://doi.org/10.6028/NIST.SP.800-181.

NHI. 2000. "Principles of Adult Learning and Instructional Systems Design." *NHI Instructor Development Course* 14 (4): 2–8. https://doi.org/Principles%20of%20Adult%20Learning%20&%20Instructional%20Systems%20Design%0AThree.

"No Free Lunch in Search and Optimization." 2018. Wikimedia Foundation. 2018. https://en.wikipedia.org/wiki/No_free_lunch_in_search_and_optimization.

Norman, Geoff. 2010. "Likert Scales, Levels of Measurement and the ``laws'' of Statistics." *Advances in Health Sciences Education* 15 (5): 625–32. https://doi.org/10.1007/s10459-010-9222-y.

"Octalysis / Gamification Building Developing Online Tool - by Yukai Chou." n.d. Accessed January 22, 2019. https://www.yukaichou.com/octalysis-tool/.

Octalysis Group. 2018. "A Comprehensive List of 90+ Gamification Cases with ROI Stats." 2018. https://yukaichou.com/gamification-examples/gamification-stats-figures/?utm_source=Octalysis+Happenings&utm_campaign=b31294c7d0-EMAIL_CAMPAIGN_2017_09_25&utm_medium=email&utm_term=0_55a1efbedd-b31294c7d0-135764341&mc_cid=b31294c7d0&mc_eid=2a4efc5205#.Wu8p.

Oltramari, Alessandro, Lorrie Faith Cranor, Robert J. Walls, and Patrick D. McDaniel. 2014. "Building an Ontology of Cyber Security." *Semantic Technologies for Intelligence, Defense, and Security*. https://www.semanticscholar.org/paper/Building-an-Ontology-of-Cyber-Security-Oltramari-Cranor/11469352326248de37b299d6f49b0b5b107a083b.

"Path of Exile." 2013. Grinding Gear Games. 2013. https://www.pathofexile.com/game.

Peker, Yesem Kurt, Lydia Ray, Stephanie Da Silva, and Nathaniel Gibson. 2016. "Raising Cybersecurity Awareness among College Students." *Journal of The Colloquium for*

Information System Security Education (CISSE), no. September: 1–17.

Phillips, Anna C, Lucy K Lewis, Maureen P Mcevoy, James Galipeau, Paul Glasziou, Marilyn Hammick, David Moher, Julie K Tilson, and Marie T Williams. 2014. “A Delphi Survey to Determine How Educational Interventions for Evidence-Based Practice Should Be Reported : Stage 2 of the Development of a Reporting Guideline” 14 (1): 1–11. https://doi.org/10.1186/1472-6920-14-159.

Prömel, Hans Jürgen., and Angelika. Steger. 2002. *The Steiner Tree Problem : A Tour through Graphs, Algorithms, and Complexity*. Vieweg+Teubner Verlag.

Reith, Mark. 2016. “Forging Tomorrow’s Air, Space, and Cyber War Fighters: Recommendations for Integration and Development.” *Air and Space Power Journal* 30 (4): 96–107.

Reith, Mark, Seeley Penecost, Daniel Celebucki, and Robert Kaufman. 2017. “Operationalizing Cyber: Recommendations for Future Research.” *12th International Conference on Cyber Warfare and Security 2017*, 1–12. https://doi.org/10.1089/acm.2015.0314.

Reith, Mark, Eric Trias, Chad Dacus, Seth Martin, and Landon Tomcho. 2018. “Rethinking USAF Cyber Education & Training.” *13th International Conference on Cyber Warfare and Security ICCWS*. Academic Conferences and Publishing International Limited, 439–47.

Roberson, Darryl L., and Matthew C. Stafford. 2017. “The Redesigned Air Force Continuum of Learning: Rethinking Force Development for the Future.” *LeMay Center for Doctrine Development and Education*. Maxwell Air Force Base, Alabama. https://media.defense.gov/2017/Dec/05/2001852390/-1/-1/0/LP_0001_ROBERSON_STAFFORD_REDESIGNED_AIR_FORCE.PDF.

Sandvine Incorporated ULC. 2016. “2016 - Global Internet Phenomena - Latin America & North America.” https://www.sandvine.com/hubfs/downloads/archive/2016-global-internet-phenomena-report-latin-america-and-north-america.pdf.

Schmidt, Lara, Caolionn O Connell, Hirokazu Miyake, Akhil R Shah, Joshua William Baron, Geof Nieboer, Rose Jourdan, et al. 2015. *Cyber Practices: What Can the U.S. Air Force Learn from the Commercial Sector?* http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA621837.

Segal, David, and Mady Wechsler. 2004. “America’s Military Population.” *Population Bulletin* 59 (4): 1–44. https://www.prb.org/americasmilitarypopulationpdf627kb/.

Shannon, C. E. 1951. “Prediction and Entropy of Printed English.” *Bell System Technical Journal* 30 (1). John Wiley & Sons, Ltd: 50–64. https://doi.org/10.1002/j.1538-7305.1951.tb01366.x.

Shernoff, David J., Mihaly Csikszentmihalyi, Barbara Schneider, and Elisa Steele Shernoff. 2014. "Student Engagement in High School Classrooms from the Perspective of Flow Theory." In *Applications of Flow in Human Development and Education*, 475–94. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-9094-9_24.

"Shortest Path Problem." 2018. Wikimedia Foundation. 2018. https://en.wikipedia.org/wiki/Shortest_path_problem.

Singer, P.W, and Allen Friedman. 2014. *Cybersecurity and Cyberwar: What Everyone Needs to Know*. New York: Oxford University Press. https://doi.org/10.1016/S1353-4858(14)70039-X.

Sobiesk, Edward, Jean Blair, Gregory Conti, Michael Lanham, and Howard Taylor. 2015. "Cyber Education: A Multi-Level, Multi-Discipline Approach." *Sigite'15*, 43–47. https://doi.org/10.1145/2656450.2656478.

"Steiner Tree Problem." 2018. Wikimedia Foundation. 2018. en.wikipedia.org/wiki/Steiner_tree_problem.

"Sushi Industry Statistics." 2017. Statistic Brain. 2017. https://www.statisticbrain.com/sushi-industry-statistics/.

Talbi, El-Ghazali. 2009. *Metaheuristics : From Design to Implementation*. John Wiley & Sons. https://www.wiley.com/en-us/Metaheuristics%3A+From+Design+to+Implementation+-p-9780470278581.

"The Elder Scrolls | Skyrim." 2011. Bethesda. 2011. https://elderscrolls.bethesda.net/en/skyrim.

Tomcho, Landon, Alan Lin, David Long, Mark Coggins, and Mark Reith. 2019. "Applying Game Elements to Cyber ELearning: An Experimental Design." *14th International Conference on Cyber Warfare and Security*.

Tomcho, Landon, and Mark Reith. 2019. "Engaging Airmen with Cyber Education and Training." *Journal of The Colloquium for Information System Security Education*, no. 06.02: In Press.

Vagias, Wade M. 2006. "Likert-Type Scale Response Anchors," 4–5.

Vallerand, Robert J, Luc G Pelletier, and Richard M Ryan. 1991. "Motivation and Education: The Self-Determination Perspective AU - Deci, Edward L." *Educational Psychologist* 26 (3–4). Routledge: 325–46. https://doi.org/10.1080/00461520.1991.9653137.

Walliman, Dominic. 2009. "Domain of Science." 2009. https://www.flickr.com/photos/95869671@N08/.

Welch, B. L. 1947. "The Generalization of 'Student's' Problem When Several Different Population Variances Are Involved." *Biometrika* 34 (1–2). Oxford University Press: 28–35. https://doi.org/10.1093/biomet/34.1-2.28.

Werbach, Kevin, and David Hunter. 2012. *For the Win: How Game Thinking Can Revolutionize Your Business*. Philadelphia, PA: Wharton Digital Press.

Wingo, J. 2017. "Personal Interview as the USAF Cyberspace Operations Officer Career Field Manager. Used with Permission."

Xu, Jiefeng, Steve Y. Chiu, and Fred Glover. 1996. "Using Tabu Search to Solve the Steiner Tree-Star Problem in Telecommunications Network Design." *Telecommunication Systems* 6 (1). Baltzer Science Publishers, Baarn/Kluwer Academic Publishers: 117–25. https://doi.org/10.1007/BF02114289.

Yannakogeorgos, Panayotis A, and John P Geis. 2016. *The Human Side of Cyber Conflict Organizing , Training , and Equipping the Air Force Cyber Workforce*. Air University Press, Air Force Research Institute.

Zelikovsky, Alexander Z. 1993. "A Faster Approximation Algorithm for the Steiner Tree Problem in Graphs." *Information Processing Letters* 46 (2). Elsevier: 79–83. https://doi.org/10.1016/0020-0190(93)90201-J.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
21-03-2019	Master's Thesis	August 2017 – March 2019

TITLE AND SUBTITLE	
Motivating Airmen to Engage with Technical Education: Experimentation and Analysis Using Modern Gamification Techniques	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Tomcho, Landon G. M., 2d Lt, USAF	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/ENG) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865	AFIT-ENG-MS-19-M-061
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Intentionally left blank	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The development and integration of computer systems into today's society and the subsequent growth of cyber as a warfighting domain has led to changes in military and civilian conflict. Several traits unique to cyber, including disruption and fast pace of change, has led to issues never before seen in the military environment, especially with educating and training. A new approach that leverages crowd-sourced content has been proposed. This approach relies on motivating military members to voluntarily engage with technical (cyber) education.
The application of gamification, a design practice aimed at increasing user engagement by targeting core motivators in humans, in the military context is presented in this paper. The adaptation and evaluation of unique game elements onto the platform is also discussed. A human-subject study involving a survey and engagement-tracking experiment is implemented. Results are analyzed using visualization software and a novel framework we created.
We then present results explaining what core drives motivate military members on average and within subgroups. We also show that engagement data can be attributed to motivation levels. Finally, we present recommendations to military leadership and education platform designers based on our findings before discussing ideas for future work.

15. SUBJECT TERMS
Cyber education and training, gamification, software design, human subject experimentation

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Lt Col Mark A. Reith, AFIT/ENG
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	279	(937) 255-6565 x4603 mark.reith@afit.edu

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