3-2022

Exploring the Value of Training System Alternatives Through Methods Technologies, and Relevant Attributes

Brian R. Novitsky

EXPLORING THE VALUE OF TRAINING SYSTEM ALTERNATIVES THROUGH METHODS, TECHNOLOGIES, AND RELEVANT ATTRIBUTES

THESIS

Brian R. Novitsky

AFIT-ENV-MS-22-M-243

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-MS-22-M-243

EXPLORING THE VALUE OF TRAINING SYSTEM ALTERNATIVES THROUGH METHODS, TECHNOLOGIES, AND RELEVANT ATTRIBUTES

THESIS

Presented to the Faculty

Department of Systems 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 Systems Engineering

Brian R. Novitsky

March 2022

AFIT-ENV-MS-22-M-243

EXPLORING THE VALUE OF TRAINING SYSTEM ALTERNATIVES THROUGH METHODS, TECHNOLOGIES, AND RELEVANT ATTRIBUTES

Brian R. Novitsky

Committee Membership:

Dr. M. E. Miller
Co-Chair

Maj J. P. Kristbaum, PhD
Co-Chair

Lt Col P. M. Beach, PhD
Member

Dr. M. F. Schneider
Member

AFIT-ENV-MS-22-M-243

Abstract

Skilled maintainers are being lost at a higher rate than average of all enlisted career fields within the Air Force. Although incentive programs can provide some retention, attrition rates may continue to vary, placing the readiness of the warfighter in jeopardy. Additionally, the technologies used to train maintainers have evolved substantially, including augmented reality (AR) and virtual reality (VR) solutions, among others. These technologies have recently demonstrated value as novel ways to increase training effectiveness, providing avenues for maintainers to achieve greater skill levels at a faster rate than traditional platforms. Through semi-structured subject matter expert (SME) interviews, this qualitative research investigated the relationships between training technologies, methods, and student learning objectives to build a tool for organizations deciding between training alternatives. The results provided important attributes of training system alternatives to use in a total value function for cost-utility estimation and information on common problems requiring a training-based solution. Key findings also included SME input about the best training technology combinations for each training method and data demonstrating that value is a function of learning objectives. Together, these findings give maintenance organizations an initial roadmap on how to decide between training system alternatives using the value-focused thinking (VFT) framework.

For my wife, who has shown me love and patience, my parents, who set the foundation for me to be successful, and God, who is faithful, always.

Acknowledgments

I would like to thank my faculty advisor, Dr. Michael Miller for your sound guidance throughout this research effort. Thank you to my committee members Dr. Michael Schneider, Maj Joseph Kristbaum, and Lt Col Paul Beach for your valuable input. Your insight and experience were greatly appreciated. Thank you, MSgt James Earley, for being a true wingman to me during our time at AFIT. Finally, I would also like to thank the 711th HPW/RHB for sponsoring this thesis.

Brian R. Novitsky

Table of Contents

Page

Abstract iv

Acknowledgments vi

Table of Contents vii

List of Figures x

List of Tables xii

I. Introduction 1

General Issue 1

Problem Statement 3

Research Objectives 4

Research Focus 4

Investigative Questions 4

Methodology 5

Assumptions/Limitations 5

Implications 6

Preview 7

II. Literature Review 8

Chapter Overview 8

Initial Skills and On-the-Job Training 8

Traditional Training 9

Virtual Reality 11

Augmented Reality 13

Training Methods Overview 14

Platform Advantages 15

Platform Disadvantages 18

Trainee Productivity and Trainee Efficiency 21

Training Costs 23

Cost-Effectiveness Analysis of Training 24

Value-Focused Thinking 26

VFT for Training Programs 28

Training Taxonomies 34

Taxonomy Application 38

MPEET Model 39

Summary 45

III. Methodology 47

Interview Methods 47

Approach 48

Interviewee Details 50

Interview Content 51

Data 54

Implementation 59

Assumptions 61

Data Analysis 62

Summary 63

IV. Analysis and Results 64

Chapter Overview 64

Problem Identification 64

Create Objectives Hierarchy....64

Develop Evaluation Measures....69

Objectives Hierarchy Weights....73

Create Value Functions....82

Alternative Generation....97

Tool Formulation....101

Summary....117

V. Conclusions and Recommendations....118

Chapter Overview....118

Conclusions....118

Significance of Research....124

Limitations and Future Research....125

Summary....126

Appendix A: Interview Handout....127

Bibliography....141

List of Figures

Figure 1 AF Aircraft Maintainer Loss Rates (GAO, 2019) 1

Figure 2 Virtual Reality Continuum (Milgram & Kishino, 1994) 11

Figure 3 Effectiveness by Years of Service (Manacapilli, Bailey, Beighley, Bower, & Bennett, 2007) 22

Figure 4 Generalized CEAT Model (Simpson H. , 1995) 25

Figure 5 Value Focused Thinking Framework adapted from (Shoviak, 2001) 28

Figure 6 Sample Objectives Hierarchy 30

Figure 7 Direct vs. Proportional Scoring Example 31

Figure 8 Training Taxonomy Summary 38

Figure 9 MPEET Model (Brown, 2014) 40

Figure 10 VFT Inputs and Outputs 48

Figure 11 Interview Block Definition Diagram 52

Figure 12 Psychomotor Value Function Example 58

Figure 13 Objectives Hierarchy and Problems 68

Figure 14 Histogram of Attribute Importance Ratings 75

Figure 15 Question 10 Results - Boxplot of Attribute Importance Ratings 76

Figure 16 Question 11 Results - Boxplot of Cost Type Rankings 82

Figure 17 Question 14 Results - Cost Value Function 83

Figure 18 Cognitive Value Function - Module 1 84

Figure 19 Cognitive Value Function - Module 2 85

Figure 20 Cognitive Value Function - Module 3 86

Figure 21 Affective Value Function - Module 1 87

Figure 22 Affective Value Function - Module 2 88

Figure 23 Affective Value Function - Module 3 89

Figure 24 Psychomotor Value Function - Module 1 90

Figure 25 Psychomotor Value Function - Module 2 90

Figure 26 Psychomotor Value Function - Module 3 91

Figure 27 Histogram of High Competency Level Responses.. 93

Figure 28 Histogram of Objective Competency Level Responses 94

Figure 29 Step 1 Problem BDD Example.. 104

Figure 30 Step 2 Objectives Hierarchy BDD Example 104

Figure 31 Step 2 Objectives Traced to Problems BDD Example.................................. 105

Figure 32 Step 2 Completed Objectives Traced to Problems Example.......................... 105

Figure 33 Step 3 Attributes BDD Example 106

Figure 34 Value Function Example 111

Figure 35 Utility Visualization Example 113

Figure 36 Total Value Function BDD 119

List of Tables

Page

Table 1 Training Methods 14

Table 2 Platform Advantages 18

Table 3 Platform Disadvantages 20

Table 4 Costs of Training (Decker & Campbell, 1996) 23

Table 5 Orlansky's Cost-Effectiveness Decision Diagram (Simpson H. , 1995) 26

Table 6 Weighting Example 32

Table 7 The Cognitive Domain (Iowa State University, 2012) 35

Table 8 The Affective Domain (O'Neill & Murphy, 2010) 36

Table 9 The Psychomotor Domain (International Civil Aviation Organization, 2016) ... 36

Table 10 The Interpersonal Domain (International Civil Aviation Organization, 2016) . 37

Table 11 Instructional Strategy Selection Chart adapted from Clark (Clark, 1999) 39

Table 12 Training Compatibility Matrix Example 41

Table 13 Value Attribute Sources 49

Table 14 Training Methods and Technologies 54

Table 15 Question 1 Results - Problem Identification 65

Table 16 Question 2 Results - Objectives 66

Table 17 Question 3 & 4 Results – Training Attributes 70

Table 18 Question 5 & 6 Results - Training Attributes 71

Table 19 Question 7 Results - Decision-Making Tools 72

Table 20 Question 8 Results - Additional Attributes 73

Table 21 Question 9 Results - Attribute Importance (Most-Least) 73

Table 22 Question 10 Results - Attribute Ratings 74
Table 23 Summary Statistics for Attribute Importance Ratings 75
Table 24 Question 10 Results – Greenhouse-Geisser 77
Table 25 Attribute Pairwise Comparisons 78
Table 26 Confidence Intervals for Attributes 79
Table 27 Question 10 Results – Shapiro-Wilk Normality Test 80
Table 28 Question 11 Results - Cost Type Rankings 81
Table 29 Question 12 Results - Unimportant Cost Types 81
Table 30 Question 13 Results - Utility Ratings of Cost 82
Table 31 Module Goals with Domain Objectives 83
Table 32 Cognitive Levels of Competency 84
Table 33 Affective Levels of Competency 86
Table 34 Psychomotor Levels of Competency 89
Table 35 Descriptive Statistics for Objective and High Competency Levels 93
Table 36 Summary Statistics for High and Objective Values 94
Table 37 Frequency of Differences 95
Table 38 Sign Test Results for Value Difference 95
Table 39 Question 16 Results - Value Function Reasoning 96
Table 40 Question 17 Results - Interactive Domain Functions 97
Table 41 Question 18 Results - Additional Methods and Technologies 98
Table 42 Question 19 Results - Possible Training Technologies for Use with Methods . 99
Table 43 Question 19 Results - Best Combinations of Technologies Per Method 99
Table 44 Question 20 Results – Mean Technology Value Per Attribute 101

Table 45 Example Problem 103

Table 46 Step 3 Evaluation Measures 107

Table 47 Attribute Utility Functions 109

Table 48 Objective Competency Level Example 110

Table 49 Attribute Weights Example 111

Table 50 Best Training Technology List 112

Table 51 Training System Alternatives Example 114

Table 52 Attribute Inputs 116

Table 53 Alternative Scoring Example 116

EXPLORING THE VALUE OF TRAINING SYSTEM ALTERNATIVES THROUGH METHODS, TECHNOLOGIES, AND RELEVANT ATTRIBUTES

I. Introduction

General Issue

Maintainers within the Air Force are essential to flight operations, ensuring the readiness and agility of the warfighter through aircraft repair and scheduled maintenance. A 2019 U.S. Government Accountability Office (GAO) report asserts that "aircraft maintenance is the Air Force's largest enlisted career field, accounting for about a quarter of its active duty enlisted personnel" (GAO, 2019). The same study also highlights the increased loss rates of 5- and 7-level maintainers, those individuals who have greater levels of on-the-job experience and expertise. These experts are lost at a rate that is above the average of enlisted attrition.

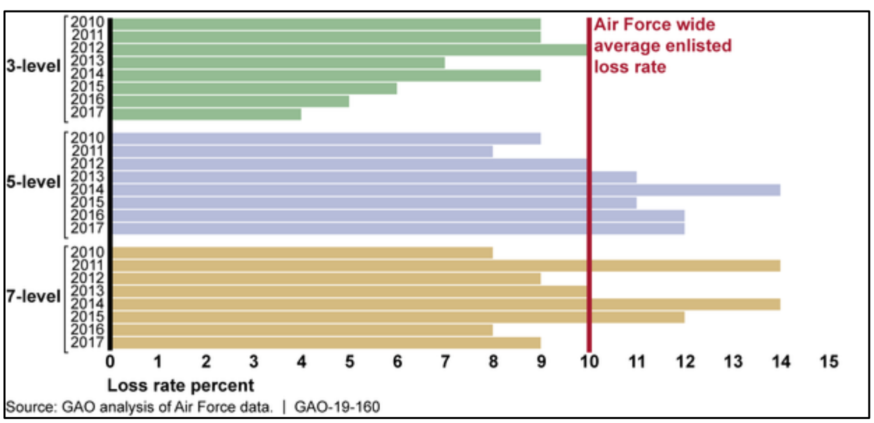


Figure 1 AF Aircraft Maintainer Loss Rates (GAO, 2019)

is retained to achieve desired system availability rates despite the higher rate of turnover in experts in the maintenance career field?

As an alternative to relying on retainment of experts, another solution may lie in rapid development of novices into experts. Documentation such as the Career Field Education and Training plan (CFETP) indicate that there is some room for improvement. Throughout several of the CFETP's within the maintenance field, there are metrics such as "Average Sew-on" and "Earliest Sew-on", indicating the average and earliest amount of time in service to attain the corresponding grade. These metrics are indicative of the variability of the time it takes to train an airman. Throughout multiple CFETP's in the aircraft maintenance career field, an upgrade to a 7-Level Craftsman has an earliest sew-on of 3 years and an average sew-on of 7.5 years (DAF, CFETP 2A3X3B, 1998). This indicates that there is a 4.5-year gap between the most proficient and average performers in the maintenance field. It is plausible that by implementing more effective training programs, the average airman can attain the requisite skills to advance to a higher grade more quickly to ensure that there is a higher number of qualified maintainers available to fill losses.

Training can take various forms, such as formal classroom education, traditional computer-based training, real-world simulation, sifting through manufacturers' technical data, phone calls to a Subject Matter Expert (SME), or having the physical presence of an expert with the apprentice technician as he or she performs a maintenance action. New technologies such as Augmented Reality (AR), Virtual Reality (VR), and other Mixed Reality (MR) platforms have the potential to change the way the Air Force does training, enabling the user to learn by "doing" rather than simply viewing, as well as enabling

virtual connection to SME support. These platforms fall under the umbrella term of extended reality (XR). Multiple career fields across the Air Force have already created and implemented XR content within their training curricula accordingly (Goldstein, 2020). In particular, AR and VR training platforms have been promoted as cost-effective means of increasing training effectiveness, but a concrete way of predicting their utility on a case-by-case basis remains elusive. The reality faced by many training organizations "is that the military currently lacks a true cost benefit model accounting for all variables. This is a common issue amongst several professional fields, not just the military" (Lyndall III, 2020).

Problem Statement

The greatest advantages and drawbacks of each training platform have yet to be identified and formed into a decision-making tool. Even though there are many options for delivering training, it is still the responsibility of the training provider to choose the form and method of training that is most beneficial given the training provider's available resources. In choosing a training method, it is necessary to justify investments in compatible training technologies. Traditional methods of training can require constant SME attention or additional hardware replicas meant solely for instructional purposes. AR/VR technologies bring additional content delivery requirements and sustainment considerations in addition to the initial cost of acquiring the necessary hardware. Faced with a plethora of ways to train today's Airmen, training units need to be familiar with the aspects of various training platforms to select the method and technology combination

that will result in the greatest benefit to the desired skill set of the trainee while balancing resource requirements.

Research Objectives

This research seeks to discover which attributes contribute to the overall value of a training alternative, to demonstrate that this value is dependent upon the desired level of trainee competency, and to find which training method-technology combination has the highest value, enabling trainers to strategically develop training solutions. The goal of this effort is to create a tool from these findings that will aid training organizations in selecting the training solution that will best satisfy its requirements.

Research Focus

The focus of this research shall be on training programs for Air Force maintainers. The current study will focus on the types of proficiencies that are enabled by various combinations of training methods and technologies. The Value-Focused Thinking (VFT) framework is used as a guide to traditional top-down systems engineering, capturing the real problems, associated objectives, relevant evaluation measures, and generating value functions for training attributes. Together, the resulting tool will inform the creation of prospective training programs and aid individuals when recommending an accompanying training solution.

Investigative Questions

To satisfy the objective of this exploratory research, it would be helpful to start with asking, how does the Air Force currently deliver training to its technicians? In addition to training delivery, there is also the aspect of, what constitutes training

effectiveness? Next, what do each of these training platforms provide to its trainees in terms of skill or experience? Also, what experience can be gained by leveraging relatively new forms of training such as those enabled by AR/VR technologies? Lastly, what costs or resources are typically involved with each training method? This examination of utility and resource commitment will provide a vision for how the Air Force should train its Airmen, the proficiencies which can be gained by each training aid, and how to build a model to decide between training alternatives.

Methodology

Since this research is exploratory in nature, the methodology shall consist of a literature review followed by semi-structured interviews with trainers administered via video conference. The structure of interview questions follows the first four steps of the VFT process: problem identification, creating an objectives hierarchy, developing evaluation measures, and creating value functions (Shoviak, 2001). The intent of this methodology is to utilize this information to discover the most valuable attributes of training and to be able to predict the cost-effectiveness of training alternatives. From there, a decision tool can be created that is compatible with the VFT framework.

Assumptions/Limitations

It is assumed that training programs enlist a combination of training media and methods to teach a single skill. Thus, it is important to characterize the components of a skill and relate the contribution of each training medium accordingly. Some limitations of this study include technology bias and small sample sizes. Technology bias may affect the attitude and opinions of interview respondents in a way that is adversarial to

implementing new technology in training. Conversely, proponents of AR/VR technologies in training may be biased toward implementing these aids where it may not be appropriate, with the sole justification of novelty or general appeal. Additionally, the valued attributes of training are specific to the field of maintenance, which involves a unique combination of skills; therefore, the opinions gathered during the interview phase may only be generalizable to maintenance-specific skills. This study is being performed under time constraints as it is a requirement for graduation. These time limitations will affect the number of samples that can be generated from interviews. Thus, this research effort is designed to focus on generating a roadmap to be augmented someday by future research.

Implications

Implications to the Air Force include more informed decision-making, more effective training, and the way the creation of training alternatives is executed. It is expected that not all types of training objectives will be satisfied by the implementation of AR/VR technologies. Likewise, there may be some skills that are very difficult or cost prohibitive to teach to the necessary degree of proficiency without utilizing AR/VR technology. In this respect, it is not the goal of this research effort to promote one type of training over another, but to build a utility with which organizations can weigh the advantages and drawbacks of each training method-technology combination. The end goal is to promote the creation of more effective training to develop today's Airmen, filling the gaps left behind by attrition of experienced technicians. The current mission of the Air Force is already being accomplished through advanced technologies that were

foreign to its users just a few years ago. As technologies like AR and VR are being implemented and becoming widely accepted, the Air Force needs to have an agile stance on adapting new technologies for meeting objectives and extending capabilities. The creation of a utility is important because it provides a point of interest through which both trainers and management can communicate. Decision-making aids promote agility through common understanding, allowing the user to make better decisions faster.

Preview

This research effort is composed of four additional chapters. Chapter 2 covers a review of literature which provides information on comparisons of AR, VR, and traditional training platforms, trainee productivity, Value-Focused Thinking, and methods for evaluating cost-effectiveness of training alternatives. Chapter 3 describes the methodology of the research, which includes semi-structured interviews, followed by the results in Chapter 4. Conclusions and recommendations are presented in Chapter 5.

II. Literature Review

Chapter Overview

This chapter reviews past findings on training platforms and metrics of evaluation relevant to this research effort. The review includes arguments for and against each training platform, definitions of training effectiveness, an estimation of productivity acquired by Air Force maintainers during initial skills training, and methods for estimating the cost-effectiveness of training. The literature review reveals a lack of decision-making tools for integrating AR/VR training platforms into professional training. Prior research has focused on qualitative and experimental methods to determine differences in trainee outcomes with alternate training platforms. To describe alternative structured approaches to training platform selection, the chapter includes a discussion of Value-Focused Thinking. A review of learning taxonomies is included to describe desired trainee competency levels when evaluating training alternatives. The chapter concludes with a summary of the most relevant decision-making tool discovered during the review, a Methodology to Predict and Evaluate the Effectiveness of Training. Combining aspects of this methodology with Value-Focused Thinking may be useful for building a tailorable decision model.

Initial Skills and On-the-Job Training

To illustrate a setting that is relevant to maintainers, it is important to understand two generalized stages of training in which the possible platforms are employed: initial skills training (IST) and on-the-job training (OJT). IST is the first phase wherein basic skills are learned for a specific career field. For aircraft maintainers, IST is made up of

formalized “schoolhouse” training, including aircraft-specific and active flight line (“hot” location) training. During IST, Airmen achieve a 3-level skill set, making them mission ready and able to perform tasks unsupervised. As stated in the Career Field Education Training Plan (CFETP) for F-16’s and F-117’s, “A task certified apprentice means the individual can complete the task utilizing tech data, but may not meet local standards for speed” (DAF, CFETP 2A3X3B, 1998). The CFETP lists all tasks in which an airman must become proficient. Formal IST is administered by dedicated trainers.

OJT is not as well-defined as IST because the technician is continuously learning new skills by being subjected to a variety of tasks. The CFETP defines OJT as “Hands-on, over-the-shoulder training at the duty location used to certify personnel for both skill level upgrade and duty position qualification” (DAF, CFETP 2A3X3B, 1998). It requires supervision by a more experienced, more proficient technician who is trained on a particular task. Although the trainee’s skills are being expanded by the supervisor, the supervisor forfeits the benefit of productivity. To put it simply, the supervisors are not completing tasks they are highly skilled at while they train the apprentice technician. Therefore, the organization loses the benefit of having a skilled worker while OJT is pursued.

Traditional Training

This research effort includes a review of traditional training, AR training, and VR training platforms. Traditional training can take many forms and can be administered almost anywhere, including classrooms, job sites, or from the trainee’s home. It is non-specific with regard to the materials required for training, allowing the instructor flexible

technology options compared to AR/VR platforms. Due to the hands-on nature of the maintenance career field and the intricacies of the physical objects that maintainers interact with, visual aids (e.g., diagrams or power point images), technical data (e.g., manuals), or physical artifacts typically accompany a lesson. Traditional training also includes online learning applications. Although some have referred to online learning as the virtual classroom, the word "virtual" does not carry the same connotation as in the term "virtual reality." In traditional training, an instructor is least restricted by the methods he or she may utilize to accomplish a training objective. For example, a lecture requires no materials, just the proper level of knowledge and the ability to speak. As a result, the only direct costs associated with this form of training are the cost of the trainer and the cost of material preparation. However, for a more costly solution, a physical artifact may be brought into the training environment, such as a model or a fully functioning system, designated for training purposes. For this option, the training organization needs to consider the cost of the physical artifact (e.g., a component or small system) in addition to the salary of the trainer and material preparation. Having an accessible system model (e.g., an aircraft) for training purposes is very costly to an organization. Thus, other means of training, such as computer-based or virtual programs, can provide the ability to interact with a system without the cost associated with the purchase of a new system (Bartley & Golek, 2004).

In the virtual classroom, a wide variety of educational diagrams and video content can be viewed by as many people as possible. However, web-based education often removes the benefit of having in-person cues which are present in class or work settings (Kraiger, 2008). Because these social cues are removed and a larger class size is possible,

the trainer may have difficulty pursuing adaptive learning, wherein a trainer assesses and adjusts the training strategy in real time to compensate for student learning style, allowing for greater training effectiveness. This example helps to illustrate some tradeoffs present in pursuing a traditional training platform. Even within the category of traditional training, the individual methods that are employed carry cost and performance tradeoffs. Thus, it is important that trainers construct a curriculum with these tradeoffs in mind to pursue a training solution that will be the best value to the organization. To add even greater complexity, modern training organizations will need to weigh the factors associated with adopting AR/VR technologies.

Virtual Reality

As opposed to traditional training where events are experienced in the physical world, VR exists on the other side of the virtual spectrum (Milgram & Kishino, 1994).

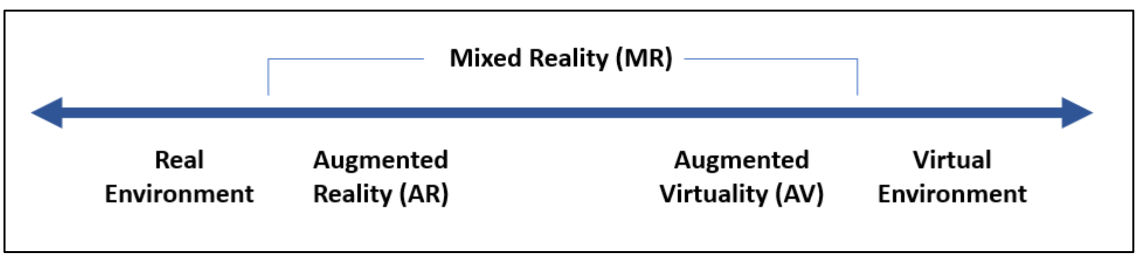


Figure 2 Virtual Reality Continuum (Milgram & Kishino, 1994)

VR is exhibited when the trainee experiences and interacts with an environment in which all objects are virtual, as "VR technology generally uses a headset, blocking out visual stimulus from the real world" in which the person resides (Kaplan, et al., 2021). A typical apparatus for VR training includes a VR headset and hand controls. The hand controls allow the user to interact with objects in the virtual world. The virtual world allows

objects to be manipulated or taken apart with ease and allows the behavior of the physical world (e.g., gravitational force) to be adjusted to the trainer's preferences. These displays may provide different views to the user's two eyes, providing stereoscopic views of objects, allowing the trainees to perceive objects as three-dimensional (VRTL Academy, 2018). Further, as the field of view of the display is controlled in a VR headset, the perceived size of objects in the virtual reality environment can be controlled to mimic the perceived size of objects in the real world. Therefore, objects may appear as they do in the real world while having the added advantage of augmentation such as a colorful halo which surrounds components of interest. A trainer in the physical world may monitor the movements and behaviors of the trainee by viewing a projection of what the trainee sees. However, interaction with a physical trainer can be limited as the trainee cannot view objects in the real world. To compensate for the lack of communication with a physical trainer, a virtual avatar may be created to guide the student, either allowing the trainer to be represented in the virtual environment or providing a representation of an artificial avatar which participates in the training. Virtual avatars have been proposed as a cost-saving measure as they may potentially replace a trainer, at least in some conditions, or to provide customized 1:1 training if the avatar is able to suit the specific needs of different learning styles (Rupp, Gibbons, & Snyder, 2008). Although human-human teaming in VR has been implemented for some tasks, it is not a prevalent feature. Many can watch what is happening in the virtual environment, but only one user can use one VR headset at a time. VR hardware is not as abundant as traditional training materials, particularly in current maintenance training settings. There are instances of VR where training can be distributed to both VR headsets and hardware such as desktop computers. In this case,

trainees still interact with the virtual content, even though they may lose the stereoscopic view of the objects and some of the tactile benefits.

Augmented Reality

Augmented reality (AR) lies on a virtuality spectrum between reality and VR (Milgram & Kishino, 1994). The user can see the real world, interact with real objects, and interact with other humans in a real physical environment. A user typically wears a headset or may view objects through a tablet computer or cellphone which hosts the AR visualization and virtual information. There are multiple forms of trainee-accessible hardware which can host AR applications; thus, AR is generally more accessible to training organizations than VR. As for the benefits of AR, it was best said by Webel et al.:

> The main advantage of using Augmented Reality for training is that the trainee can interact with real world objects and simultaneously access the virtual information for guidance. Therefore, the trainee can easily create a mapping between the training and the real-world task. Additionally, the trainee can perform the actual task while accessing additional training material (Webel, et al., 2013).

With appropriate hardware and AR software, these systems can overlay virtual objects onto real objects, aligning objects in the virtual and real world in near real-time and creating the appearance that the virtual object is in the real world. Because the AR user can interact with objects in the real world, productivity is achievable. Some industrial settings in the private sector have adopted the use of AR to guide workers through a variety of tasks and procedures to produce value by enhancing productivity. These systems may rely either on automated presentation of the AR information or enable sharing of the real-world environment with experts to provide personalized

interaction during task performance. The main cost tradeoff that exists with both VR and AR technologies is that it generally costs more money to produce more virtual content or content of a higher fidelity (Day, 2020; Fade, 2019).

Training Methods Overview

Traditional training, VR training, and AR training are umbrella terms which characterize the delivery media of training. These are the vessels through which one to many training methods may be administered simultaneously. The International Civil Aviation Organization (ICAO) published a "Taxonomy to Assist in the Identification of Instructional Methods." It lists and defines learning behaviors, instructional methods, and factors associated with different learning environments (International Civil Aviation Organization, 2016). The instructional methods in Table 1 can be combined and administered with E-learning methods to create blended training.

Table 1 Training Methods

Training Methods	
Instructional Methods	**Categories of E-Learning**
Lecture	Delivery over time
Drill and Practice	Information Sharing
Demonstration	Knowledge building
Discussion	Skill development
Interactive Instruction	Individual or group learning
Skill Development and Integration	Low/high technology
Case Studies	Games
Simulation	Synchronous/Asynchronous
Role Play	Self-Instruction
Games	Virtual Classroom
	Virtual Worlds and Simulations
	Webinars
	Mobile Learning
	Levels of Achievement

One can observe that these methods may be administered through traditional, AR, and VR training platforms. For example, if an AR headset guides a user through a table assembly task, the AR headset could provide a simulation of a finished table (simulation, high technology, information sharing, individual), a video on how to position a partially-constructed table to assemble components (demonstration, knowledge building), and trainee-activated overlays with a virtual instructor commanding the trainee through the process (simulation, self-instruction, demonstration, skill development and integration, asynchronous). The same table assembly task can also be trained using traditional training platforms, although the setting, context, number of trainees, physical trainer, etc. may differ. It is the responsibility of training organizations to determine the platform and methods which are to be used to accomplish training objectives. These combinations of training method and training technologies are created by considering the advantages and disadvantages associated with each training system alternative.

Platform Advantages

To determine if a strong advantage of one training platform over the other exists, the benefits of each training platform were researched. Benefits of traditional training include the opportunity for peer-to-peer learning, the presence of nonverbal cues, and the ability of an instructor to control the complexity of content delivery to facilitate adaptive learning (Kraiger, 2008). The presence of a competent instructor who may informally assess the skill level of trainees can adjust the way information is presented in real time to achieve learning objectives. Although the focus of his research was on primary and secondary education, Blatchford suggests that smaller class sizes can allow instructors to

individualize instruction to a greater extent (Blatchford, Bassett, & Brown, 2011). Thus, there is an inherent balance that must be optimized to leverage both class size and individualized instruction. Gavish notes that additional training time is necessary for familiarization with XR platforms, thus some traditional training methods boast an immediate usability over XR platforms (Gavish, et al., 2015). Unless the use of XR platforms is sufficiently frequent, time spent training its use could be a deciding factor in cost comparisons.

AR and VR platforms carry many shared benefits. Kaplan et al. and Webel et al. present many of these shared benefits, including the ability to:

- replicate dangerous or cost-prohibitive scenarios
- create scenarios that have not yet been encountered
- update training quickly and efficiently
- free up physical equipment for use
- provide immediate feedback on a task
- increase the speed of performing maintenance tasks
- allow effective transfer of skills from a virtual environment to the physical environment (encoding specificity) (Kaplan, et al., 2021)
- remove the need for a physical instructor (Webel, et al., 2013).

Although some AR and VR solutions apply similar hardware, each platform has distinct advantages. Gavish and colleagues carried out an experiment that evaluated the performance of individuals as they performed a 25-step industrial maintenance assembly task on various training platforms. This research discovered that AR scored significantly

better than VR in 6 of 9 metrics. Specifically, AR resulted in lower practice time, greater ease of performing task, higher efficiency, quicker error recovery, comfortability of system experience, and ease of use (Gavish, et al., 2015). Gavish and colleagues also noted that participants had a lower first-time unsolved error rate when using AR than when following a traditional paper procedure (Gavish, et al., 2015). Webel asserts that the greatest advantage of AR is the ability to affect physical objects while having access to technical information, ensuring both tactile feedback and access to topical knowledge (Webel, et al., 2013). In a meta-analysis of AR in educational settings, learning gains and motivation were the most reported benefits of AR (Garzon, Pavon, & Baldiris, 2019). Table 2 summarizes the advantages presented in this section.

Table 2 Platform Advantages

Traditional	AR	VR	Source
Ability to train large class sizes			Bartley & Golek, 2004
More opportunities for peer-to-peer learning			Kraiger, 2008
Presence of nonverbal cues			Kraiger, 2008
Ability to pursue adaptive learning through trainer-metered content complexity			Kraiger, 2008
Individualized Instruction			Blatchford et al, 2011
No need to train for platform use			Gavish et al, 2015
	Easily replicate dangerous or cost-prohibitive scenarios	Easily replicate dangerous or cost-prohibitive scenarios	Kaplan et al, 2021
	Create new scenarios	Create new scenarios	Kaplan et al, 2021
	Quickly and efficiently update training	Quickly and efficiently update training	Kaplan et al, 2021
	Free up physical equipment for use	Free up physical equipment for use	Kaplan et al, 2021
	Effective transfer of skills from a virtual environment to the physical environment (encoding specificity)	Effective transfer of skills from a virtual environment to the physical environment (encoding specificity)	Kaplan et al, 2021
	Increased speed of performing maintenance tasks	Increased speed of performing maintenance tasks	Ganier et al, 2014
	Provide immediate feedback without instructor present	Provide immediate feedback without instructor present	Haque & Srinivasan, 2006
	Not necessary to have an instructor available	Not necessary to have an instructor available	Webel et al, 2013
	Lower platform practice time than VR		Gavish et al, 2015
	Greater ease of performing task than VR		Gavish et al, 2015
	Greater efficiency of task performance than VR		Gavish et al, 2015
	Quicker recovery from mistakes than VR		Gavish et al, 2015
	Greater comfortability of system experience than VR		Gavish et al, 2015
	Easier to use than VR		Gavish et al, 2015
	Lower first-time unsolved error rates than VR		Gavish et al, 2015
	Ability to affect physical objects while having accessibility to technical information		Gavish et al, 2015
	Higher learning gains		Garzon et al, 2019
	Greater Motivation		Garzon et al, 2019

Platform Disadvantages

Keesling presents inadequacies of both traditional training and AR training platforms obtained from the maintenance community, AR solution providers, and AR users (Keesling, 2019). A common drawback to traditional training identified by these groups includes information delivery that is inadequate for training purposes, requiring additional input from experts to provide context to diagrams within manuals and procedures. Similar results are presented in a study comparing performance of a

maintenance assembly task after being trained by traditional and AR methods. Webel finds that traditional documentation often fails to demonstrate pertinent skills and knowledge a technician requires to complete a task (Webel, et al., 2013). Keesling also presents multiple challenges to AR implementation, including content creation, hardware, network infrastructure, and organizational challenges (Keesling, 2019). All participants held the belief that the proper place for AR was in the training environment, however it is unclear whether the definition of a training environment also includes OJT. Since training may be administered during IST and OJT, it is assumed that AR can be used for both settings. In addition, the number of students that can be trained using AR or VR is limited by the number of headsets that can be used at one time. Gavish indicates that there is an additional training period for AR and VR that is not present for the use of traditional training materials, stemming from new interactions while learning to use the hardware efficiently (Gavish, et al., 2015). In addition, AR developers must develop curriculum which can incrementally reduce the trainee's dependency on the visual features it provides, encouraging the trainee to learn rather than rely on reference materials (Webel, et al., 2013).

Drawbacks to VR training include the inability to use the platform in an OJT context, the necessity of developing everything within the virtual environment, and the lack of interaction with a physical trainer. Many of the same drawbacks experienced by AR users can be extrapolated to VR users due to the similarities of the hardware required for each solution. "The variability in visual quality of different XR products, lag and tracking problems, and the potential for simulator sickness are all sources of limitation

that may diminish training efficacy" (Kaplan, et al., 2021). Table 3 summarizes the disadvantages presented in this section.

Table 3 Platform Disadvantages

Traditional	AR	VR	Source
Inadequate information delivery			Keesling, 2019
Manuals fail to properly demonstrate skills and knowledge for a technician to complete a task			Keesling, 2019
	Content creation	Content creation	Keesling, 2019
	Hardware challenges	Hardware challenges	Keesling, 2019
	Network infrastructure challenges	Network infrastructure challenges	Keesling, 2019
	Organizational acceptance	Organizational acceptance	Keesling, 2019
	Environmental suitability	Environmental suitability	Keesling, 2019
	Inability to train large class size simultaneously	Inability to train large class size simultaneously	Keesling, 2019
	Longer platform training period	Longer platform training period	Gavish et al, 2015
	Need to incrementally reduce trainee dependency on visual features	Need to incrementally reduce trainee dependency on visual features	Webel et al, 2013
	Variability in display visual quality	Variability in display visual quality	Kaplan et al, 2021
	Lag problems	Lag problems	Kaplan et al, 2021
	Tracking issues	Tracking issues	Kaplan et al, 2021
	Simulator sickness	Simulator sickness	Kaplan et al, 2021
		Using VR to replace video lessons or other types of instructional media may not yet be useful	Parong & Mayer, 2021
		Inability to use platform in OJT setting	
		Development of virtual content	
		Lack of interaction with a physical trainer	

Because existing research indicates a wide array of tasks performed with, as well as benefits and drawbacks of each platform, it is difficult to provide a generalized recommendation of one platform over another for a given a task to be trained. A common thread is that research which compared traditional to AR/VR methods was highly task-specific, sometimes with conflicting outcomes. Such was the case with encoding specificity, or the idea that virtual acquisition of skill can transfer reasonably well to physical productivity. The National Commission on Military Aviation Safety (NCMAS)

cited the lack of physical hands-on training as having a negative effect on skill retention (National Commission on Military Aviation Safety, 2020). This finding indicates the need for determining when to implement various training technologies associated with each platform.

Trainee Productivity and Trainee Efficiency

Before going further, it is important to define terms like “productivity” and “training efficiency.” For the purposes of this research, productivity is defined as an individual’s rate of completing maintenance activities, often measured in units such as output of services per hour of labor. Training efficiency is defined as how well or how fast one can become proficient to perform tasks without assistance. An increase in training efficiency occurs when an individual can obtain a desired level of proficiency in less time or obtain a greater level of proficiency in comparable time (Kraiger, 2008). In other words, greater efficiency relates to a higher ratio of proficiency over time. An excerpt from a 2007 RAND study states that training experts “defined a fully mission effective person as 100-percent productive when the individual was considered a “go-to” person by the leadership” (Manacapilli, Bailey, Beighley, Bower, & Bennett, 2007). This is a necessarily loose definition due to the breadth and depth of knowledge required to perform the job of an aircraft maintenance technician. This research also modeled the acquisition of productivity throughout the IST and OJT phases with the goal of finding a more efficient balance of time spent in formalized training with respect to the time spent in on-the-job training (Manacapilli, Bailey, Beighley, Bower, & Bennett, 2007). Ninety percent of survey responses from the RAND study estimate the productivity of a special

vehicle maintenance airman at the end of IST to be about 10-40% (Manacapilli, Bailey, Beighley, Bower, & Bennett, 2007). Surveys of subject matter experts (SME's) indicated that the mean time to 100% proficiency ("fully mission-effective") for two groups of aircraft maintainers was about 4.5 years. Figure 3 illustrates the mission effectiveness for special vehicle maintainers versus years of service (YOS).

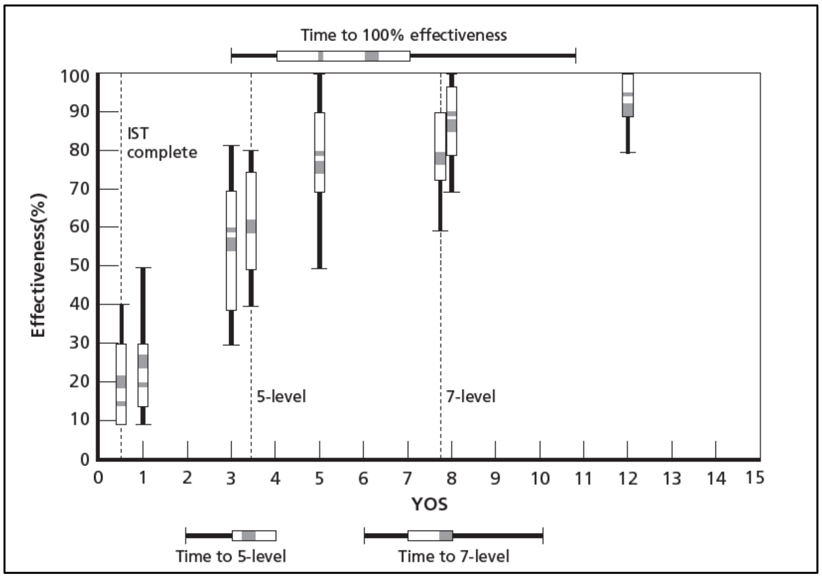


Figure 3 Effectiveness by Years of Service (Manacapilli, Bailey, Beighley, Bower, & Bennett, 2007)

This figure is representative of the variability of effectiveness versus time to attain 5- and 7-level maintainer status. Across all skill levels, maintainer effectiveness varies by a large margin. This may indicate that not all current training programs are effective for all students and that productivity is not a well-defined metric. Because productivity itself is not well-defined, even within the maintenance career field, it indicates the need for more specific measures of trainee performance and efficiency.

Training Costs

Cost is a primary metric used for the basis of organizational decisions; however, many organizations fail to pursue due diligence when calculating the true cost of training (Decker & Campbell, 1996). Both direct and indirect costs are to be considered. Direct costs include personnel, external training services, training development, instructional materials, equipment, facilities, and travel (Decker & Campbell, 1996). Indirect costs include overhead and general & administrative (G&A) (Decker & Campbell, 1996). Costs and associated examples are presented in Table 4.

Table 4 Costs of Training (Decker & Campbell, 1996)

Costs		Examples
Direct	Personnel	Wages and benefits of trainers and trainees
	External Training Services	Externally produced training materials, delivery costs
	Training Development and Instructional Material Prep	Costs of training development effort, costs of course content preparation, supplies related to prep
	Instructional Materials	Instructional materials, books, pencils, paper
	Equipment	Rented or purchased hardware, hardware maintenance
	Facilities	Rental of training facilities, facility maintenance
	Travel	Air fare, housing, per diem
Indirect	Overhead	Materials and labor not directly related to training
	General & Administrative	Everyday organizational expenses, auditing expenses, legal expenses, internet

Additionally, it is helpful to track the effects of training so curriculum-based changes can be assessed. However, this has been an issue within the aircraft maintenance community. The 2020 NCMAS report on aviation losses found that aircraft maintenance training units have not been collecting reliable data to assess skill retention before and after virtual training (National Commission on Military Aviation Safety, 2020). Without

the justification of both cost and training effectiveness metrics, programmatic changes have been perceived as unfavorable. Even if the training effectiveness metrics are in place, a higher-level system view of training reveals that stakeholder expectations matter as well (Buede & Miller, 2016). In the aircraft maintenance community, operational units believe that the VR and AR training methods implemented during IST have resulted in less-skilled schoolhouse graduates (National Commission on Military Aviation Safety, 2020). Schoolhouse units' careful consideration of the needs of the operational units are required to ensure that the introduction of new training methods reduce the transition of training from IST to on-the-job performance. Therefore, cost should not be the only driver of change.

Cost-Effectiveness Analysis of Training

Although no specific tool was discovered for relating the cost-effectiveness of XR-based training alternatives within the maintenance community, a 1995 report identifying documented cost-effectiveness analysis of training (CEAT) methods was discovered. The CEAT process follows five generalized steps, including formulating assumptions, determining alternatives, conducting Cost Analysis (CA) and Training Efficiency Analysis (TEA), comparing and selecting alternatives, and conducting a sensitivity analysis (Simpson H. , 1995). This framework can be tailored to suit the selection of training methods which incorporate AR/VR technology. Figure 4 illustrates an idealized CEAT conceptual model.

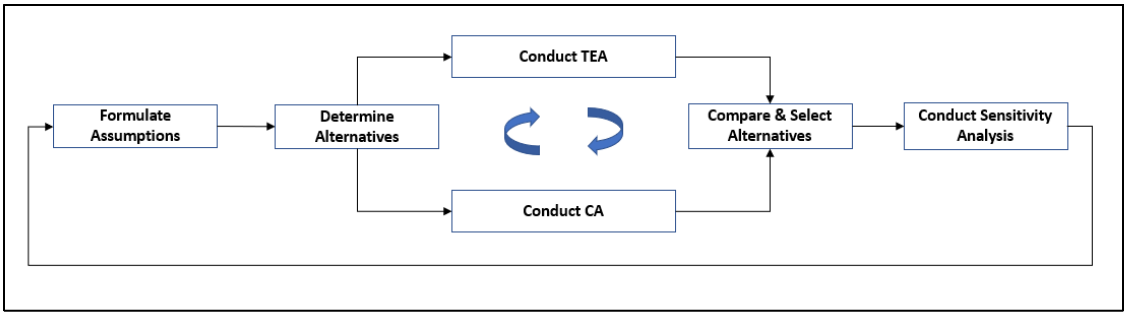


Figure 4 Generalized CEAT Model (Simpson H. , 1995)

Since AR/VR technology can be implemented in a variety of ways, the formulation of assumptions should focus on the target process and potential benefits. From there, training system alternatives can be formulated. When generating alternatives, one should consider whether the use of AR/VR enables training of a task that otherwise would be too expensive to safely replicate using traditional methods. CA and TEA are performed concurrently, although a framework particular to maintenance for each does not yet exist. The CEAT method requires a binary comparison of alternatives where each alternative is ranked against each other alternative based on cost and effectiveness as less, same, or more. When comparing and selecting alternatives, the CEAT model references Orlansky's simple decision matrix relating cost and effectiveness, presented in Table 5. As shown, a technology is adopted if it is lower in cost while delivering the same or higher effectiveness. The technology is also adopted if it is equal in cost but more effective.

Table 5 Orlansky's Cost-Effectiveness Decision Diagram (Simpson H. , 1995)

	Effectiveness		
Cost	Less	Same	More
Less	Uncertain	Adopt	Adopt
Same	Reject	Uncertain	Adopt
More	Reject	Reject	Uncertain

This matrix raises the need for clearly defined objectives of both cost and training effectiveness, as similar estimates between a new and existing solution may not provide sufficient justification to incur the cost and risk associated with a program overhaul. However, the reliability of these estimates for a new technology is likely lower than for a well understood technology. The sensitivity analysis portion of CEAT requires data to be gathered after training method implementation. The training organization should consult with outside stakeholders in operational units to gather trainee performance data to inform training system implementation. One of the drawbacks of CEAT identified by Simpson is the inability to develop a generalized cookbook to perform CEAT (Simpson H. , 1995). Perhaps by narrowing the scope of tool application to training within the aircraft maintenance community and refining its application in this constrained domain, the resulting approach may be applicable in this narrower context.

Value-Focused Thinking

A common approach to training platform selection is alternative-focused thinking (AFT). AFT consists of generating solution alternatives, then deciding between the alternatives. The problem with AFT is that it is too narrow, shutting out possible alternatives though selection of favorites and leading to objectives that are means-oriented rather than fundamental (Keeney, 2009). In systems engineering terms, AFT is a

bottom-up approach, starting with the range of alternatives rather than the objective. Conversely, value-focused thinking (VFT) is a top-down approach. "Value-focused thinking essentially consists of two activities: first deciding what you want and then figuring out how to get it" (Keeney, 2009). It solves both problems of AFT by establishing a fundamental objective and allowing a wide array of alternatives to be evaluated which may satisfy the fundamental objective. The generalized CEAT model in Figure 4 loosely follows this process by first formulating assumptions, then determining the alternatives. While the CEAT model reveals the need for both cost and training effectiveness considerations when deciding between training alternatives, it does not address the weighted value of either in the eyes of the stakeholder. For example, in Table 5, Orlansky's diagram indicates the uncertainty of a resulting training solution that has greater cost and greater effectiveness. How could the better alternative be selected if only two possible alternatives were produced to satisfy the same objectives, but they both would result in an increased cost while improving effectiveness? Without clearly established weighted value functions, this could not be accomplished, as effectiveness is only meaningful when gauged by the stakeholder's construct of the relevant training objectives. Shoviak applies the tenets of VFT into a ten-step framework as illustrated in Figure 5.

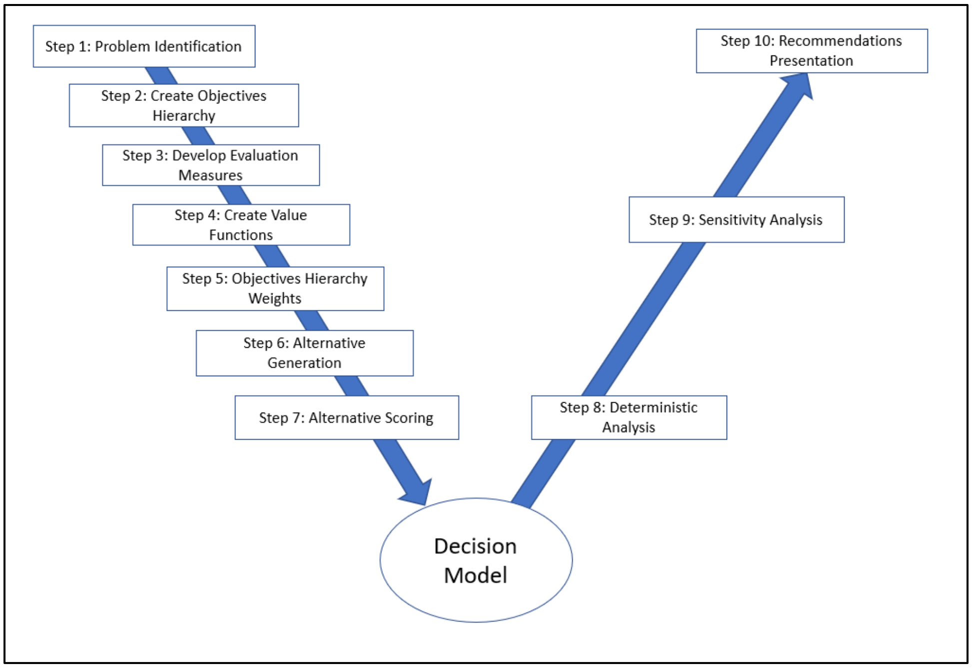


Figure 5 Value Focused Thinking Framework adapted from (Shoviak, 2001)

Shoviak's framework can help create the objectives, institute evaluation measures, and weigh objectives of each training organization to facilitate the analysis of alternatives, resulting in training implementation that is informed by stakeholder values, rather than a means-focused alternative. By utilizing this framework, useful metrics can be developed beyond the general category of proficiency, which will help organizations better estimate training effectiveness.

VFT for Training Programs

The VFT framework gives a pathway to complexity management, ensuring that an intricate problem can be parsed into manageable pieces. One accomplishes this through decomposition and hierarchy. Decomposition is the process of breaking down a problem or entity into smaller pieces and hierarchy is ordering these pieces into levels or

ranks (Cameron, Crawley, & Selva, 2016). The VFT framework starts with problem identification in Step 1. A problem is usually identified by a triggering event, such as trends in tracked metrics, results of a formal report, identification of a new threat, or even informal comments made by individuals within an organization (United States Department of Defense, 2001). The problem statement should answer the questions:

“a. Who are the personnel that are performing a task inadequately?

b. What is the exact performance problem?

c. When and where is the task performed incorrectly?” (United States Department of Defense, 2001)

After the correct problem has been identified, the VFT framework recommends creating an objectives hierarchy. The objectives hierarchy is built by considering the goals of the organization and stakeholders, then arranging them in hierarchical fashion. At the top of the hierarchy is a general objective, such as “Provide best possible training program.” One method of building a hierarchy involves simply breaking down the general objective into smaller, more specific objectives and so on, which is called decomposition. Two rules when organizing an objectives hierarchy include ensuring completeness of the higher-ordered level by defining a broad set of specific objectives and ensuring no redundancy of objectives on the same level (Shoviak, 2001). A sample objectives hierarchy is shown in Figure 6, which shows two levels of decomposition below the primary objective.

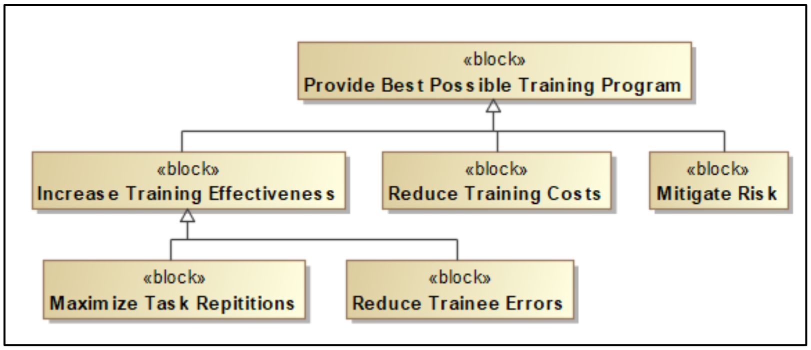


Figure 6 Sample Objectives Hierarchy

Although the decomposed objectives in this figure do not entirely capture the primary objective at the top, they do not overlap.

Step 3 of the VFT process is to develop evaluation measures. Evaluation measures consider the scale by which achievement of an objective will be measured (Shoviak, 2001). This step also considers the range of values that can be attained within each measure. This is especially important if the evaluation measure is created to measure something that is not on a common scale, such as risk or knowledge. Each measurement category shall provide a range of values for every possible input. These evaluation measures can be related or traced to an objective within the objectives hierarchy from Step 2.

Step 4 is to create the value functions to be able to score alternatives. In the direct assessment method, the form of the value function is determined by SME experience (Shoviak, 2001). In the case of proportional scoring, the output of a value function may be assumed to follow a linear trend, however this does not account for stakeholder preference (Brown, 2014). An example of direct assessment versus proportional scoring is depicted in Figure 7.

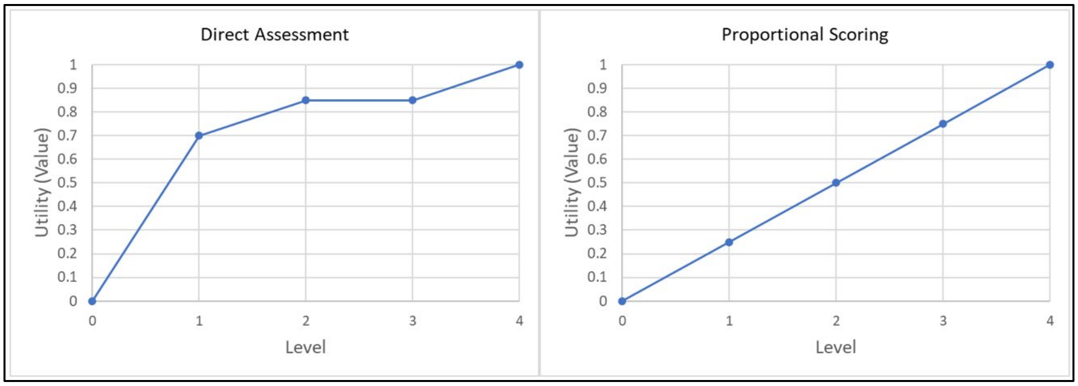


Figure 7 Direct vs. Proportional Scoring Example

A more effective predictor of a value function may be test results from actual trainee performance, but because this effort is an initial and exploratory one, it is beyond the scope of this research effort. The individual value functions, also called attributes, shall be incorporated into one total value function by normalizing each output on a scale from 0 to 1. The normalization equation is a ratio of the difference between two sets of values:

$$y_n = \left(\frac{y - y_{lowest}}{y_{highest} - y_{lowest}}\right) \tag{1}$$

Where:
y_n = normalized value output
y_{lowest} = lowest value in the range of data
$y_{highest}$ = highest value in the range of data

This nondimensional scale allows the value outputs of each function to be summed, which will be beneficial in the following step.

Step 5 is completed by assigning weights of importance to each value category (Shoviak, 2001). Continuing with the hypothetical example, the tier 2 value categories

(training effectiveness, cost, and risk) are shown in Table 6 as if the decision-maker decided to make cost five times as important as risk and training effectiveness four times as important as risk. Note that the summation of importance should be equal to 100%.

Table 6 Weighting Example

Category	Weight of Importance
Risk	x
Cost	5x
Training Effectiveness	4x
Sum	100%

Solving for x in terms of percent weight gives:

$$x + 5x + 4x = 100\%$$
$$\Rightarrow 10x = 100\%$$
$$\Rightarrow x = 10\%$$

Entering the value of x back into the table, risk makes up 10%, cost makes up 50%, and training effectiveness makes up 40% of the total value function. Each value category should relate to an objective within the objectives hierarchy. These value categories can then be broken down into smaller, independent components, called value attributes.

Step 6 includes generation of training system alternatives. The key to performing this step resides in the adherence to best practices of training system design while considering the effects of combining various training technologies to accomplish each training method. Three important best practices include breaking down complex tasks into smaller pieces, sequencing learning objectives, and sequencing training technologies. Complex tasks must be broken down into smaller segments so trainees do not get overwhelmed by either the course material or the complex technologies used during

training (Mulders, Buchner, & Kerres, 2020). This implies that both training materials and content can be managed to an optimum level to promote the satisfaction of learning objectives (LOs). Secondly, training system alternatives should be designed so that modules have proper sequencing of LOs (United States Department of Defense, 2001). A basic example is training a runner. Running is a complex movement that first requires the subject to stand, then walk, then run, then improve upon form. It would not be wise to start training advanced running techniques to someone who is learning to stand. Thus, training modules must start at the basics and advance incrementally. The third best practice is to sequence the use of training technologies. Training technologies such as slideshows, VR headsets, chalkboards, or simulators, for example, support different learning levels by enabling a unique combination of sensory stimuli for the purpose of accomplishing LOs (United States Department of Defense, 2001). For example, an aspiring pilot does not immediately train in a real aircraft. Not only would the many forms of stimuli provided by the real aircraft quickly overwhelm the trainee, but a great deal of risk is involved to both the training system and the trainee due to lack of skill. Thus, managing lesson complexity on an individual and sequential basis is assisted by the selection of proper training methods and corresponding technologies, through which LOs may be attained.

The final step before decision analysis, Step 7, is to score the alternatives. Alternatives are scored according to the overall utility each training system provides across all attributes that the decision-maker deems relevant to accomplishing training objectives. Brown illustrates a total utility function used to score alternatives (Brown, 2014):

$$Total\ Value\ Function = \sum_{i=1}^{n} w_i U_i(x_i) \quad (2)$$

Where:
n = number of included value attributes
w = weight of importance of the value attribute (%)
U = utility function
x = input for each utility function

The highest output of the total value function among the alternatives is generally considered the winner, although analysis is to be accomplished before a recommendation is made.

The analysis portion of the VFT process is beyond the scope of this research, which includes Steps 8, 9, and 10 of the VFT framework. Step 8 requires ranking the alternatives by the output of the decision model (Shoviak, 2001). By varying the weight component of the total value equation for different attributes, the decision maker can view changes to the rank order in Step 9 (Shoviak, 2001). Step 10 of the VFT process is to present the results and make recommendations to the decision maker (Shoviak, 2001).

Training Taxonomies

An important part of communicating objectives includes using verbiage that can be easily understood by stakeholders. When building, revising, or overhauling training programs, the establishment of LOs is important to express the desired outcomes specific to the trainee. These types of objectives can be traced to outcomes like productivity or acquired skills. LOs may be defined with the help of taxonomies, which are systems of classification. The ICAO synthesized taxonomy findings into four categories that are directly relevant to training objectives, including cognitive, affective, psychomotor, and

interactive domains (International Civil Aviation Organization, 2016). The cognitive domain was initially created by Bloom et al. in 1956 and later revised by Anderson & Kratwohl in 2001 to align with modern educational objectives. The cognitive domain is organized into six categories of increasing skill level related to mental skill. Remember, understand, and apply are skills that are drawn from existing knowledge while the levels of analyze, evaluate, and create imply a level of skill beyond the available existing knowledge (International Civil Aviation Organization, 2016). Table 7 categorizes skills that are within each sub-level.

Table 7 The Cognitive Domain (Iowa State University, 2012)

Level	Definition	Common Verbs
Create	Put elements together to form a coherent whole; reorganize into a new pattern or structure.	Generating, Hypothesizing, Planning, Designing, Producing, Constructing
Evaluate	Make judgments based on criteria and standards.	Checking, Coordinating, Detecting, Monitoring, Critiquing, Testing, Judging
Analyze	Break down information into component parts.	Differentiating, Organizing, Distinguishing, Selecting, Structuring Attributing, Parsing
Apply	Carry out or use a procedure in a given situation.	Executing, Carrying Out, Implementing, Using
Understand	Construct meaning from instructional messages, including oral, written, and graphic communication.	Interpreting, Exemplifying, Classifying, Summarizing, Inferring, Comparing, Explaining
Remember	Retrieve relevant knowledge from long-term memory.	Recognizing, Identifying, Recalling, Retrieving

The affective domain is also hierarchical in nature, including levels of characterizing, organizing, valuing, responding, and receiving. This domain deals with factors that affect one’s behavior, attitude, and values. Table 8 provides definitions and common verbs associated with each level (O'Neill & Murphy, 2010).

Table 8 The Affective Domain (O'Neill & Murphy, 2010)

Level	Definition	Common Verbs
Characterizing	To act consistent with the internalized values.	Discriminate, Question, Revise, Change
Organizing	To begin to harmonize internalized values.	Arrange, Combine, Compare, Balance, Theorize
Valuing	Being willing to be seen as valuing certain ideas or material.	Justify, Propose, Debate, Relinquish, Defend, Initiate
Responding	Committing to the ideas, etc. by responding to them.	Answer, Recite, Perform, Report, Select, Follow, Explore, Display
Receiving	Developing awareness of ideas and phenomena.	Ask, Follow, Reply, Accept, Prefer

The next taxonomy is the psychomotor domain, which is especially relevant to aircraft maintenance training due to the physical nature of the skills that are employed to perform a task. This domain ranges from the awareness brought about by human senses to coming up with new actions to solve a problem (Simpson E. , 1971). Table 9 provides the levels of the psychomotor domain along with their definitions and follows the taxonomy established by Simpson.

Table 9 The Psychomotor Domain (International Civil Aviation Organization, 2016)

Level	Definition
Origination	The ability to develop new actions of behavior patterns by adapting already highly developed skills.
Adaptation	The ability to modify actions to meet different or unusual requirements.
Complex Responding/ Mastery	Skillful performance that is fully integrated and automatic. Implies a high level of accuracy and proficiency.
Basic Proficiency	The ability to perform with some confidence and proficiency but without mastery.
Guided Responses	Early stages in skill acquisition involving imitation and trial and error.
Set	Readiness to act. Becoming ready to respond to different situations.
Perception (Awareness)	The ability to use sensory cues to guide motor ability, from sensory stimulation to action.

The final taxonomy considered in this research is the interpersonal domain, which includes skills necessary for various social interactions. Unlike the other domains, the interpersonal domain is not hierarchical. Its categories and definitions are represented in Table 10.

Table 10 The Interpersonal Domain (International Civil Aviation Organization, 2016)

Level	Definition
Seeking/Giving Information	Seeking or offering clarification of facts or opinions to / from individuals.
Proposing	Putting forward a new concept, suggestion or course of action that can be actioned.
Supporting	Conscious and direct declaration of support or agreement with another person or his concepts.
Including	Direct and positive attempt to involve another group member.
Summarizing	Summarizing or restating in a compact form the content of previous discussions or considerations.
Disagreeing	Conscious, direct and reasoned declaration of difference of opinion, or criticism of another person's concepts.

Figure 8 summarizes the taxonomies covered in this section in the form of ordered lists. The sublevels within each domain can be incorporated into practical statements expressing a desired knowledge, attitude, motor skill, or interaction type that will serve as the goal of training. These goals are applied to each module to communicate the expected level of achievement in each domain. A module may have more than one LO, but at this point it is not clear how to best utilize various training method and technology combinations for maximum trainee achievement.

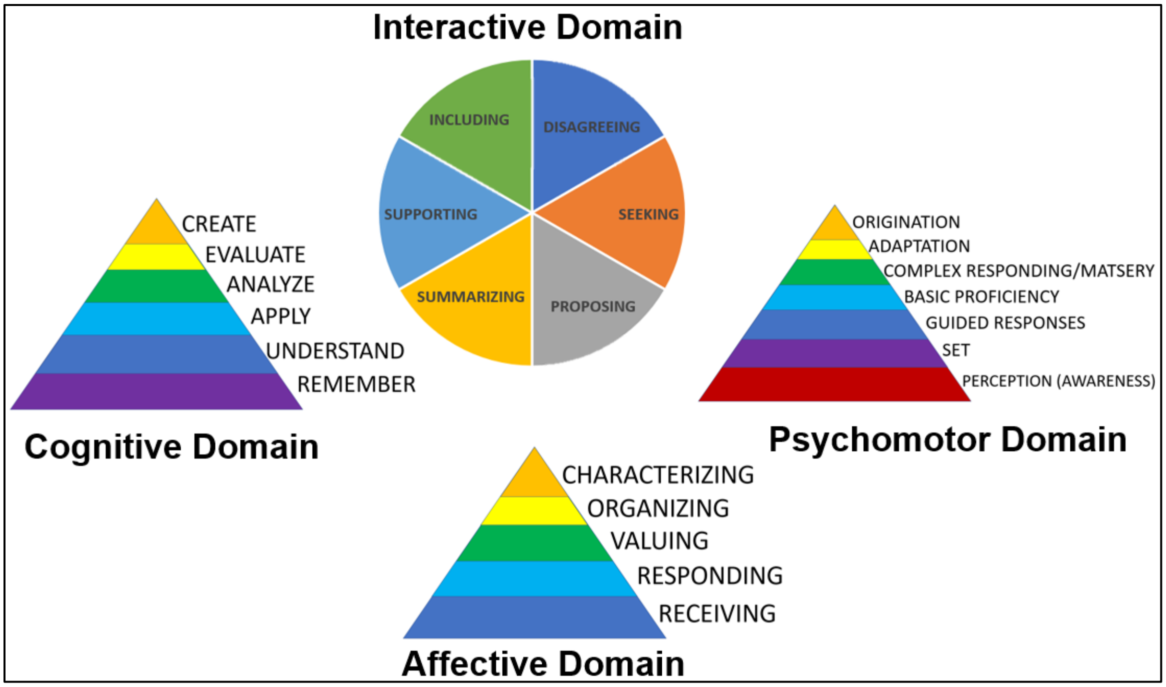


Figure 8 Training Taxonomy Summary

Taxonomy Application

Because a module can have many LOs, it would be ideal to be able to train all LOs using a single method. Clark indicates that this is not always possible by aligning instructional strategies or methods to various LOs across the cognitive, affective, and psychomotor domains (Clark, 1999). This implies that the application of a certain method or instructional strategy can only lead to the achievement of specific LOs. It also means that to achieve higher competency levels within each domain, several methods may need to be utilized throughout a training program. Table 11 shows Clark's relations between training method or instructional strategy and three domains. The entries below each domain have been adapted to use the nomenclature reviewed earlier in this chapter.

Table 11 Instructional Strategy Selection Chart adapted from Clark (Clark, 1999)

Instructional Strategy (Training Method)	Cognitive Domain	Affective Domain	Psychomotor Domain
Lecture, reading, audio/visual, demonstration, guided observations, or question and answer period	1 - Remember	1 - Receiving	1 - Perception 2 - Set
Discussions, multimedia computer-based training, Socratic didactic method, reflection. Activities such as surveys, role playing, case studies	2 - Understand 3 - Apply	2 - Responding	3 - Guided Responses 4 - Basic Proficiency
OJT, practice by doing (some direction or coaching required), simulated job settings (to include computer-based training simulations)	4 - Analyze	3 - Valuing	5 - Complex Responding
Use in real situations. Also may be trained by using several high level activities coupled with OJT	5 - Evaluate	4 - Organizing	6 - Adaptation
Normally developed on own (informal learning) through self-study or learning through mistakes, but mentoring or coaching can speed the process	6 - Create	5 - Characterizing	7 - Origination

It should be noted that some of the instructional strategies do not prescribe a specific solution regarding training media or technology. For the trainer responsible for choosing the content delivery methods, this can be both helpful and frustrating. While it is useful to be able to apply self-paced slide shows, demonstrations, or AR/VR platforms within a training curriculum, each delivery method carries consequences to implementation. Thus, training outcomes not only need to be balanced with the selection of the proper training method, but training technology, program costs, performance, risk, and schedule considerations.

MPEET Model

A common concern when restructuring training programs is predicting the effectiveness of the new training solution. This is especially true within the setting of

government acquisitions, as the bid that fulfills the baseline requirements for the lowest cost is typically the winner. How can the Air Force predict training requirements and their relative value before a new training program is implemented? As a solution to this problem, Brown introduced a Methodology to Predict and Evaluate the Effectiveness of Training (MPEET) (Brown, 2014). “MPEET uses primary elements of learning theory and instructional design to predict the cost-effectiveness of a training program, and recommends training alternatives based on decision-maker preferences for each of the cost and effectiveness criteria” (Brown, 2014). This five-step model follows the same basic top-down structure as the VFT model, starting with writing objectives, generating alternatives, and then evaluating to determine the best solution. The process is illustrated in Figure 9.

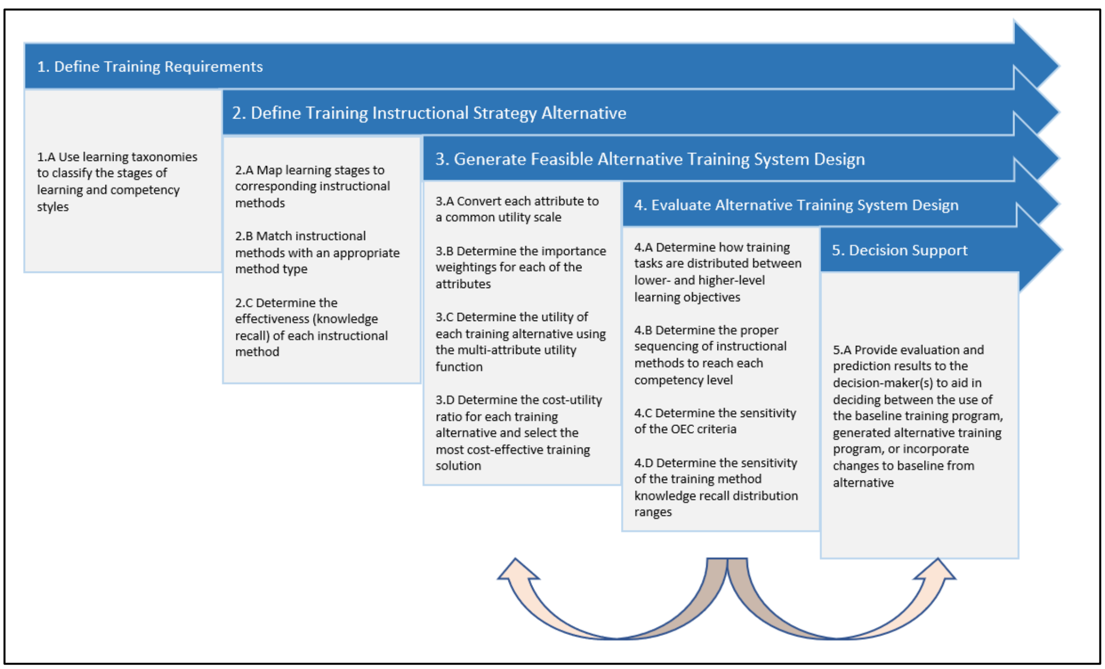


Figure 9 MPEET Model (Brown, 2014)

The first step in MPEET is to define the training requirements in the applicable domains to classify learning stages and competency of the trainee. This involves selecting an appropriate learning taxonomy (cognitive, psychomotor, affective, or interactive) that aligns with LOs.

The second step of MPEET is to define training instructional strategy alternatives, which includes accomplishing three subtasks. The first subtask involves mapping learning stages presented in each taxonomy to their corresponding instructional strategies. The recommended method for performing this subtask is to use a compatibility matrix, through which the user may visualize compatible alternatives. The second subtask for step 2 involves expanding the compatibility matrix to contain media types in addition to the training methods. The example in Table 12 provides a compatibility matrix of training media (technology) and training methods. The number "1" is an indicator of compatibility.

Table 12 Training Compatibility Matrix Example

	Training Technology				
Training Method	**Chalkboard/ Whiteboard**	**Oral Description**	**Augmented Reality**	**Virtual Reality**	**PowerPoint Slides**
Lecture	1	1			1
Demonstration	1	1		1	1
Exhibit	1	1			1
Questioning	1	1			1
Seminar	1	1		1	1
Discussion	1	1			1

The final subtask is to determine the effectiveness of each training method. Specifically, a subject matter expert (SME) provides an estimate of knowledge recall in

units of percent. The SME is to provide the minimum, maximum, and most likely value estimates to create a triangular distribution. The intent of providing estimates with a triangular distribution is to be able to include a plausible range of values when the mean and standard deviation are unknown.

Step 3 of MPEET involves generating feasible alternate training system designs, which is accomplished through four subtasks. Brown assumes that the overall effectiveness model is influenced by and composed of ten attributes grouped into three main categories: learning objectives, instructional strategies, and criticality ratings. The first subtask is to define functions for each attribute and to normalize possible values on a scale from 0.00 to 1.00. After the functions for each attribute are established, the overall importance of each attribute is to be weighed. All weights should add up to 100%. Once the attributes are weighted, the attribute values are entered to calculate the overall utility value of each alternative. Then, the cost-utility ratio (CUR) is calculated for each alternative by taking the ratio of training lesson cost to the utility function outcome. The smallest CUR is considered the most cost-effective solution.

Step 4 of the MPEET model is to evaluate the alternative training system design by following four subtasks. The first task within this step is to determine how training is distributed among learning objectives. It is important to remember that LOs are characterized by the ordered levels of difficulty within each learning domain. Brown suggests that varying levels of LO competency, low to high, should be distributed across a training program according to best practices (Brown, 2014). The next subtask is to determine the sequence of instructional methods to reach the proper competency level. Sequence of training competency is particularly important for training media such as

virtual reality, which Brown would consider a higher-efficiency training method. This idea is aligned with the recommendation of "learning first, immersion second," which suggests that traditional training technologies such as slide shows or assigned readings be used to train initial content in preparation for more complex technologies such as AR or VR (Mulders, Buchner, & Kerres, 2020). One way to visualize this process is to plot the competency in the order that a skill is trained. If the slope of the competency increases chronologically, it is indicative of an effective training program. Brown also advocates for observing how instructional methods are paired with LOs to ensure that lower competencies are trained with passive methods and higher competencies are trained with active methods. "Examples of active instructional methods are class and group discussions where students are engaged in the discussion, practice, and eventually peer-to-peer teaching" (Brown, 2014). The next two subtasks determine the sensitivity of the overall evaluation criterion (OEC) and the knowledge recall distribution ranges. This is performed by varying the weights of the OEC and the knowledge recall ranges of each training method to determine if varying these factors result in a significant change to the training system design.

The final step in the MPEET model is to present the results to the instructional design team to check if adjustments to the alternative are necessary, then to forward the results to the decision-making entity. Overall, the MPEET model incorporates sound traditional top-down systems engineering (TTDSE) principles, instructional design, and learning theory to create a decision model that can be utilized to predict the performance of a training system solution. This is especially useful when no performance metrics are available for a novel training alternative.

Although the MPEET model incorporates some important considerations for training systems like having one training method per module and incorporating various taxonomies to create learning objectives, it has some drawbacks. These drawbacks include the assumption that the effectiveness of each training method is constant, only 9 attributes make up the total utility function, and only the highest-order competency level is given the most value. Each training method was assigned an estimated level of effectiveness (defined as knowledge recall) by a SME (Brown, 2014). This estimation was based on one expert's experience from a West Point study, although it was not referenced in the dissertation. Because the level of effectiveness was constant, possible moderating effects such as trainer access, information access, or immersion were unaccounted for. Moving forward, it is necessary to determine how training effectiveness is viewed by SME's and constructed by training organizations.

Secondly, only nine attributes made up the total value function. The nine attributes chosen for the value function are training method, resource cost, asset cost, affective competency, cognitive competency, psychomotor competency, difficulty level, importance level, and frequency level. Although some of the attributes may be necessary to include, such as the trainee competency levels in various domains, some attributes are missing from the list and others may be omitted in future value functions. Attributes that organizations may value include number of possible task repetitions, user safety, interactive competency level, and service life (Murty, Djang, Butler, & Laferriere, 1995; Kaplan, et al., 2021; International Civil Aviation Organization, 2016; Decker & Campbell, 1996). This list of proposed attributes is not exhaustive; however, it reveals the

need to incorporate additional categories into the total value function. The complete list of investigated attributes is listed in Table 13.

Third, the MPEET model assigns the greatest utility value to the highest-order competency level in each domain. This assumption is directly against a sensitivity analysis subtask objective, which seeks to ensure that the curriculum is organized in a sequence of increasing trainee competency. For example, if the goal of a module is to train "understand the operating principle of a torque wrench," there would not be any added utility value if a trainee could create a torque wrench. The cognitive level of "Create" certainly implies a high level of skill, but in the view of strategically sequenced modules for competency acquisition, this skill level may not always be the most desirable.

Summary

The apparent gap in research exists in an applied cost-benefit tool which serves to relate the estimated costs of implementing a training solution to anticipated performance gains. Performance is relative to each training organization and task, so some aspects of any decision tool will require tailoring. Due to the broad scope of available training methods, available technologies, and the vast problem space to be solved by training programs, the variables associated with training alternative selection are to be investigated. As new training technologies become more accessible, trainers and decision-makers within the DoD will need a method of justification that can serve as a common reference while preserving organizational requirements. Because training programs build the trainee skill to a desired level of competency, multiple training

platforms may need to be leveraged during a training program. Therefore, any analysis tool should not seek to determine which training platform should be developed, but when should each training platform be incorporated into a training regimen.

III. Methodology

Interview Methods

The research effort took a qualitative approach due to the form of the research questions and the lack of existing data on the topic. Qualitative research is typically exploratory in nature, based off preliminary information or data that is discovered by the researcher (Patten, 2009). A qualitative approach was important because those involved in the study provided information from an individual experience, which varied amongst participants. Findings were based upon respondents' current understanding of the topic at a specific point in time in a particular context (Merriam, 2002). Thus, it was important to obtain information from multiple sources. In this case, interviews included trainers with differing experiences to create an understanding of the research topic from multiple vantage points within the maintenance realm (Merriam, 2002).

To gain insight to these varied experiences, a semi-structured interview technique was used. Semi-structured interviews are often face-to-face and recorded so further examination of responses can be performed (Patten, 2009). However, interviews via phone call or video conferencing may also be used. Semi-structured interviews do not need to be entirely scripted, which can benefit a research effort by revealing more threads of information relevant to the topic (Patten, 2009). When different groups are to be interviewed, semi-structured interviews allow for slight improvisational changes to the interview script to make a question relevant for a specific party. Prior to interviews, it was anticipated that some participants would be unfamiliar with the learning domains and the method of scoring training system alternatives. The semi-structured interview

technique allowed for clarifying information to be given for greater participant understanding of the prompts.

Approach

The methodology portion of this research was primarily based off the VFT and MPEET models (Shoviak, 2001; Brown, 2014). The flow of interview questions came from the VFT model, which follows a top-down process. This includes steps for identifying the problem, finding the organizational objectives which characterize the solution, finding the pertinent measurement categories which are used to evaluate alternatives, and creating the value functions. This information along with content found in the literature review helped to inform a systems modeling language (SysML) model and a blueprint for an initial total value function. These inputs and outputs are represented in Figure 10.

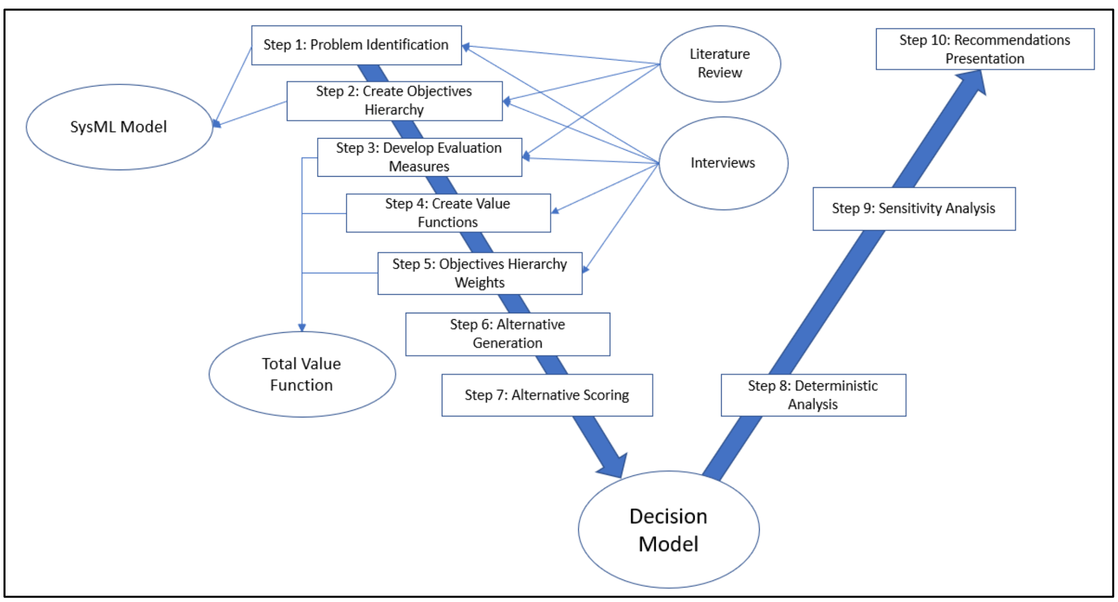


Figure 10 VFT Inputs and Outputs

Portions of the MPEET model were used as an inspiration for this research methodology due to its structural alignment with portions of the VFT model and its use of a total value function to calculate overall utility of training alternatives. MPEET's overall effectiveness model is influenced by LOs, instructional methods, and training technology. However, it did not consider some attributes that training organizations value, for example, user safety, immersion, and distance access. A list of the included attributes is presented in Table 13.

Table 13 Value Attribute Sources

Value Attribute	Source
Cost	(Decker & Campbell, 1996)
Service Life	
Immersion	(Lackey, Salcedo, Szalma, & Hancock, 2016)
User Safety	(Kaplan, et al., 2021)
Availability	
Schedule	(Buede & Miller, 2016)
Cognitive Competency	(International Civil Aviation Organization, 2016)
Psychomotor Competency	
Interactive Competency	
Affective Competency	
Trainer Access	(Webel, et al., 2013)
Information Access	
Tactile Feedback	
Repetitions	(Murty, Djang, Butler, & Laferriere, 1995)
Distance Accessibility	(Grossman, Oglesby, & Salas, 2015)

The methodology used in this research effort differed from the MPEET model in attribute decomposition and selection. Both cost and training effectiveness attributes are broad categories that can be broken up into constituent parts. A detailed look at what makes up cost and effectiveness is important for revealing what maintenance trainers value. As for the selection of attributes, this methodology differed from MPEET in that it

was up to the maintenance trainers to add in any attributes that were missing from their definition of a total value function. By challenging interviewees to respond to the questionnaire with a new focus on training methods, technologies, and LO competency thresholds of achievement, the research effort provides information on pertinent attributes, value functions, and technology combinations used for training module development.

Interviewee Details

The ideal interview candidate has worked in a training organization as a decision-maker or in a training role, had experience deciding between training alternatives, had a broad range of experience with training methods and technologies, and had more than two years of experience developing or assessing training systems. It was assumed that such a person had informed opinions on various contributors to the overall value of a training alternative. An interviewee did not need to fulfill all the requirements, but it was essential that he or she had knowledge of training within the maintenance field to increase the external validity of the study. These requirements were generated to describe the SME that Brown utilized to predict training effectiveness for various training methods (Brown, 2014). Although Brown's SME resource had over 25 years of experience, the requirement of two years' experience was used to broaden the potential pool of participants. Of 10 leads provided through personal reference or snowball sampling, 5 were interviewed. The interviews were administered from December 17th, 2021 to January 10th, 2022 via Zoom for government.

Of the 5 Airmen interviewed, respondents 1, 2, 4, and 5 had experience performing at least an informal decision analysis to choose the best training alternative for a given module. All respondents met the requirement of having an Air Force training background within the maintenance field. Additionally, all respondents had at least some experience with AR and VR technologies. Three participants were directly involved in the creation and delivery of training programs for a schoolhouse. The other two participants focused specifically on the development and use of training technologies across multiple maintenance applications. One issue was encountered during interviews. The recording for the first half of the Respondent 1 interview was lost due to a computer crash midway through the interview, however responses were recorded in the researcher's notes.

Interview Content

To provide the participants with a relevant context before responding to any questions, a simple block definition diagram was presented and explained. The diagram was built using Cameo Systems Modeler and is depicted in Figure 11.

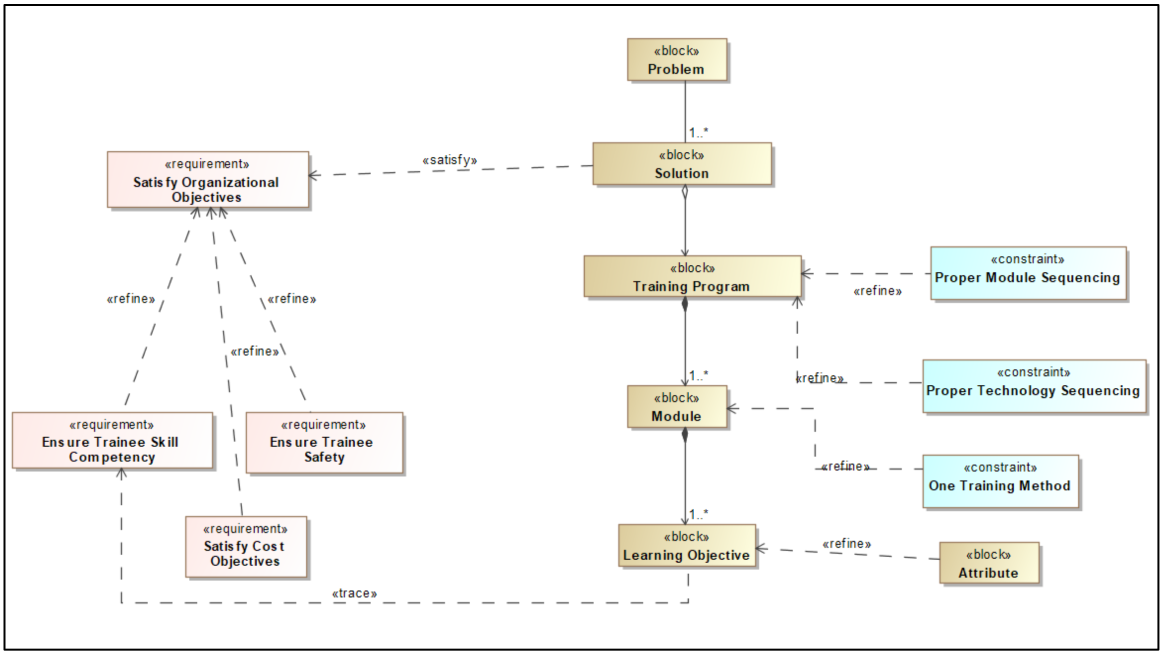


Figure 11 Interview Block Definition Diagram

Its intent was to illustrate how solutions fulfill organizational objectives, to show that modules are composed of at least one learning objective, and to communicate that there are constraints associated with building a sound training program. The interview consisted of 20 questions, 16 of which relate to attributes and their value functions. These individual value functions provide a basis for future utility estimates of training alternatives. The remaining four questions were designed to reveal common problems which have resulted in a training solution, organizational objectives that were met by implementing a training solution, and relationships between training methods and training technologies.

Several definition tables were utilized throughout the interview, including tables for various value attributes, cost types, learning domains, training methods, and training technologies. Sources found during the literature review phase provided the attribute

selections and definitions. Table 13 contains the value attributes utilized for questions 8-10 along with the sources from which they were found. Attributes that were not found in literature were not included in the interview script, however, respondents were given the opportunity to add value categories to ensure a complete representation of value attributes. The cost value attribute was broken down into the direct and indirect categories as discussed in Chapter 2, using the categories and examples given by Decker and Campbell (Decker & Campbell, 1996). The cognitive, affective, psychomotor, and interactive learning domain definitions were also pulled directly from Table 7 thru Table 10 in Chapter 2. Although training technologies and training methods were covered at length in previous chapters, MIL-HDBK-29612-2A contains useful definitions that specifically apply to training within a military context (United States Department of Defense, 2001). Thus, categories and definitions from this document were used to preface questions 18 through 20, with some additions. Additional categories include AR, VR, live artifacts, demo hardware, technical information, and real-life performance. These additions were motivated by the need to represent current training technologies and to illustrate distinctions between training methods (ex. performance versus real-life performance). Training methods and training technology categories used in the interview are represented in Table 14.

Table 14 Training Methods and Technologies

Training Methods	Training Technologies
Lecture	Chalkboard/ Whiteboard
Demonstration	Oral Description
Exhibit	Augmented Reality
Questioning	Virtual Reality
Seminar	PowerPoint Slides
Discussion	Computer-based lessons
Performance	Interactive computer courseware
Case Study	Simulator
Real life performance	Models
Indirect Discourse	Live artifacts
Assigned Reading	Demo Hardware
Small Group Method	Technical Information
Student Query	Workbooks/ Handouts

The complete interview handout is provided in Appendix A: Interview Handout. After an initial pilot study to refine the questionnaire and to ensure interviewee understanding, the interview period was projected to take 1-2 hours. On average, the interviews took 2.5 hours to gather the required data.

Data

The data that was produced from the interviews included unstructured verbal responses, ratings of attributes, participant-created value plots, compatibility matrices, fill-in-the-blank questions, and percent value estimates. The two objectives considered when processing data included finding an answer to the research questions and representing findings in a way that would be useful to trainers, like a roadmap that could walk them through the first seven steps of the VFT process. The first research question to be answered was to find which attributes contributed to the total value of training. The data gathered to answer this question included unstructured verbal responses and ratings

of importance. The verbal responses were reduced to a single word or short phrase that could communicate the meaning of the described attribute. This was accomplished through identifying key words and phrases throughout the interview, marking them down, and interpreting the context (Merriam, 2002). Although the interview materials contained multiple hypothesized attributes, the questions that were asked were meant to bring any additional value categories to light. A total of 6 additional attributes were discovered that were not initially included in the definition tables.

The most-least rating method was used to determine the order of importance of each attribute. This was due to the time limitation of the interview and the number of attributes to be considered. The most-least rating method requires the participant to choose the most important and least important attributes, then rate each attribute individually on a predetermined importance scale. The purpose of this was to force the participant to make differentiations or distinctions between the levels before assigning a rating, which may have provided more quality responses for analysis (McCarty & Shrum, 2000). Research shows that this method produces better results than a simple rating method, providing more differentiation or utilization of more rating levels, and less end-piling of results (McCarty & Shrum, 2000). A most-least rating method rather than a ranking method was chosen because of the intended use of the attributes. Since the attributes that make up the total value function can be assigned the same weight of importance by decision-makers, the respondents were able to do so indirectly. Although respondents were not assigning weights to each attribute, it is plausible that the importance rating indicated high and low weights in addition to distinguishing between unimportant attributes. However, it was assumed to be more likely that the results would

only point toward the attributes which should have been included in the total value function calculation.

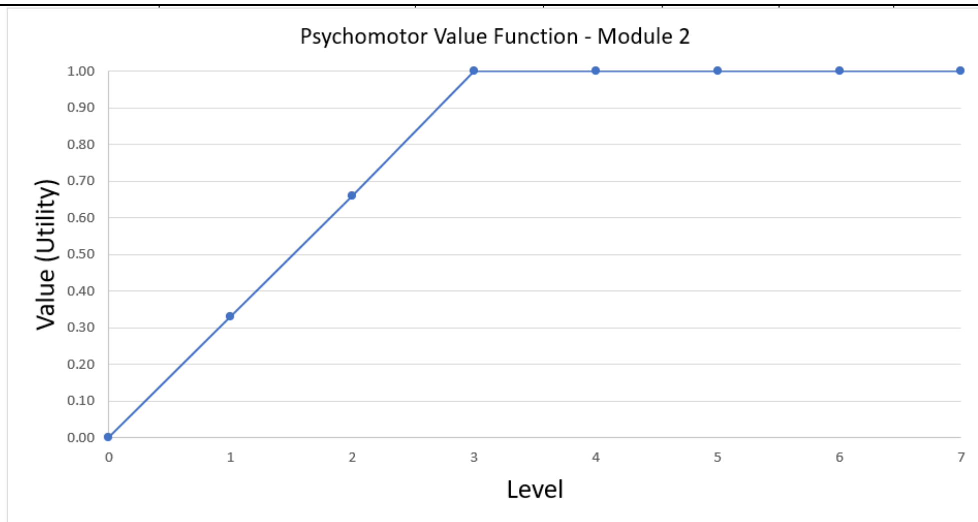
The list of attributes was randomized a total of 12 times to obtain an ordered list for the expected maximum number of participants. The intended effect of randomizing the list of attributes was to reduce bias incurred from ordering. The ratings were based on a scale from 0-10, with 0 indicating the least important and 10 indicating the most important. Participants were allowed to respond with integers only. This rating scale was selected to provide consistency and separation. The scales in this interview each had a maximum value of 1, 10, or 100 so that the participant could have some sense of proportionality when responding across all questions. It was anticipated that many attributes could have the same rating, so if the respondent utilized all levels, only 4 attributes could have the same rating. Verbal responses and ratings answered the research question by ensuring the most valuable attributes are considered while revealing those that would likely be highly influential in alternative selection, as indicated by the participants.

Questions 11 and 12 were about the cost attribute. Ranking was utilized for these responses because there were only 9 types of cost that needed to be ranked. Additionally, participants were asked to identify which types of cost were not important to include in decision analysis. The purpose behind these questions was to investigate the ordering of importance and to see which costs were typically incurred and evaluated when implementing a training program.

The second research question explored whether value is a function of learning objectives. The data gathered to answer this question included several value plots on cost

and trainee learning objectives along with a verbal response to clarify graphical trends. To answer whether value is a function of learning objectives, the plots were created in a deterministic fashion. In other words, the participant was given the opportunity to assign the value of each data point on a scale from 0.00 to 1.00, rather than assuming an increasing or decreasing value trendline proportional to the x-axis values or LO domain levels. Hermans and Cunningham recommend that when scoring values on a scale, expert opinion should be used, as they are “able to evaluate both new and existing solutions along a range of technical criteria. They may also be able to handle scoring both tangible and intangible quantities (Hermans & Cunningham, 2018).”

For the cost value plot, the question was designed to expose the participant to plotting value functions and, if the alternative costs were normalized properly, to provide an exemplary function that is almost proportional except for one data point. The intent was to provoke the participant to think deterministically for the value plotting questions. LO value functions were built by participants in random order. For each LO plot, two trends that may be qualitative indicators of value as a function of LOs. The first is that the objective level of competency for each of the learning domains would be valued highest. The second is a decrease in the slope of the trendline with the objective level of competency as the inflection point. For example, assume the objective level of psychomotor competency is “Guided Responses” as expressed in the prompt for module two. Figure 12 depicts both indicators, with “Guided Responses” marked as both the highest value and as the inflection point where the slope has decreased.



Level	0	1	2	3	4	5	6	7
	None	Perception (Awareness)	Set	Guided Responses	Basic Proficiency	Complex Responding/ Mastery	Adaptation	Origination
Value	0.00	0.33	0.66	1.00	1.00	1.00	1.00	1.00

Figure 12 Psychomotor Value Function Example

The highest possible value is clearly marked at 1.00, however, it should be noted that there is no established, consistent indicator for what slope decrease might mean to the participant. To ensure that the trends of the value plots aligned with the intent of the participant, a verbal response was gathered from the participant as to why or why not the highest level of competency was the most desirable.

The third research question was to find the best or highest-valued combination of training method and training technology. The data gathered to answer this question included a compatibility matrix and a fill-in-the-blank section. The reasoning behind the use of a compatibility matrix is that it helped to narrow the technologies that were particular to the training method before the selection of the best combination of technologies. The compatibility matrix listed the training method along the rows and training technology along the columns. If the participant entered a “1” in a single cell of

the array, the corresponding row (method) and column (technology) were deemed compatible with one another. Then, once the matrix was filled out, the best combination of compatible technologies was selected by the participant via fill-in. The matrix responses from all participants were tallied and entered in the corresponding cells in the results section. The cell counts indicated how many participants agreed as to which training technologies are compatible with each training method. The fill-in section was compared between participants to find a recommendation for the best combination amongst compatible technologies for each training method.

In addition to the research questions answered, the last interview question obtained the percent value estimates of training technologies for various attributes. The purpose of this question is to set the groundwork for future research on attributes of training and to determine if the utilization of certain training technologies could result in higher value to training organizations. This question can also be used to determine the utility of each technology across important attribute categories.

Implementation

The culmination of the data processing effort provides a roadmap to evaluate training alternatives, specifically by illustrating common problems, identifying objectives of maintenance organizations, and building an initial total value function highlighting important attributes. Both the problems and objectives can be represented in a single block definition diagram (BDD). The advantage of listing both problems and objectives in a BDD is that specific objectives can trace to problems so the viewer can see what the objectives are fulfilling. By creating models in SysML, organizations can refer to these

diagrams to discover the types of problems requiring a training solution and have an example of an objectives hierarchy. This is essential for those who do not have experience or knowledge of systems engineering processes as it gives information on the initial steps. By viewing the diagrams as a starting point, users may refine or make edits to listed problems and objectives to better fit the needs of their organizations.

The total value function mentioned in the literature review section is important because it incorporates pertinent measures of training system alternatives by which the alternatives may be scored. It is also the primary decision tool in the model because its output will theoretically indicate the best training system alternatives. There are two important aspects of a total value function, the attributes which the stakeholders wish to incorporate, and the weight of the attributes. Attribute selection is the easier of the two, as the stakeholders need only to consider which they think are important. When given a list of attributes, it may be easier for users of the tool to think of additional categories that are relevant to their objectives since some have already been defined. Weighting the attributes is more difficult, as perceived importance of each attribute among multiple decision-makers may vary. Constraints such as cost, and schedule may be disqualifying factors for some training system alternatives even if the calculated total value is highest. It should be noted that the tool this research provides is incomplete in view of the overall VFT framework, but it may be useful as a starting point for calculating cost-utility of training alternatives and considering alternative training method-technology combinations.

Assumptions

This research is most useful under a specific set of assumptions, starting with accurate problem identification. Problem identification is an important first step because it identifies the issue that is being experienced and documents it for all stakeholders to see, thereby establishing a common ground (Buede & Miller, 2016). It is essential that the “right problem” is established so that resources are not put toward solving a different problem. In this case, the right problem is one that is caused by training. The second assumption is that problems not caused by training would be ruled out. It follows that to pursue a training solution, a training-based problem is to be fully defined, including potential causes outside of training. Outcomes not caused by training problems may still need to be addressed, but they are only related to this research as they define the boundary between what is and is not relevant to training. The third assumption is for the training organization to gain input from all relevant stakeholders on the objectives hierarchy. These objectives are to be agreed upon so that important outcomes may be realized. The fourth assumption is that training system design best-practices will be followed for the purposes of scoring and generating alternatives. Good training system design includes employing one training method per module, the proper sequencing of modules, and proper technology sequencing, as illustrated in Chapter 2. By fulfilling these assumptions, it establishes a training solution as the most effective course of action and lays the groundwork for integration with steps in the VFT model.

Data Analysis

The primary method used for analyzing verbal response data was keyword identification. Other responses were represented through frequency distribution tables, bar charts, scatter plots, and comparing mean and median responses. The goal of these representations was to identify trends that would inform results applicable to answering the research questions. The first research question sought to find the importance of attributes related to training. This was answered by repeated measures analysis of variance (ANOVA). This method compared the mean importance rating for each attribute across all responses to determine if any attribute would be important to consider in a total value function. Repeated measures ANOVA has three assumptions:

1. Observations are independent.
2. The data are normally distributed.
3. Equal variance between combinations of groups (Sphericity).

The hypotheses for the repeated measures ANOVA are:

> H_0: The population means are equal with respect to importance rating.
> H_a: At least two population means are significantly different with respect to importance rating.

The second research question sought to determine whether utility value is a function of LOs. Respondents provided utility value ratings for each competency level with respect to each of the 9 treatments. An objective competency level and high competency level was identified for each treatment and the difference in value ratings was compared via sign test. The sign test has four assumptions:

1. The dependent variable (utility value) is measured on a continuous level.
2. The independent variable (treatment) should consist of matched pairs.
3. The paired observations are independent of one another.
4. The distribution is continuous.

The hypotheses for the sign test are:

> H_0: The median of differences between the objective competency rating and the high competency rating is equal to 0.
> H_a: The median of differences between the objective competency rating and the high competency rating is not equal to 0.

The third research question sought to determine the best combination of training technology and training method, which was accomplished by qualitative analysis of frequency and rating tables. The experimental alpha level chosen for this research was 0.1 due to its exploratory posture.

Summary

The semi-structured interviews resulted in many data types, each requiring unique processing considerations to obtain an answer for each research question. A total of 5 participants were interviewed due to the difficulty of finding experts who fit the interviewee description. Statistical methods used include repeated measures ANOVA and sign test. This qualitative research applies findings to create a SysML model and an initial guide to total value function creation. This total value function may be used to calculate an initial estimate of the cost-effectiveness of training system alternatives.

IV. Analysis and Results

Chapter Overview

Data collected during interviews was analyzed to produce an objectives hierarchy and statistically significant results. The statistical methods employed during analysis included repeated measures analysis of variance (ANOVA) and the sign test at an alpha level of 0.1. Qualitative trends in data were also observed, yielding results that were relevant to answering the three primary research questions.

Problem Identification

Respondents were asked to provide problem areas which led their organizations to revise or implement training programs. Responses are provided in Table 15. Some of the most common problems were related to unfavorable training outcomes, lack of trainee engagement, outdated equipment, environmental hazards, and disparity between schoolhouse (IST) training and flightline training (OJT).

Create Objectives Hierarchy

Step 2 in the VFT framework is the creation of an objectives hierarchy. Respondents were asked to provide the main objectives of revising some aspect of a training program. The results are listed in Table 16.

Table 15 Question 1 Results - Problem Identification

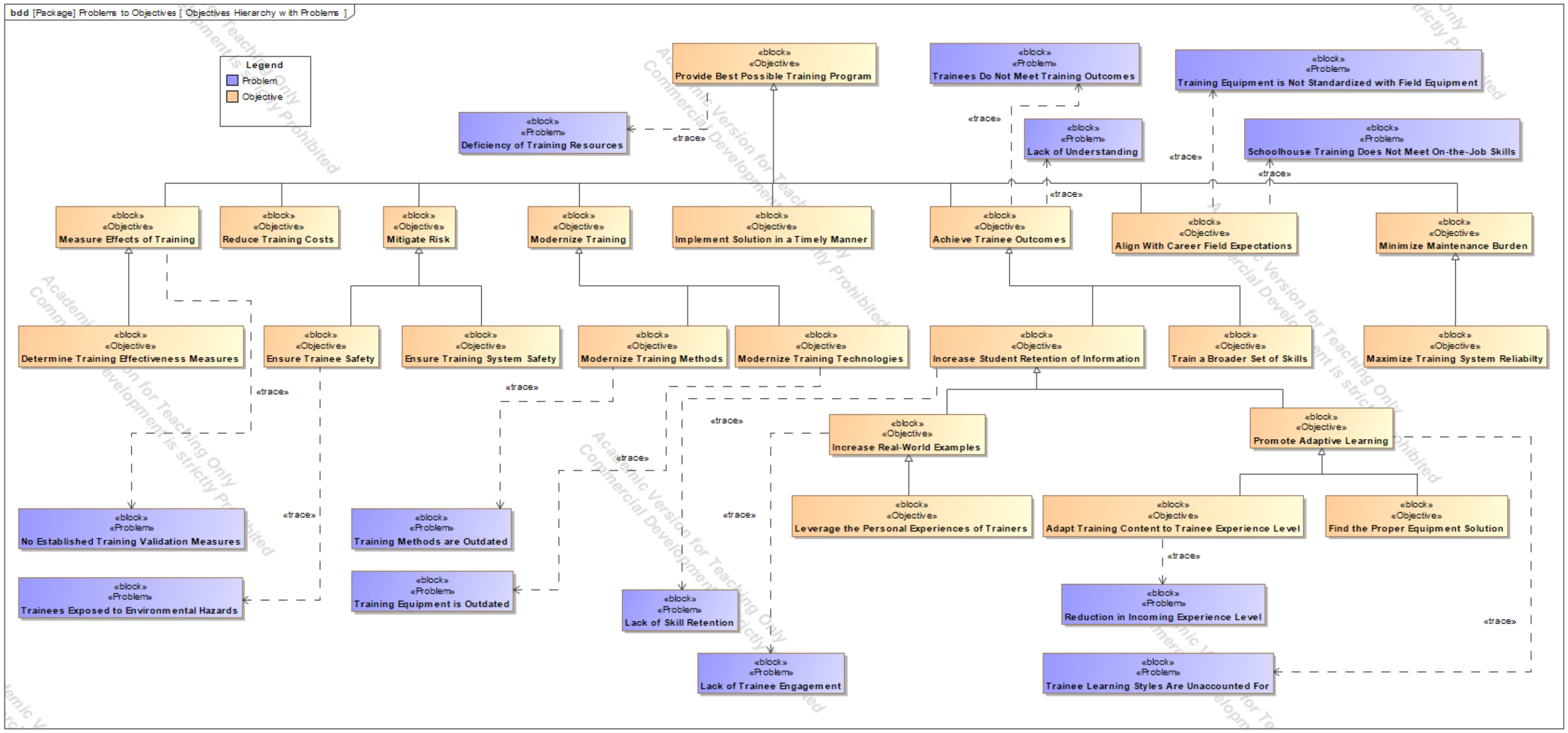
Respondent	Problem	Description
1	**No established validation measures for training**	Not having measures to determine that training was accomplished effectively and efficiently.
	Trainee learning styles are unaccounted for	The adaptability of training to various learning styles.
	Deficiency in training resources (includes both trainer competency and other resources)	Trainers are unable to train well.
2	**Training Equipment - Age of Technology**	"It's been equipment. Either the equipment is old, or it's not what the field has"
	Training Equipment - Standardization	
	Outcomes of Training	"The career field has a training deficiency"
	Technician lack of understanding	The technicians "didn't understand how to set it up" (referring to following an Air Force procedure for making diagnostic measurements)
	Safety (training equipment hazards, environmental hazards)	Participant illustrated an example of the need to substitute non-hazardous fluid for jet fuel during training.
3	**Training aids/methods outdated**	"We haven't changed anything in decades, it feels like, [sic] and the same training that I went through is the same training my dad went through."
	Trainee learning styles differ	"...the trainees that are coming in are not someone [sic] who learns like that"
	Schoolhouse training does not reflect On-the-job skills	"When I was on the flightline... what's been said before in the past is that 'the way they taught you in tech school is wrong, and what we're going to do is dump all that information and we're going to teach you how to do it now.'"
4	**Retention of trainee skill**	"The biggest reasons [sic] for altering training material and the method is the retention of the student. In a lot of cases, when we first started out, the students weren't grasping the concepts because they were assuming they had a level of understanding which they didn't have... You have to adjust the material to bridge the gap in a lot of cases"
	Reduction of experience level in incoming students	"The experience level of the students has been reducing over time, so what worked two years ago, when I had more experienced technicians coming through, doesn't necessarily work today. So, you have to adapt live with the overall level of understanding that your students have, and believe it or not, it does change from year to year because of the people you're getting in"
5	**Lack of trainee engagement**	"For those who want to learn, it's very easy for them to sit and listen... Being able to keep their interest and still be able to teach them what they need to know. Everything that we're told to teach isn't everything they need to learn."
	Survey results indicate performance issues	"GAS - Graduation Assessment Surveys that field supervisors are supposed to fill out and give back to the schoolhouse"
	Informal conversations indicate performance issues or unfavorable outcomes	Respondent cited an example of conversations had with another technician when he or she leaves a duty station to become a trainer at a schoolhouse for the career field.

Table 16 Question 2 Results - Objectives

Respondent	Objectives	Description
1	Determine effectiveness measures	"What do we value as effectiveness?"
	Find the proper hardware/equipment solution	"Did it [equipment] satisfy the need for us?"
2	Train to a lesser level of understanding of a broader range of topics	Participant gave an example (career field specific) about switching from teaching advanced skills to a broader set of skills.
	Achieve trainee outcomes	"The career field has a training deficiency"
	Ensure trainee safety	Participant illustrated the need to substitute non-hazardous fluid for jet fuel during training.
3	Tailor training equipment to airman preference	"Break the system… break how tech school is being taught… Our job is to change as much as we can to try to get something new in there, to get people to realize that the training that is being taught currently is not optimizing the Airmen that are coming in"
	Promote adaptive learning	"...try to give the airman as many options to learn as possible"
4	Increase student retention of information	"The main objective is always for student retention, so in any cases, [sic] if I'm noticing any issues where I'm measuring consistently that they're not retaining the information, ultimately, that's what's required in the revision… We survey each objective in the course."
	Adapt training content to experience level	"The experience level of the students has been reducing over time, so what worked two years ago when I had more experienced technicians coming through, doesn't necessarily work today. So, you have to adapt, live, with the overall level of understanding that your students have, and believe it or not, it does change from year to year because of the people you're getting in."
	Align with career field expectations	"…keeping current with what the career field's expectations are"
5	Leverage the personal experiences of trainers	"One of the goals as being an instructor is to also roll in our personal experience so they understand what they may possibly see in the field. So, we tend to use a lot of personal or hearsay examples of things we know of while we're teaching certain subjects."
	Increase real-world examples	"Making things relevant for the students so that they would stay interested, and also relevant to what we were using in the field."
	Modernize training technologies	"Our main goal now for the unit is to help modernize things now for students that are in school now, like teenage students now. But me and the commander are trying to view things even further down to elementary school where they're getting iPads and stuff like that... and if they were to come into the Air Force now, they'd be getting recycled paper book, not even books, like stapled together papers."

Common objectives expressed by respondents included overcoming trainee deficiencies, adapting training to suit student needs, increasing the relevancy of the training content, and improving/measuring the effectiveness of training. The culmination

of information from the first two steps of the VFT process resulted in a SysML BDD, represented in Figure 13.


bdd [Package] Problems to Objectives [Objectives Hierarchy with Problems]
Legend
Problem
Objective
Provide Best Possible Training Program
Trainees Do Not Meet Training Outcomes
Training Equipment is Not Standardized with Field Equipment
Deficiency of Training Resources
Lack of Understanding
Schoolhouse Training Does Not Meet On-the-Job Skills
Measure Effects of Training
Reduce Training Costs
Mitigate Risk
Modernize Training
Implement Solution in a Timely Manner
Achieve Trainee Outcomes
Align With Career Field Expectations
Minimize Maintenance Burden
Determine Training Effectiveness Measures
Ensure Trainee Safety
Ensure Training System Safety
Modernize Training Methods
Modernize Training Technologies
Increase Student Retention of Information
Train a Broader Set of Skills
Maximize Training System Reliabilty
Increase Real-World Examples
Promote Adaptive Learning
Leverage the Personal Experiences of Trainers
Adapt Training Content to Trainee Experience Level
Find the Proper Equipment Solution
No Established Training Validation Measures
Training Methods are Outdated
Trainees Exposed to Environmental Hazards
Training Equipment is Outdated
Lack of Skill Retention
Reduction in Incoming Experience Level
Lack of Trainee Engagement
Trainee Learning Styles Are Unaccounted For
«trace»

The BDD represents an objectives hierarchy that has specific objectives traced to the problems encountered by respondents. The problems are stereotyped and represented in an indigo color. Objectives provided by respondents are stereotyped and represented in orange.

Develop Evaluation Measures

Step 3 of the VFT framework requires the development of evaluation measures. The term "attribute" was used during the interview to represent the categories by which training alternatives would be evaluated. Questions 3 through 6 provided two scenarios designed to evoke multiple attributes of training. Scenario 1 asked the respondents which attributes they would consider if the training alternatives included classroom delivery versus hands-on delivery with training equipment. The results along with each respondent's most important attribute are indicated in Table 17.

Attributes that were considered most important by the respondents included equipment, cost, training content, effectiveness at emulating the real world, and safety. Attributes that had not been considered by the researcher included complexity of the task, effectiveness at emulating the real world, intensity of training experience (Law of Intensity), and instructor attitude.

Scenario 2 asked the respondents which attributes they would consider if training delivery options included a dedicated piece of training equipment or a VR system. The results and the respondent's opinion of which attribute is most important is listed in Table 18.

Table 17 Question 3 & 4 Results – Training Attributes

Respondent	Attributes to Consider**	Quotation
1	Resources	
	Facilities	
	Environment	
	Method Suitability	
	Equipment*	
	Presentation Devices	
	Time	
	Money	
	Manning	
	Transfer/Likeness	"Hands-on is the way to go"
	Student Need	Participant had "held a class with trainees focusing on learning styles" in the past.
	Student level of experience	
2	Complexity of the task	
	Performing tasks how they are performed on the job	"…the hands-on… for our career field, would [sic] use that more"
	Safety	"There are some things where I would consider classroom would probably work, like safety"
	Cost*	"Cost would be one of the attributes that would be impactful. If we don't have the funds to do some of the hands-on [training], then we're going to have to find something else"
	Adaptability to trainee learning styles	"...reaching the learner [is the most important facet of training effectiveness] because every learner is different"
3	Performing tasks how they are performed on the job	"Maintenance is a very hands-on kind of job, so we want to make sure that we always kept anything that had to do with hands-on training"
	Sequence of Training	"...have them understand the concept behind it first before we go to a hands-on piece."
	Equipment availability	
	Training Content*	"...what is being taught."
4	Laws of Learning - Merrill's Principles of Instruction	"Activate knowledge, demonstrate, integrate, and apply"
	Applicability to real-world application	"When I'm designing a real-world curriculum, how close can I get to the real thing is the first thing you need to consider, because it shows the students that the information matters. This is something you are truly going to be using when you go back to whatever it is"
	Law of Intensity	
	Effectiveness at emulating the real world*	
	Instructor attitude	"The attitude the instructor has toward the information matters a lot… the feedback that you get in the course surveys that we do, and in the objective [module] surveys, the passion of the instructor makes a major difference on their perception of the information. You can get the exact same information, but if the instructor isn't buying what he's selling, don't expect the students to buy it."
5	**Safety***	
	What the training module Is	
	Importance of the Task	
	Cost	"It's more worth it to have dedicated training equipment."
	Risk of error	"I want to make sure they're doing it right so that they don't screw up or hurt somebody or themselves"

* Attribute was designated as the most important

** Attributes without accompanying quotation or explanation were provided by the respondent verbatim

Table 18 Question 5 & 6 Results - Training Attributes

Respondent	Attributes to Consider **	Quotation
1	Availability	
	Equipment*	
	Ability	"VR is best to have if you don't have the ability to do so otherwise"
2	Programmed scenarios	"The benefits of virtual reality. Is it going to be the same experience in the virtual reality as it would be with the dedicated training equipment? Sometimes virtual reality can get really, really close to being the same as the equipment, but then once you're actually at the equipment, you have all these things that didn't come up in the virtual reality. If you can't make it [real enough], you are losing time instructing the differences between the two"
	Fidelity of training technology *	"It wouldn't break the dedicated piece of training equipment. It would act the same so you don't have the training gap with the virtual reality… you pick up both, essentially. If virtual reality could reach the same competency as the dedicated training equipment."
3	Location of Training	"With developing VR systems with tech schools… the main reason why we got into it was to give more context to the airman, exactly what they're going to be experiencing… give visual aids to that airman in a virtual environment, where I don't have to take them out to the flightline to experience it"
	Content	"What content are you looking to train on?"
	Skill Transfer	VR "should never replace the hands-on experience, especially in maintenance"
	CFETP Line Items*	"CFETP line items are the most important factor"
4	**Safety***	"Can I teach them real-world… Virtual reality is great when you can't teach them the real thing… Think of air traffic controllers. You can't have a student land real aircraft. There's too much risk there, so virtual reality, in that case, would be an option because it's the closest thing I can get to the real world. I can give them a virtual reality simulation of an airfield because it's just not safe."
	Cost	"VR, in my opinion… unless you have a lot of money, is not a really good solution in most cases… because [of] the cost of the system."
	Fidelity of training technology	
5	Fidelity of training technology	"If the virtual reality system could be up to my standard, I would go for that all day long" What defines this airman's standards are "basically the tactile feedback you need to actually be working with something… fidelity of controls"
	Task Requirements	"The task you're training"
	Scalability of equipment	Referring to maintenance work such as changing a tire on a plane, "virtual reality is better for that since we can't have full planes… all over the place for every student"
	Importance of the Training*	"Importance of the training… overrides everything"

* Attribute was designated as the most important
** Attributes without accompanying quotation or explanation were provided by the respondent verbatim

The attributes considered most important included equipment, fidelity of training technology, CFETP line items (training objectives found in official documentation), safety, and importance of the training. Most respondents brought up attributes related to the fidelity of the training technology.

Question 7 extended the two scenarios by asking the respondents about tools they had used to assist with decision-making. Respondents 1 and 2 utilized informal practices when deciding between training alternatives, citing need as an important driver. Respondent 3 helped to develop training but was not involved in decision-making. Respondents 4 and 5 cited both formal and informal frameworks as tools to aid in training development. The results are listed in Table 19.

Table 19 Question 7 Results - Decision-Making Tools

Respondent	Tool Name	Description
4	Task Analysis as part of the ADDIE process (Analysis, Design, Development, Implementation, Evaluation)	Breaking down tasks into smaller components to make the content trainable in a step-by-step fashion.
5	Risk Management Framework (RMF) for Air Force Information Technology (IT)	AFI 17-101 published 6 February, 2020. An Air Force Instruction on how to manage risks associated with cybersecurity across a system's lifecycle.
	Cost-benefit analysis	A spreadsheet owned by a decision-maker with limited access by others involved in the process.

The respondents were presented with the list of attributes from Table 13 along with their definitions. Question 8 asked respondents to review the list and provide additional attributes. The results are listed in Table 20.

Table 20 Question 8 Results - Additional Attributes

Respondent	Attribute	Attribute Description
1	Learning Style Adaptability	The ability to adapt the training content to multiple learning styles.
2	N/A	N/A
3	Retention	The level of performance retained by the trainee from the time of training until a skill is performed.
4	Instructor Attitude	The level of enthusiasm or belief in the purpose of the training content by the instructor.
	Fidelity	"How close it is [sic] to the real-world application"
5	Security	The level of security of the training system from an information technology (IT) perspective
	Ease of Use	"The ability [of the trainee] to use the tools that the career field needs"

Objectives Hierarchy Weights

Questions 9 and 10 implemented the most-least rating method, requiring respondents to select the most and least important attribute from Table 13, then rate each attribute on a scale from 0 to 10 where 0 represents the least importance and 10 represents the most importance. Attributes were presented in random order to the respondents. The results are listed in Table 21 and Table 22.

Table 21 Question 9 Results - Attribute Importance (Most-Least)

	Attribute				
Importance	Respondent 1	Respondent 2	Respondent 3	Respondent 4	Respondent 5
Most	Immersion	Availability	Cognitive Competency	Immersion	User Safety
Least	Cost	Schedule	Cost	Schedule	Distance Accessibility

Only Respondent 5's identification of safety as the most important factor was the same in Question 4 and Question 9. The remainder of respondents selected an attribute from the initial list as the most important.

Table 22 Question 10 Results - Attribute Ratings

	Rating					
Attribute	**Respondent 1**	**Respondent 2**	**Respondent 3**	**Respondent 4**	**Respondent 5**	**Average**
Cost	5	9	6	10	6	7.20
Immersion	9	8	10	10	3	8.00
User Safety	9	9	9	10	10	9.40
Schedule	3	0	5	10	8	5.20
Cognitive Competency	7	9	10	10	8	8.80
Psychomotor Competency	7	8	8	10	5	7.60
Interactive Competency	6	8	10	4	7	7.00
Affective Competency	6	8	8	8	6	7.20
Trainer Access	7	8	10	8	4	7.40
Information Access	8	6	10	6	5	7.00
Repetitions	7	7	10	8	4	7.20
Tactile Feedback	8	7	9	6	7	7.40
Service Life	6	6	7	7	2	5.60
Distance Accessibility	4	4	9	2	3	4.40
Availability	9	10	8	10	9	9.20

The mean number of ratings utilized by the respondents was 7 of 11 possible ratings. The lowest rating given was by Respondent 2, who gave the schedule attribute a rating of 0, which was consistent with their response in Question 9 as the least important attribute. Respondents 1, 3, 4, and 5 were not consistent when rating their least important attribute in Question 10. The attribute of least importance for these respondents, as expressed by the ratings was schedule, schedule, distance accessibility, and service life, respectively. There may have been some confusion about what the question was asking

for respondents 3 and 4, who both expressed that cost should not be considered an important factor, despite their awareness that it was typically used as a constraining metric when evaluating training alternatives. Figure 14 illustrates the distribution of attribute importance ratings, which is highly skewed toward the high values on the scale. Table 23 lists the summary statistics.

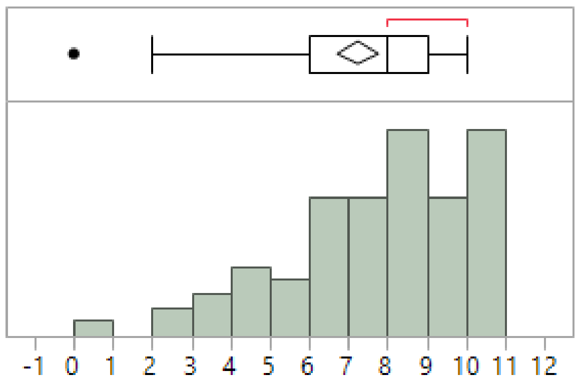


Figure 14 Histogram of Attribute Importance Ratings

Table 23 Summary Statistics for Attribute Importance Ratings

Summary Statistics	
Mean	7.24
Std Dev	2.3356231
Upper 90% Mean	7.6892
Lower 90% Mean	6.7908
N	75
Skewness	-0.821276
Kurtosis	0.2585503
Median	8
Interquartile Range	3

Figure 15 is a boxplot of ratings of importance for each attribute. Availability, cognitive competency, and user safety appear to be the most highly rated.

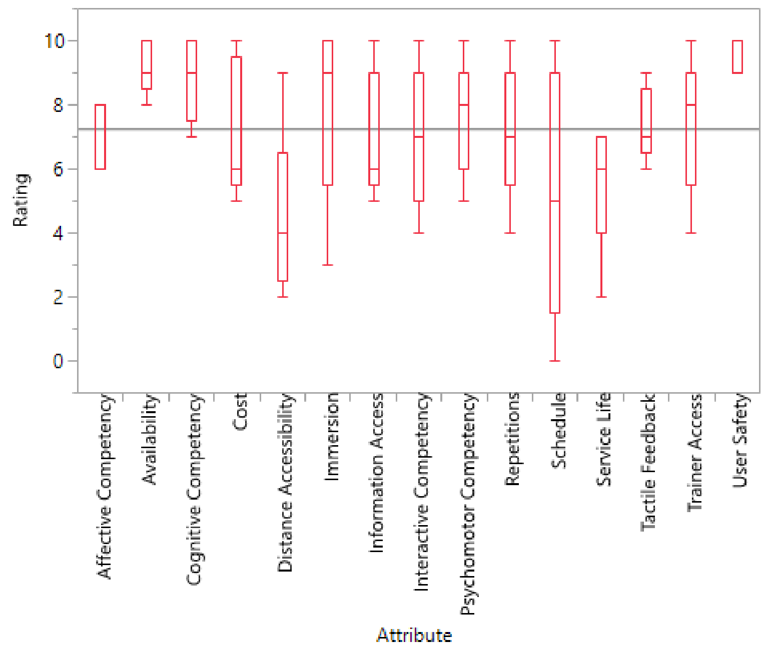


Figure 15 Question 10 Results - Boxplot of Attribute Importance Ratings

Table 24 illustrates the results of the repeated measures ANOVA Greenhouse-Geisser test. This test was used due to sphericity not being satisfied. The test fails to reject the null hypothesis that the means of each attribute with respect to importance rating are equal at an alpha level of 0.1.

Because the result of this test was close to rejection, pairwise comparisons of attributes were also calculated, which resulted in no statistically significant pairings at the experimental alpha level due to the Bonferroni correction. Without the Bonferroni correction, there are several significant pairings. Significant pairings may be useful for future investigations. These pairings are indicated in Table 25 by the presence of a p-value.

Table 26 provides 90% confidence intervals for each attribute. User safety had the highest mean rating and distance accessibility had the lowest mean rating. Standard error

was small compared to the mean. Ideally, mean values of importance for each attribute within the value function could function as weights of importance for step 5 of the VFT process.

Table 24 Question 10 Results – Greenhouse-Geisser

Tests of Within-Subjects Effects						
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
attribute	Sphericity Assumed	133.680	14	9.549	2.675	.005
	Greenhouse-Geisser	133.680	2.767	48.320	2.675	.101
	Huynh-Feldt	133.680	9.593	13.935	2.675	.015
	Lower-bound	133.680	1.000	133.680	2.675	.177
Error(attribute)	Sphericity Assumed	199.920	56	3.570		
	Greenhouse-Geisser	199.920	11.066	18.066		
	Huynh-Feldt	199.920	38.373	5.210		
	Lower-bound	199.920	4.000	49.980		

Table 25 Attribute Pairwise Comparisons

Attribute Pairwise Comparisons (No Bonferroni Correction)																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Cost	1			.074		.099										.047
Immersion	2													.004	.066	
User Safety	3	.074			.055				.020		.090		.047	.024	.022	
Schedule	4			.055												.092
Cognitive Competency	5	.099							.003			.099		.016	.020	
Psychomotor Competency	6													.011	.094	.099
Interactive Competency	7														.012	
Affective Competency	8			.020		.003								.078	.073	.022
Trainer Access	9													.009	.034	
Information Access	10			.090											.012	
Repetitions	11					.099								.016	.038	
Tactile Feedback	12			.047											.018	
Service Life	13		.004	.024		.016	.011		.078	.009		.016				.021
Distance Accessibility	14		.066	.022		.020	.094	.012	.073	.034	.012	.038	.018			.035
Availability	15	.047			.092		.099		.022					.021	.035	

Table 26 Confidence Intervals for Attributes

Estimates				
attribute	Mean	Std. Error	90% Confidence Interval	
			Lower Bound	Upper Bound
Cost	7.200	.970	5.133	9.267
Immersion	8.000	1.304	5.220	10.780
User Safety	9.400	.245	8.878	9.922
Schedule	5.200	1.772	1.422	8.978
Cognitive Competency	8.800	.583	7.557	10.043
Psychomotor Competency	7.600	.812	5.868	9.332
Interactive Competency	7.000	1.000	4.868	9.132
Affective Competency	7.200	.490	6.156	8.244
Trainer Access	7.400	.980	5.311	9.489
Information Access	7.000	.894	5.093	8.907
Repetitions	7.200	.970	5.133	9.267
Tactile Feedback	7.400	.510	6.313	8.487
Service Life	5.600	.927	3.623	7.577
Distance Accessibility	4.400	1.208	1.824	6.976
Availability	9.200	.374	8.402	9.998

The first assumption of repeated measures ANOVA is most likely satisfied as all observations are independent of one another. The assumption of sphericity was violated, forcing the use of the Greenhouse-Geisser test. The result of the Shapiro-Wilk normality test is listed in Table 27 for each population. The assumption of normality is satisfied for most attributes but is not satisfied for 4 of 15 attributes.

Table 27 Question 10 Results – Shapiro-Wilk Normality Test

	Shapiro-Wilk Test		
	Statistic	df	Sig.
Cost	.871	5	.272
Immersion	.776	5	.050
User Safety	.684	5	.006
Schedule	.982	5	.945
Cognitive Competency	.902	5	.421
Psychomotor Competency	.963	5	.826
Interactive Competency	.999	5	1.000
Affective Competency	.684	5	.006
Trainer Access	.932	5	.607
Information Access	.905	5	.440
Repetitions	.951	5	.747
Tactile Feedback	.961	5	.814
Service Life	.739	5	.023
Distance Accessibility	.820	5	.117
Availability	.881	5	.314

Questions 11 and 12 asked respondents to rank the order of importance of cost types, then express which types they consider unimportant to analyze when selecting training alternatives. The results are listed in Table 28 and Table 29.

Table 28 Question 11 Results - Cost Type Rankings

		Rank					
Cost Type		Respondent 1	Respondent 2	Respondent 3	Respondent 4	Respondent 5	Mean
Direct	Personnel	1	2	2	2	1	1.6
	External Training Services	8	8	5	9	7	7.4
	Training Development and Instructional Material Prep	4	1	1	7	5	3.6
	Instructional Materials	5	4	4	6	3	4.4
	Equipment	2	3	3	3	2	2.6
	Facilities	3	5	6	1	4	3.8
	Travel	7	7	7	5	9	7
Indirect	Overhead	9	9	9	8	8	8.6
	General & Administrative	6	6	8	4	6	6

Table 29 Question 12 Results - Unimportant Cost Types

	Respondent 1	Respondent 2	Respondent 3	Respondent 4	Respondent 5
Unimportant Attributes	External Training Services		Travel		Travel
	Overhead		General & Administrative		
			Overhead		

Additionally, results from Question 11 are represented as a boxplot in Figure 16. The boxplot qualitatively indicates that personnel and equipment are the most important categories of cost to consider whereas overhead is ranked as the least important aspect of cost.

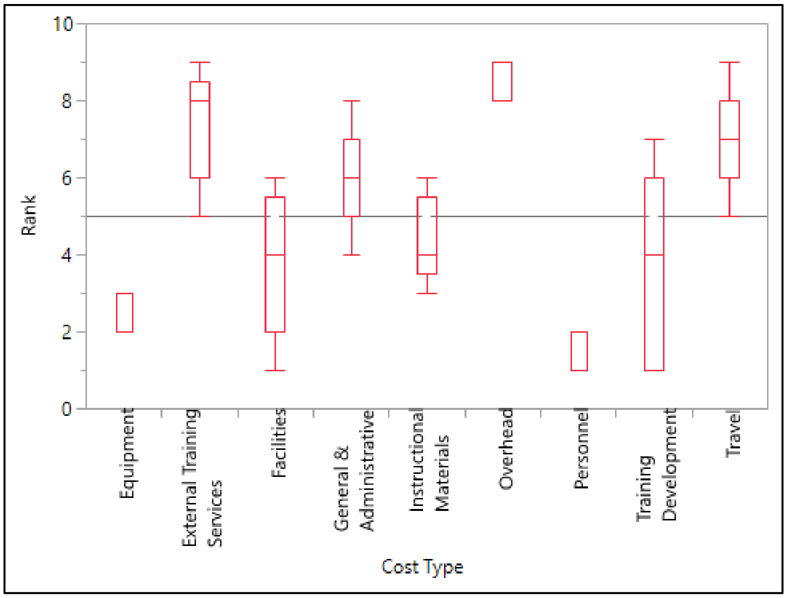


Figure 16 Question 11 Results - Boxplot of Cost Type Rankings

Create Value Functions

Questions 13 and 14 asked respondents to create a value function by rating the utility of five hypothetical training program costs, only with respect to cost. A comparison of responses to the expected value is listed in Table 30. The expected value was calculated by subtracting the normalized dollar amount across all training programs from 1.

Table 30 Question 13 Results - Utility Ratings of Cost

Alternative	Cost ($)	Expected Value (0.00-1.00)	Respondent 1	Respondent 2	Respondent 3	Respondent 4	Respondent 5
Training Program #1	100,000	1	0.90	1.00	1.00	1.00	1.00
Training Program #2	200,000	0.75	0.70	0.80	0.70	0.70	0.75
Training Program #3	300,000	0.5	0.50	0.60	0.40	0.50	0.50
Training Program #4	460,000	0.1	0.20	0.50	0.10	0.10	0.10
Training Program #5	500,000	0	0.10	0.00	0.00	0.00	0.00

All respondents constructed a monotonically decreasing value function, however only one of five respondents matched all expected values. Figure 17 graphically displays the cost value function results.

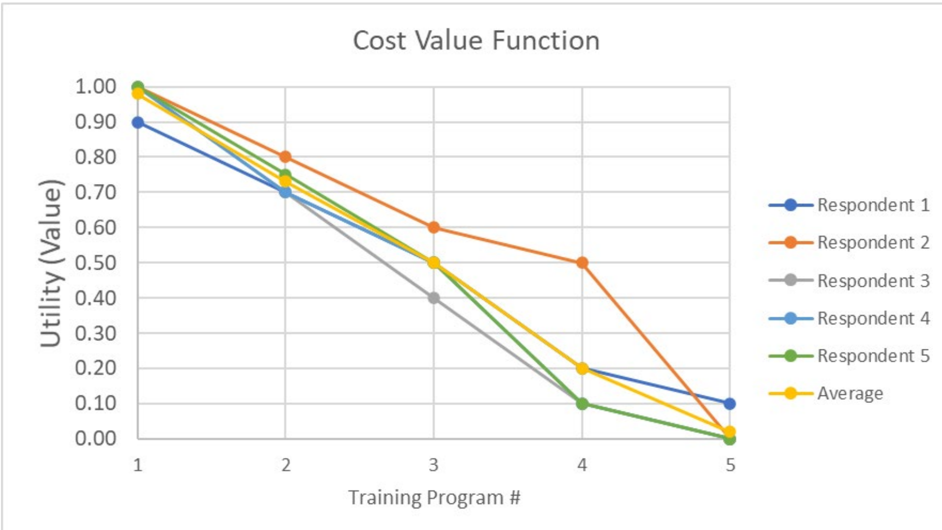


Figure 17 Question 14 Results - Cost Value Function

Question 15 asked respondents to build value functions for the cognitive, affective, and psychomotor domains based on each module goal. An objective level of competency was highlighted for each domain for each module goal, as listed in Table 31.

Table 31 Module Goals with Domain Objectives

Module	Module Goal	Objective Level of Competency	
1	Identify that a torque wrench is necessary to attach a panel.	Cognitive	Remember
		Affective	Responding
		Psychomotor	Perception (Awareness)
2	Be able to identify and use the proper torque wrench successfully given a verbal prompt by the instructor.	Cognitive	Apply
		Affective	Responding
		Psychomotor	Guided Responses
3	Recognizing improper function of a torque wrench during use and modifying it back to proper functionality through troubleshooting/repair.	Cognitive	Evaluate
		Affective	Characterizing
		Psychomotor	Adaptation

Respondents were told to imagine various nondescript training alternatives that could train to a separate level of competency within each domain. (Ex. What if a training alternative could only teach to the level of Remember? What value would you assign to that training alternative given the module goal?) The associated cognitive level is listed in Table 32 to accompany each cognitive value function in Figures Figure 18, Figure 19, and Figure 20.

Table 32 Cognitive Levels of Competency

Cognitive Level	0	1	2	3	4	5	6
	None	Remember	Understand	Apply	Analyze	Evaluate	Create

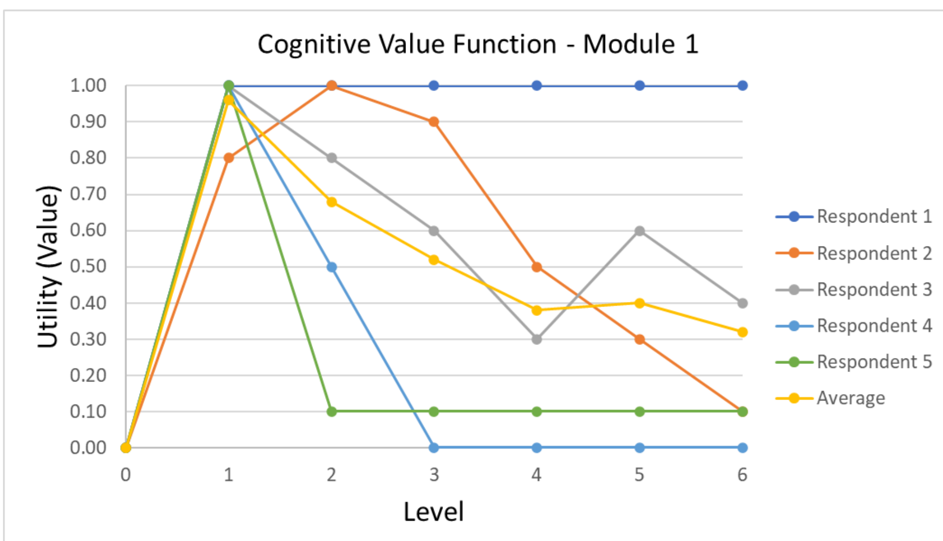


Figure 18 Cognitive Value Function - Module 1

For module 1 in the cognitive domain, four of five respondents agreed that the highest utility value was “Remember” given the module goal. Respondent 2 assigned the level of “Understand” as the highest value competency with respect to the module goal.

Additionally, four of five respondents assigned a lesser value to training targeted at the next highest level of competency compared to the level they first assigned a value of 1.

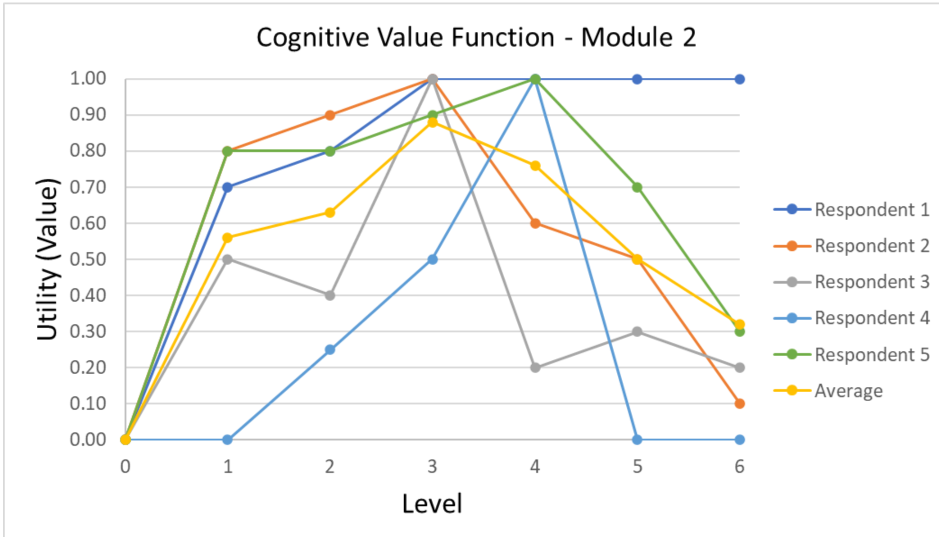


Figure 19 Cognitive Value Function - Module 2

For module 2 in the cognitive domain, three of five respondents agreed that the highest utility value was "Apply" given the module goal. The remainder of respondents expressed the level of "Analyze" as their sole level of highest value. Again, four of five respondents assigned a lesser value to training targeted at the next highest level of competency compared to the level they first assigned a value of 1.

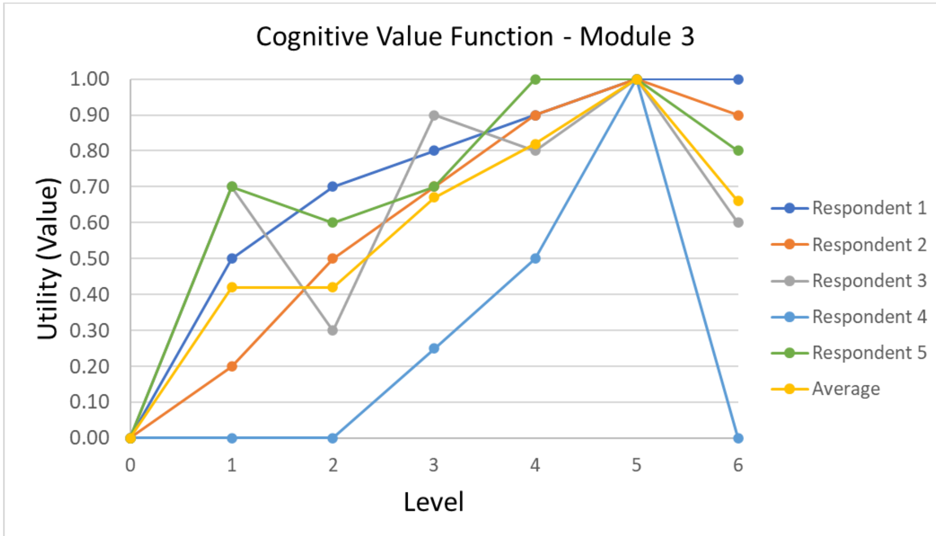


Figure 20 Cognitive Value Function - Module 3

For module 3 in the cognitive domain, four of five respondents agreed that the highest utility value was "Evaluate" given the module goal. Respondent 5 expressed the same value in training to a level of "Analyze" or "Evaluate" with respect to the module goal. Thus, only three of five respondents assigned a lesser value to training targeted at the next highest level of competency compared to the level they first assigned a value of 1.

The associated affective level is listed in Table 33 to accompany each affective value function in Figures Figure 21, Figure 22, and Figure 23.

Table 33 Affective Levels of Competency

Affective Level	0	1	2	3	4	5
	None	**Receiving**	**Responding**	**Valuing**	**Organizing**	**Characterizing**

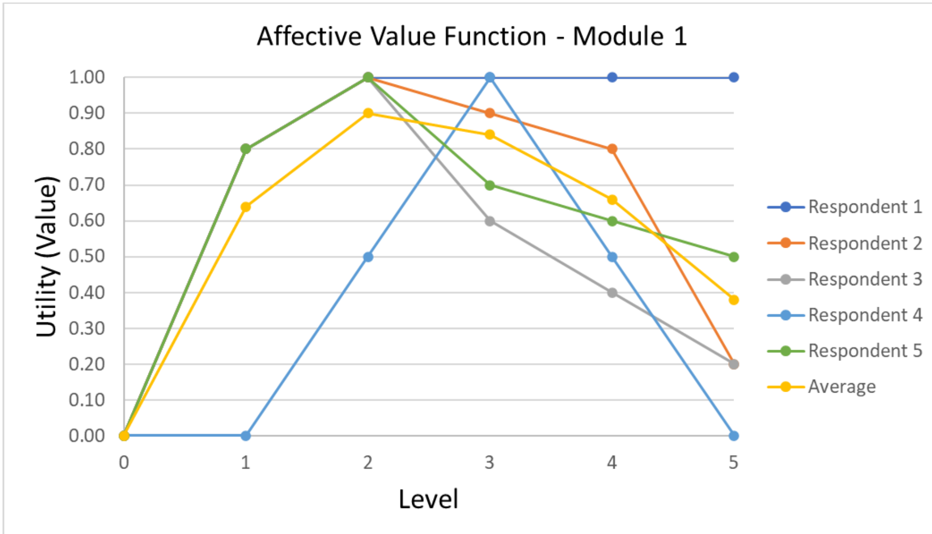


Figure 21 Affective Value Function - Module 1

For module 1 in the affective domain, four of five respondents agreed that the highest utility value was "Responding" given the module goal. Respondent 4 assigned the level of "Valuing" as the highest value competency with respect to the module goal. Four of five respondents assigned a lesser value to training targeted at the next highest level of competency compared to the level they first assigned a value of 1.

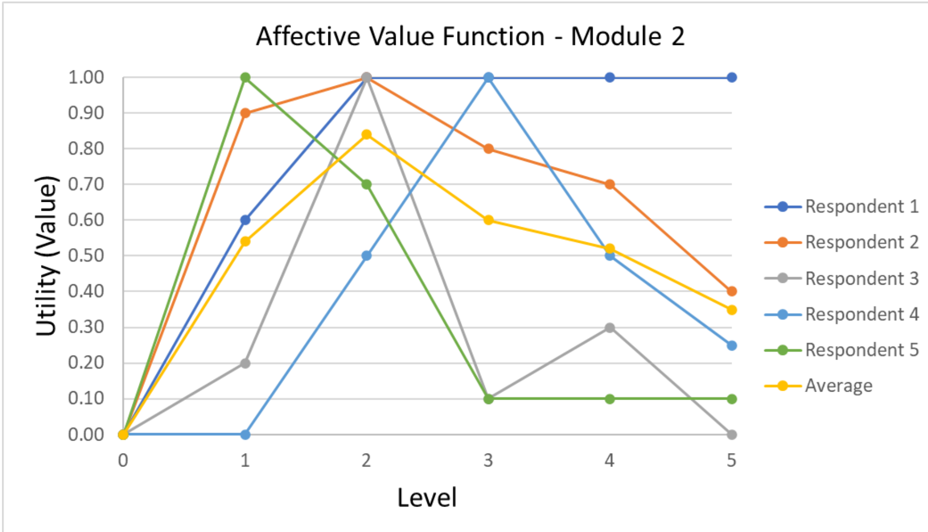


Figure 22 Affective Value Function - Module 2

For module 2 in the affective domain, three of five respondents agreed that the highest utility value was "Responding" given the module goal. Respondent 4 assigned the level of "Valuing" as the highest value, while Respondent 5 assigned a level of "Receiving" as the highest value competency with respect to the module goal. Four of five respondents assigned a lesser value to training targeted at the next highest level of competency compared to the level they first assigned a value of 1.

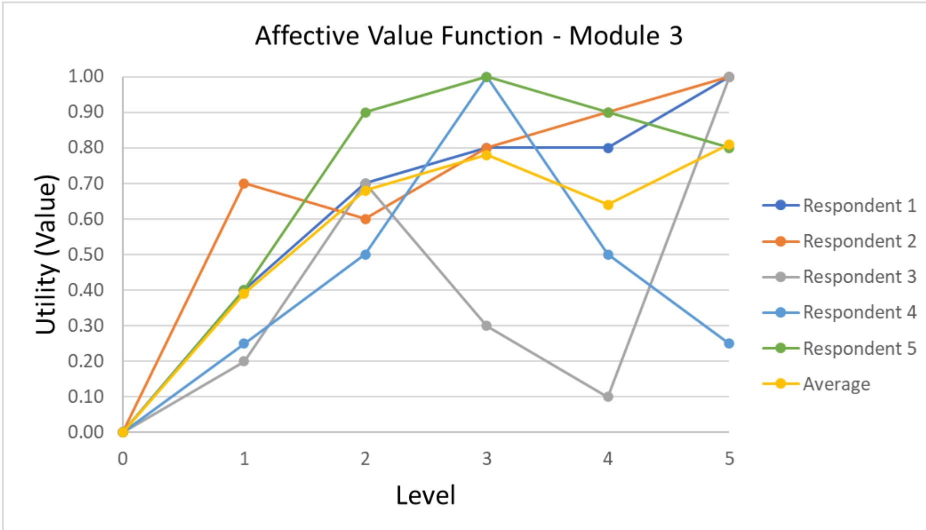


Figure 23 Affective Value Function - Module 3

For module 3 in the affective domain, three of five respondents agreed that the highest utility value was "Characterizing" given the module goal. Respondents 4 and 5 assigned the level of "Valuing" as the highest value competency with respect to the module goal. All respondents assigned only one level as having the highest value in the affective domain for module 3.

The associated psychomotor level is listed in Table 34 to accompany each psychomotor value function in Figures Figure 24, Figure 25, and Figure 26.

Table 34 Psychomotor Levels of Competency

Psychomotor Level	0	1	2	3	4	5	6	7
	None	**Perception (Awareness)**	**Set**	**Guided Responses**	**Basic Proficiency**	**Complex Responding/ Mastery**	**Adaptation**	**Origination**

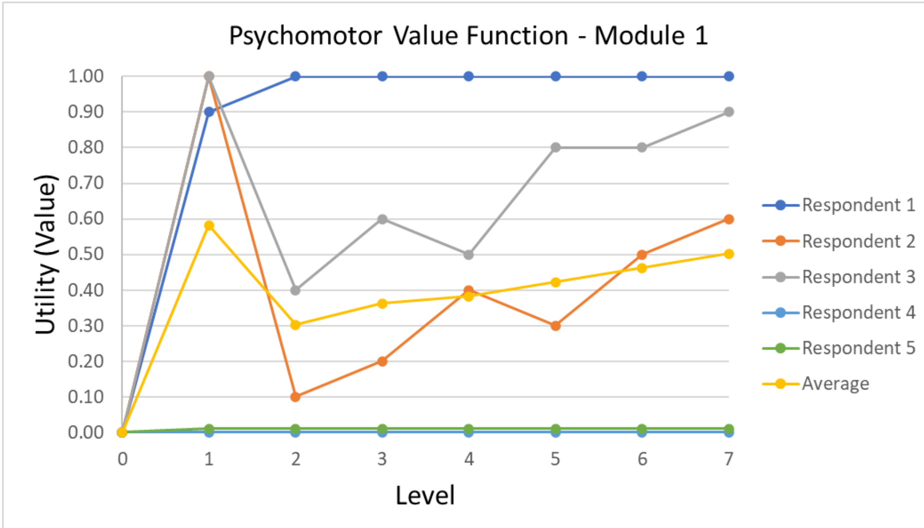


Figure 24 Psychomotor Value Function - Module 1

For module 1 in the psychomotor domain, two respondents assigned “Perception (Awareness)” the highest utility value, Respondent 1 assigned “Set” as the having the highest utility value, and the remaining respondents assigned all psychomotor competencies as having little or no value with respect to the module goal.

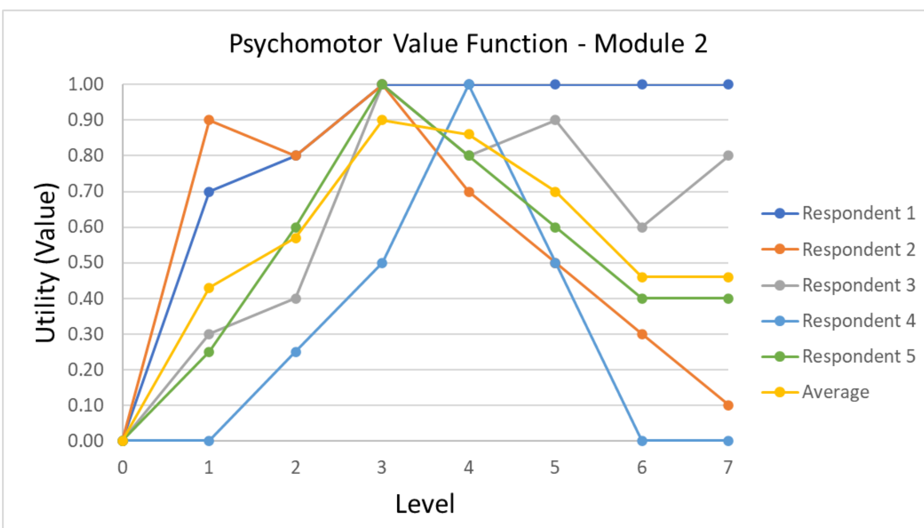


Figure 25 Psychomotor Value Function - Module 2

For module 2 in the psychomotor domain, four of five respondents agreed that the highest utility value was “Guided Responses” given the module goal. Respondent 4 assigned the level of “Basic Proficiency” as the highest value competency with respect to the module goal. Four of five respondents assigned a lesser value to training targeted at the next highest level of competency compared to the level they first assigned a value of 1.

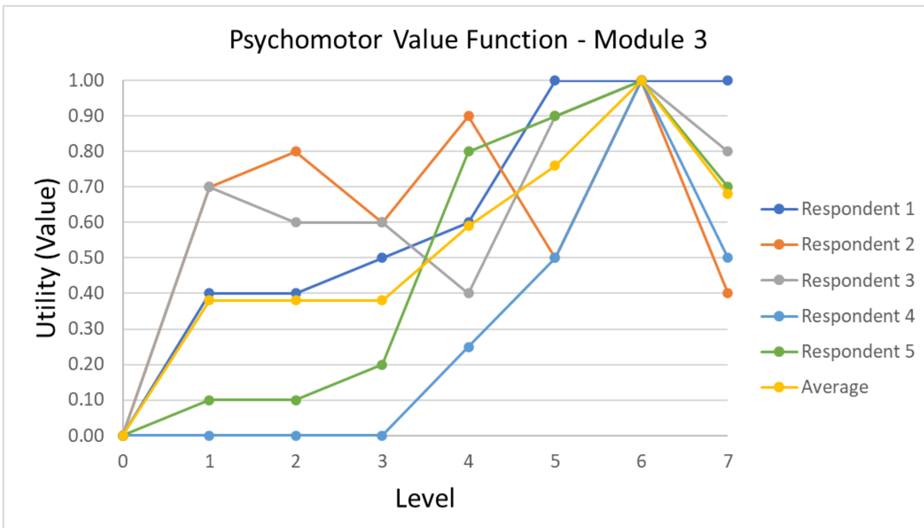


Figure 26 Psychomotor Value Function - Module 3

For module 3 in the psychomotor domain, four of five respondents agreed that the sole highest utility value was “Adaptation” given the module goal. Respondent 1 assigned the level of “Complex Responding/Mastery” as the highest value competency with respect to the module goal. Four of five respondents assigned a lesser value to training targeted at the next highest level of competency compared to the level they first assigned a value of 1. Throughout each domain and module goal, Respondent 1 set the utility value at 1 for every higher competency level than that which was first given a value of 1.

The average trendlines for each LO value function qualitatively reflect different trends than the analogous value functions within the MPEET model, which assumes its LO value functions as proportionally increasing with higher domain level competencies. Instead, the results indicated that most trendlines initially increased, peaked at or around the objective level of competency, then decreased as domain competency levels increased.

Statistical analysis of the value rating responses compared the ratings at two levels: the objective competency and the high competency. The objective competency level varied with each of the nine treatments, according to the module goal. The high competency level did not vary according to module goal. The analysis was based on the median difference between high and objective values for each domain. Initially, paired t-tests and related-samples Wilcoxon signed rank tests were utilized, however the assumptions for each could not be satisfied. The sign test was used for analysis. Note that the suffix "Obj" in the variable stands for the objective competency level whereas "Hi" stands for the high competency level. The prefixes "Cog," "Aff," and "Psy" refer to the cognitive, affective, and psychomotor domains, respectively. Descriptive statistics are listed in Table 35.

Table 35 Descriptive Statistics for Objective and High Competency Levels

Descriptive Statistics					
	N	Mean	Std. Deviation	Minimum	Maximum
CogObj	15	.9467	.13558	.50	1.00
AffObj	15	.8500	.24713	.25	1.00
PsyObj	15	.8273	.35786	.00	1.00
CogHi	15	.4333	.40649	.00	1.00
AffHi	15	.5133	.40641	.00	1.00
PsyHi	15	.5473	.38046	.00	1.00

Figure 27 is a histogram of utility value ratings for the high competency level and Figure 28 is a histogram of utility value ratings for objective competency level. The objective competency level has more utility ratings at a value of 1 compared to other values than the high competency level.

Table 36 gives the summary statistics for the high and objective competency level value ratings. The median ratings indicate that the objective competency level was rated highest, and the high competency was rated lower than the objective competency level.

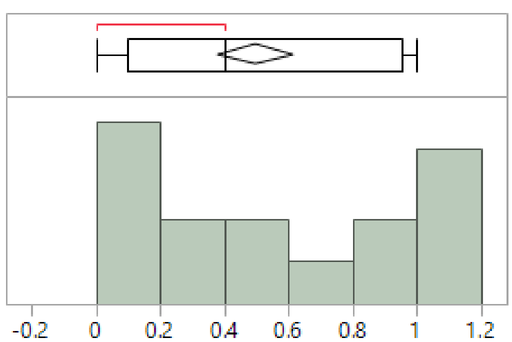


Figure 27 Histogram of High Competency Level Responses

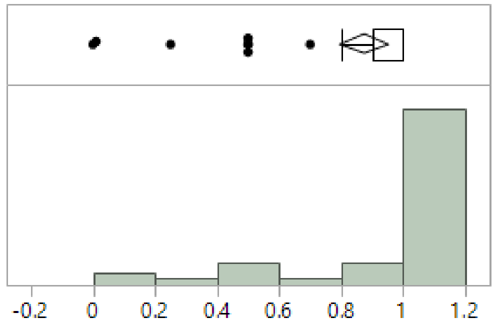


Figure 28 Histogram of Objective Competency Level Responses

Table 36 Summary Statistics for High and Objective Values

Summary Statistics		
	High	**Objective**
Mean	0.4980	0.8747
Std Dev	0.3918	0.2622
Std Err Mean	0.0584	0.0391
Upper 90% Mean	0.5961	0.9403
Lower 90% Mean	0.3999	0.8090
N	45	45
Median	0.4	1
Interquartile Range	0.85	0.1

Results of the frequencies of differences between high and objective competency levels for each domain are listed in Table 37. Frequencies indicate mostly negative differences between the high competency and objective competency level responses. Negative counts mean that the objective competency was assigned a higher value than the high competency level by the same respondent. At least one tie is present in all treatments. Ties indicate that the high competency and objective competency levels were scored identically by the same respondent.

Table 37 Frequency of Differences

Frequencies		
		N
CogHi - CogObj	Negative Differences	12
	Positive Differences	0
	Ties	3
	Total	15
AffHi - AffObj	Negative Differences	8
	Positive Differences	0
	Ties	7
	Total	15
PsyHi - PsyObj	Negative Differences	10
	Positive Differences	1
	Ties	4
	Total	15

The sign test was analyzed for each difference, with the results displayed in Table 38. The results lead to a rejection of the null hypothesis at an alpha level of 0.1 in each case. The assumptions for the sign test are satisfied due to the presence of continuous data, the use of independent matched pairs, and respondents using a continuous value scale.

Table 38 Sign Test Results for Value Difference

Test Statistics[a]			
	CogHi - CogObj	AffHi - AffObj	PsyHi - PsyObj
Exact Sig. (2-tailed)	.000[b]	.008[b]	.012[b]
a. Sign Test			
b. Binomial distribution used.			

After creating each value function, respondents were asked to provide reasoning why the highest competency level in each domain was or was not assigned a value of 1. The results are listed in Table 39.

Table 39 Question 16 Results - Value Function Reasoning

Respondent	Reasons
1	You are assigning a value of 1 "to the most relevant [competency level] to what you're trying to achieve."
2	"If you are going above that objective, you are kind of over-teaching, so you could lose [the students]"
3	"A lot of the top I didn't really care too much if the student or the trainee would reach that level [of competency] … Those ones aren't always important. It seems like the other ones in the middle are real [sic] set foundations that they needed to understand before they got to that level"
4	"It depends on what the level of training is required. The wording of the [module goal] is what determines the level of competency... because every level you go up takes more time. If it's not required of them to do the job, you're over-teaching, and over-teaching can often lead to confusion"
5	*Respondent:* "…I felt that the modules that you were trying to teach could be accomplished by a lower level or higher level depending on where I was rating it at." *Interviewer:* "So, basically the module goal was dictating that value level, correct?" *Respondent:* "Correct."

All respondents generally expressed the idea that the purpose behind assigning a utility value of 1 was based on the goal they were trying to achieve. Respondents 2 and 4 explicitly mentioned the idea of "over-teaching" as influencing their responses. Their value curves reflect a decrease in utility value for the next highest level of competency after their value assignment of 1 within each domain.

Because the interactive domain was also reviewed in the literature portion, respondents were asked how they would build value functions. Two options were given, either assign a hierarchical order to the interactive domain or evaluate each level as

needed. The hierarchical order example looked like the cognitive, psychomotor, and affective domains, where all levels were evaluated on the same value curve. The alternative example included separate hypothetical value curves for each interactive domain level. The results are listed in Table 40.

Table 40 Question 17 Results - Interactive Domain Functions

Answer	Respondent 1	Respondent 2	Respondent 3	Respondent 4	Respondent 5
A) Increasing Order of Skill				X	X
B) Evaluated as needed	X	X	X		X

Respondents 1, 2, and 3 expressed that each level within the interactive domain should be evaluated on an individual basis as needed, assuming no hierarchical order between the levels. Respondent 4 assigned an order to the levels, indicating an ordering of Seeking/Giving Information, Including, Summarizing, Supporting, Proposing, then Disagreeing (least to highest level). Respondent 5 assigned an order to some levels and determined that Including and Disagreeing should be evaluated as needed.

Alternative Generation

The remainder of the interview questions placed heavy emphasis on training methods, technologies, and their relation to overall value to gather information that could facilitate the development of training alternatives. Question 18 asked for respondents to give additional categories that were not presented in the interview. The results are listed in Table 41.

Table 41 Question 18 Results - Additional Methods and Technologies

Respondent	Missing Categories	Description	Method	Technology
1	Field Trips	Observing a real-life performance in a real environment.	X	
2	Field Experience	"[the students are] going out and… they are developing a project based off of the field experience they had"	X	
4	Teaching Interview	Conversation with a subject matter expert, similar to a panel.	X	
5	Self-Paced	The trainee holds the responsibility of observing, interacting with, and pacing through training content. This is typically subject to an instructor-imposed deadline.	X	
	Observation	Watching a performance without the intent to teach on the part of the performer.	X	
	Audiovisuals (Videos)	Audiovisual delivery of training content in the form of conventional videos, 360-degree videos, and 3-dimensional videos.		X
	Remote Operations	Real-life performance that is accomplished through remote control such as teleoperation surgeries or piloting drones.	X	X

Question 19 asked respondents to indicate whether a technology is possible to use when employing a specific training method and enter a "1" in the corresponding row and column. A "0" was entered if a technology was not able to be used with a training method. The sum of each cell for all respondents is presented in Table 42.

The technologies that were unanimously selected by respondents as being possible to employ for each method are represented in dark green with a number 5. After respondents indicated possible technologies for use during a training method, each selected the best combination of technologies for use in a single training method. The results are presented in Table 43. Numbers entered in each cell correspond to how many respondents selected each technology as part of a best combination for a single training method. For example, an entry of "0" means that none of the respondents selected the technology as part of a best combination.

Table 42 Question 19 Results - Possible Training Technologies for Use with Methods

	Training Technology												
	A	B	C	D	E	F	G	H	I	J	K	L	M
Training Method	Chalkboard/ Whiteboard	Oral Description	Augmented Reality	Virtual Reality	PowerPoint Slides	Computer-based lessons	Interactive computer courseware	Simulator	Models	Live artifacts	Demo Hardware	Technical Information	Workbooks/ Handouts
Lecture	5	5	2	2	5	1	2	1	4	4	4	5	5
Demonstration	4	3	5	4	4	5	5	5	5	5	5	3	3
Exhibit	3	4	2	2	5	3	4	3	5	5	4	3	2
Questioning	3	5	3	3	5	4	3	2	3	3	2	3	5
Seminar	3	5	1	2	5	1	3	2	4	3	3	2	2
Discussion	4	5	2	2	5	4	4	1	4	4	4	4	4
Performance	2	3	4	5	2	3	5	5	4	5	5	4	2
Case Study	3	5	1	3	5	5	3	1	2	2	2	2	5
Real Life Performance	0	3	2	2	1	1	3	4	2	4	3	3	1
Indirect Discourse	1	5	1	2	2	2	0	1	1	3	2	3	3
Assigned Reading	2	1	2	1	4	3	2	0	0	1	0	4	5
Small Group Method	3	5	2	3	4	3	3	3	3	3	2	2	3
Student Query	2	4	1	2	3	5	4	2	1	2	1	4	3

Table 43 Question 19 Results - Best Combinations of Technologies Per Method

	Training Technology												
	A	B	C	D	E	F	G	H	I	J	K	L	M
Training Method	Chalkboard/ Whiteboard	Oral Description	Augmented Reality	Virtual Reality	PowerPoint Slides	Computer-based lessons	Interactive computer courseware	Simulator	Models	Live artifacts	Demo Hardware	Technical Information	Workbooks/ Handouts
Lecture	2	3	0	0	5	0	0	0	3	2	3	1	3
Demonstration	0	3	3	1	1	0	1	2	2	3	3	2	0
Exhibit	0	3	2	1	1	0	1	1	4	2	4	1	0
Questioning	2	5	1	0	2	2	2	1	3	2	1	1	2
Seminar	2	4	1	1	3	0	1	2	2	2	2	1	0
Discussion	3	5	1	1	4	1	1	1	2	2	1	1	1
Performance	0	2	3	2	1	0	1	2	2	3	3	2	0
Case Study	1	4	1	2	1	2	1	1	2	2	1	2	2
Real Life Performance	0	3	1	0	0	0	0	1	0	4	1	1	0
Indirect Discourse	1	5	0	1	0	0	0	1	0	2	0	2	1
Assigned Reading	0	0	1	0	1	2	0	0	1	0	0	2	4
Small Group Method	1	3	0	1	2	2	2	1	1	2	1	1	0
Student Query	1	3	1	1	1	2	3	1	0	1	0	3	2

For lectures, PowerPoint slides were unanimously selected as part of a best combination of training technologies. Demonstrations did not have a single technology that was unanimously selected as best, rather oral descriptions, AR, live artifacts, and demo hardware were each selected by three different participants. Four participants selected models and demo hardware as the best technologies for exhibits. Oral descriptions were unanimously selected by participants for use in questioning. Four

participants selected oral descriptions as part of a best technology combination for seminars. The best training technologies for discussions indicate a unanimous selection of oral descriptions with 4 of 5 respondents also choosing to employ PowerPoint slides. In the context of a performance, 3 of 5 participants selected AR, live artifacts, and demo hardware. Case study was the only method where at least one participant had selected each possible technology as best for use, however oral description was the choice for 4 of 5 respondents. Live artifacts were the best technology selection for real life performances for 4 of 5 respondents. For indirect discourse, oral descriptions were unanimously selected as part of a combination of best training technologies. Workbooks/handouts were part of a best combination of technologies for 4 of 5 participants during assigned readings. Most respondents selected oral description as the best training technology for the small group method. Finally, 3 of 5 respondents selected oral description, interactive computer courseware, and technical information for student query.

Question 20 asked respondents to provide a best estimate of value for each training technology according to the level of utility it provides in each attribute category. The utility was calculated across all participant responses by normalizing the attribute columns of each participant, then calculating the mean for each corresponding cell. The result is presented in Table 44.

Table 44 Question 20 Results – Mean Technology Value Per Attribute

		Attributes								
		Immersion	User Safety	Repetitions	Tactile Feedback	Service Life	Distance Accessibility	Trainer Access	Information Access	Availability
Training Technology	Chalkboard/ Whiteboard	3	98	48	1	80	11	98	23	82
	Oral Description	5	95	13	5	76	50	100	29	88
	Augmented Reality	75	49	56	65	30	23	47	45	13
	Virtual Reality	90	47	67	57	27	36	51	56	18
	PowerPoint Slides	19	98	48	6	66	80	46	47	72
	Computer-based lessons	29	87	30	20	55	89	37	66	56
	Interactive computer courseware	37	87	50	30	55	79	38	66	42
	Simulator	85	45	40	73	25	10	49	56	20
	Models	60	48	55	55	37	16	43	45	39
	Live artifacts	92	28	53	98	46	16	48	55	37
	Demo Hardware	76	36	50	65	41	10	48	37	44
	Technical Information	48	100	63	15	42	80	32	62	85
	Workbooks/ Handouts	28	100	62	19	43	87	30	51	94

Tool Formulation

This data was gathered from SME's who had experience creating training alternatives for technicians, so the organization of the results shall provide utility to maintenance organizations that want a means to solve problems caused by training in accordance with systems engineering best practices. The VFT framework is an established, accessible 10-step process that can be used to create training solutions which address organizational issues and compare the cost-utility of alternatives. The tool this research effort provides consists of Model-Based Systems Engineering (MBSE) representations of information which are presented in SysML, a list of relevant attributes, information applicable for creating value functions for LOs, and qualitative data helpful for constructing training alternatives. The following is the recommended process for incorporating the results that follows the numbering scheme of the VFT framework:

1. Aggregate problems from stakeholders into a shared BDD using MBSE software.

2. Create a BDD representing an objectives hierarchy. Then, add problems, tracing specific objectives to each problem. Edit objectives hierarchy so that it accounts for all problems.
3. Develop any additional evaluation measures, starting with the list of identified attributes.
4. Create value functions deterministically for each LO, highlighting the objective competency level. Create value functions for all relevant selected attributes.
5. Have decision-makers assign weights of importance to each attribute.
6. Classify the current training alternative by method and technologies used. Create new alternatives with the help of technology-method compatibility matrices and technology-attribute rating table.
7. Score each training alternative according to the Total Value Function in Equation (**2**).

This process produces initial value scores for training alternatives and accompanying diagrams to enable users and decision-makers to make a case for alternative training systems based on cost-utility. All diagrams should be revised as necessary to ensure the model reflects reality, and successful implementations should be documented to inform future work.

Tool Example

This section walks through a simplified application of the tool to an example problem using the VFT framework. Steps 1 thru 3 involve the use of MBSE software with input from relevant stakeholders. When creating diagrams using MBSE software, it is recommended that naming conventions of all folders and diagrams are consistent with organizational policy. This example begins with a scenario and module description in Table 45.

Table 45 Example Problem

Problem Scenario	You are a decision-maker in charge of training for a maintenance organization. Leadership has become aware that the current system used to train a particular module is about twenty years old, and there have been few updates to the training content since. It is suspected that this is a primary contributor to the poor performance exhibited during several of the last technician audits. Additionally, the current training module requires transporting students to a location that poses various environmental health hazards. Due to the importance and limited number of the live artifacts that are used during training, there is no guarantee that extra units will be available for training purposes without at least two months' notice. To make matters worse, a trainer on your staff has just retired, leaving your organization desperate for technical skill. The current training is administered via demonstration with oral descriptions, live artifacts, and a chalkboard. Leadership is open to pursuing alternative training systems with the hope that some of these problems can be solved.
Module Description	A system component on an aircraft is to be installed by fastening both hydraulic and pneumatic pressure connectors to the proper locations as listed in the procedure, otherwise known as the Technical Order (TO). The trainee needs to be able to follow the TO, select the correct tool to make each connection, and demonstrate proficiency in performing this task with some confidence.
Training System	Method – Demonstration Technologies – Live artifacts, oral description, chalkboard A demonstration is given to a class size of 12-15 trainees at an on-site location where the primary trainer shows how to properly install the system component, explaining the "Do's and Don'ts" associated with performing the task. Hand-drawn diagrams are provided as a reference on a chalkboard. A trainer who facilitates the module is also present.

In Step 1, problems are identified and recorded in a BDD, then problem stereotypes are assigned to each block. A problem BDD for the example scenario is listed in Figure 29.

In Step 2, once problems are recorded, objectives can be created in a separate BDD. Blocks are created for each objective, stereotyped as objectives, then organized into a hierarchy through generalization relationships. The objectives hierarchy BDD is listed in Figure 30.

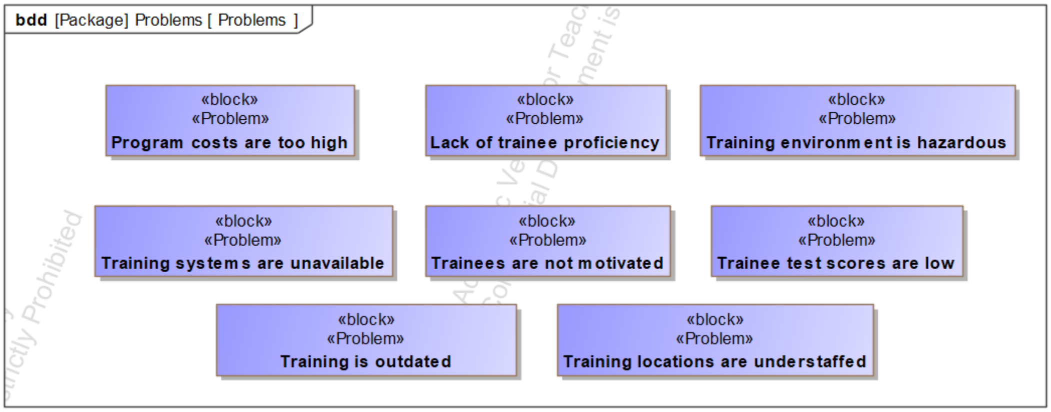


Figure 29 Step 1 Problem BDD Example

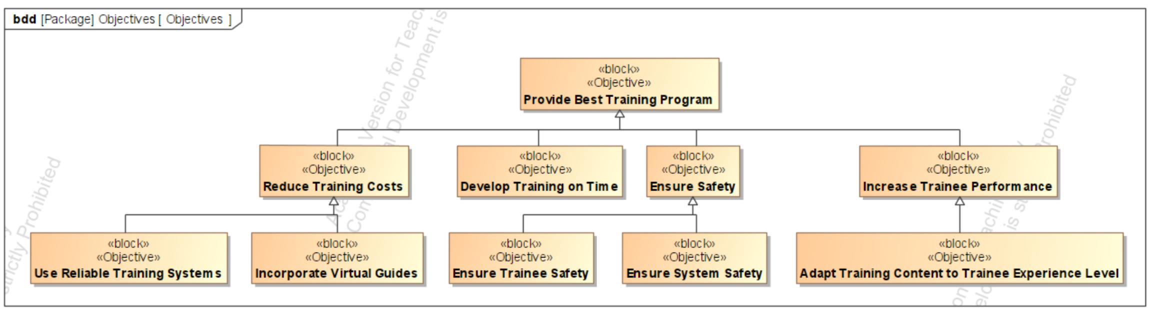


Figure 30 Step 2 Objectives Hierarchy BDD Example

To ensure that all problem areas are addressed, connect objectives within the hierarchy to each problem with trace relationships. Note that if there are problems that cannot be traced to an objective, then an objective needs to be added to the hierarchy to address the problem. An example is represented in Figure 31 where the problems "Training is outdated" and "Trainees are not motivated" cannot be traced from any existing objective, indicating the need for the addition of further objectives to the hierarchy.

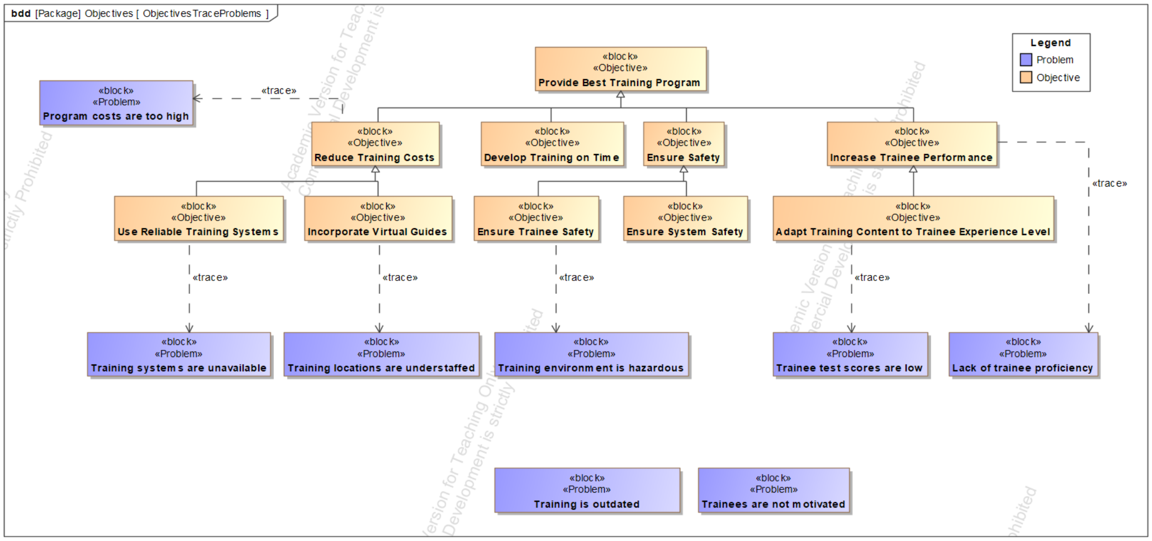


Figure 31 Step 2 Objectives Traced to Problems BDD Example

Figure 32 represents the additional objectives added to account for the untraced problems. “Modernize Training” and “Modernize Training Technologies” were objectives added to account for these problems.

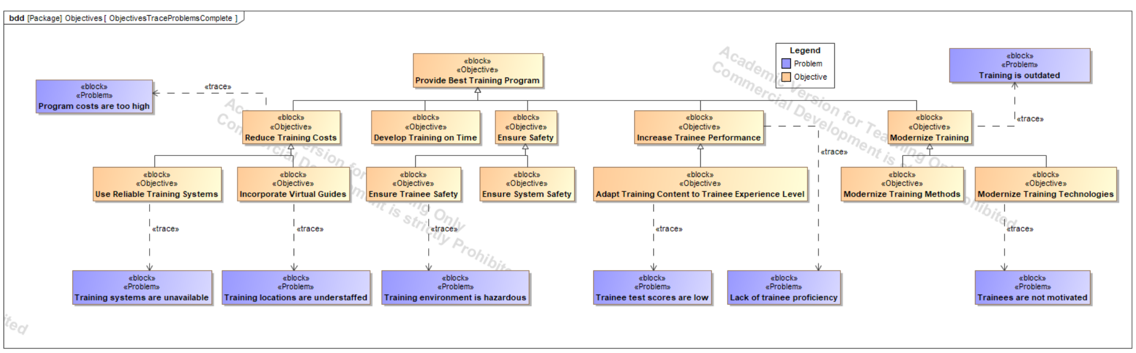


Figure 32 Step 2 Completed Objectives Traced to Problems Example

Step 3 begins by determining evaluation measures for each objective. For example, if creating an evaluation measure for the objective “Use Reliable Training Systems,” a user may reference an attribute from the initial list such as “Service Life” to evaluate the objective or create a new attribute for a different measure. The selected

attribute categories will compose the total value function. The initial list of attributes is recorded in Figure 33. Definitions of each added attribute may be recorded in a table or comments within the block specifications.

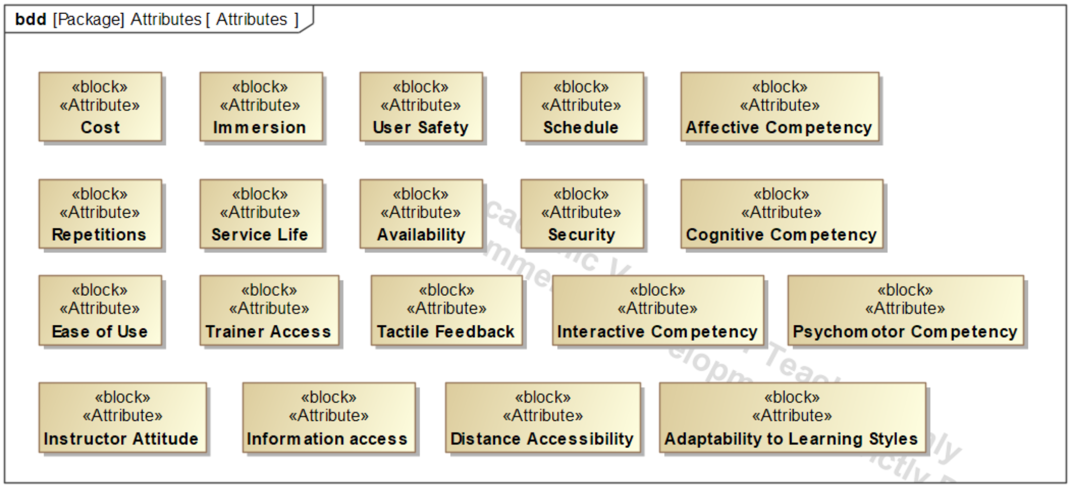


Figure 33 Step 3 Attributes BDD Example

The attributes and corresponding evaluation measures are incorporated with the problem-traced objectives and listed in Table 46. Note that the objective of "Modernize training" is broad and does not have a corresponding attribute or evaluation measure. Rather than create a new attribute to measure modernization, it may be addressed when generating training alternatives in Step 6.

Note that there is some freedom when establishing how to measure a training system alternative's utility with respect to each attribute. Use of metrics common to the Department of Defense (DoD) are preferred, such as materiel availability for the availability attribute, which is a measure of percentage of uptime of a system, or risk levels established in Air Force Instruction (AFI) 90-802 *Risk Management* (DAF, Air Force Instruction (AFI) 90-802, 2019). Ideally, an attribute should be utilized for only

one objective within the hierarchy to prevent dependencies between objectives. To fix this, objectives can be decomposed into more specific objectives, or attributes should be defined in more detail.

Table 46 Step 3 Evaluation Measures

Problem	Objective	Attribute	Evaluation Measure
Program costs are too high	Reduce training costs	Cost	Dollars ($)
Training systems are unavailable	Use reliable training systems	Availability	Materiel Availability (% uptime)
Training locations are understaffed	Incorporate virtual guides	Trainer Access	Trainer Configuration (Discrete values)
Training environment is hazardous	Ensure trainee safety	User Safety	Risk level (AFI 90-802)
Trainee test scores are low	Adapt training content to trainee experience level	Adaptability to learning styles	Adaptability level (Discrete values)
Lack of trainee proficiency	Increase trainee performance	Cognitive Competency	Cognitive Levels
		Affective Competency	Affective Levels
		Psychomotor Competency	Psychomotor Levels
		Immersion	SME-provided technology utility Ratings
Trainees are not motivated	Modernize training technologies	Instructor Attitude	Preference Selection
Training is outdated	Modernize training	N/A	N/A

For Step 4, utility functions are created for each attribute category. It is recommended that the output of the utility functions are on a scale from 0 (least valuable) to 1 (most valuable). Utility functions may be created multiple ways, such as specifying an equation, assigning utility values to discrete levels, or ranking the order of alternatives. A utility function for ranking order is shown in (3).

$$Value\ by\ Rank = 1.0 - \left(\frac{1.0}{\#\ of\ Alternatives}\right) * (System\ Rank - 1) \tag{3}$$

Assigning a utility value by rank is a quick means of value calculation that can be used for many attributes, but it is subjective in that it is dependent upon decision maker

preference rather than performance. A list of attributes and their recommended equations are presented in Table 47. For attributes that do not have a recommended equation, some may be evaluated according to technology-attribute utility values presented in Table 44. Stakeholders within the organization should develop evaluation measures and utility functions for attributes that are well-understood, relevant to accomplishing the module, and representative of the expected performance requirements.

The variables within Table 46 are mostly self-explanatory. The variables for the cognitive, affective, and psychomotor domains are more difficult to understand but are based on the information in Table 32, Table 33, and Table 34 for the domain levels. The variable "Level" refers to the competency level that a training system can train provide to a student, designated by an integer corresponding to the levels within each domain. "ObjLevel" reflects the skill within each domain that is required to successfully attain the module goal, also designated by an integer. "HiLevel" refers to the integer corresponding to the highest level within a domain. For example, there are 6 levels within the cognitive domain, so the HiLevel is equal to 6. "HiVal" refers to a utility value between 0 and 1 which corresponds to the utility value of training the highest competency level for a given module.

Table 47 Attribute Utility Functions

Attribute	Equation
Cost	$Cost = 1.0 - \frac{Cost}{Available\ Funds}$
Immersion	Primary technology attribute rating or Value by rank
User Safety	$User\ Safety = \begin{cases} 1.0, & Risk = Low \\ 0.67, & Risk = Medium \\ 0.33, & Risk = High \\ 0, & Risk = Extremely\ High \end{cases}$
Schedule	$Schedule = 1.0 - \left(\frac{Weeks\ of\ development}{Maximum\ allotted\ weeks}\right)$
Cognitive Competency	$Cognitive = \begin{cases} \frac{Level}{ObjLevel}, & Level \le ObjLevel \\ 1.0 - \frac{(Level - ObjLevel)(1.0 - HiVal)}{HiLevel - ObjLevel}, & Level > ObjLevel \end{cases}$
Affective Competency	$Affective = \begin{cases} \frac{Level}{ObjLevel}, & Level \le ObjLevel \\ 1.0 - \frac{(Level - ObjLevel)(1.0 - HiVal)}{HiLevel - ObjLevel}, & Level > ObjLevel \end{cases}$
Psychomotor Competency	$Psychomotor = \begin{cases} \frac{Level}{ObjLevel}, & Level \le ObjLevel \\ 1.0 - \frac{(Level - ObjLevel)(1.0 - HiVal)}{HiLevel - ObjLevel}, & Level > ObjLevel \end{cases}$
Interactive Competency	Value by rank for each relevant competency category
Trainer Access	$Trainer\ Access = \begin{cases} 1.0, & Real\ Trainer \\ 0.67, & Virtual\ Trainer \\ 0.33, & Virtual\ Guide \\ 0, & No\ Trainer \end{cases}$
Information Access	Primary technology attribute rating or Value by rank
Repetitions	$Repititions = \frac{Repetitions\ delivered\ by\ solution}{Desired\ \#\ of\ repititions}$
Tactile Feedback	Primary technology attribute rating or Value by rank
Service Life	Primary technology attribute rating or Value by rank
Distance Accessibility	Primary technology attribute rating or Value by rank
Availability	Primary technology attribute rating or Value by rank
Security	Value by rank
Ease of use	Value by rank
Instructor Attitude	Value by rank
Adaptability to Learning Styles	Value by rank

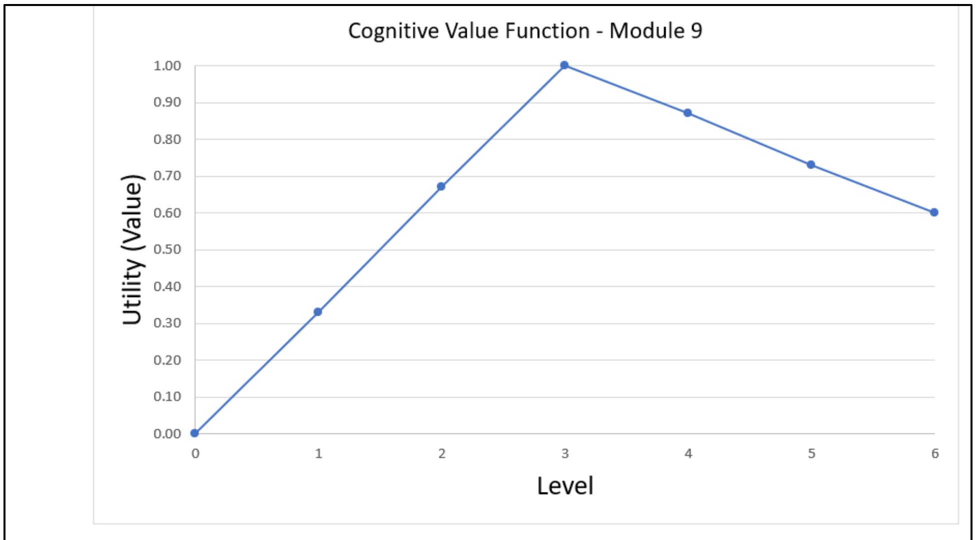
To illustrate how a module is characterized, first, the objective competency levels for included LO domains need to be identified for utility function creation. An example

of recording objective competency levels for the training module is represented in Table 48. Again, the objective competency is dependent on the module goal.

Table 48 Objective Competency Level Example

<table>
<tr><th>Module</th><th>Module Goal</th><th colspan="2">Objective Competency Level</th></tr>
<tr><td rowspan="3">1</td><td rowspan="3">Fasten all hydraulic and pneumatic connections for component install.</td><td>Cognitive</td><td>Apply</td></tr>
<tr><td>Affective</td><td>Responding</td></tr>
<tr><td>Psychomotor</td><td>Basic Proficiency</td></tr>
</table>

It is recommended that value functions for the LO domains shall be created deterministically from SME input for more thorough evaluations. However, to obtain a rough utility function for any of the LOs, the equations in Table 47 are sufficient. The equations for the cognitive, affective, and psychomotor domains reflect the qualitative trend of responses uncovered by research question two. When identifying variables for the cognitive utility function, refer to Table 32. The highest level in the cognitive domain has a corresponding integer of 6. This will be the value for the “HiLevel” variable. If the training alternative can train to the highest competency level, the utility value to the organization is 0.6. This is the “HiVal” variable, which is selected by the training organization. The value for “Level” is the competency level that the training alternative can train to. For the cognitive domain, “Level” will be an integer from zero to six as there are six levels in the domain. Zero indicates the training alternative is ineffective at training to any skill level within the domain. Finally, “ObjLevel” is equal to 3 because the level of Apply is indicated in Table 48 as the necessary skill level to perform the module goal successfully. The resulting utility function from these inputs is represented in Figure 34.



Level	0	1	2	3	4	5	6
	None	Remember	Understand	Apply	Analyze	Evaluate	Create
Value	0.00	0.33	0.67	1.00	0.87	0.73	0.60

Figure 34 Value Function Example

Step 5 is accomplished by recording the percent weight for each attribute, as expressed by the decision-maker. Table 49 illustrates the attributes and assigned weights of importance for this example. The weights sum to a value of 1.

Table 49 Attribute Weights Example

Attribute	Weight of Importance (*w*)
Cost	0.25
Availability	0.10
Trainer Access	0.05
User Safety	0.10
Adaptability to Learning Styles	0.05
Cognitive Competency	0.15
Affective Competency	0.05
Psychomotor Competency	0.15
Immersion	0.05
Instructor Attitude	0.05
Sum (Total Value)	**1.00**

In Step 6, training alternatives are created by choosing one training method and one to many training technologies from a list. The best training technology list represented in Table 50 may be used as a starting point for creating training alternatives. The items in this list were derived by referencing values of 3 or greater from Table 43. However, VR was added to the Demonstration and Performance categories to be sure that systems using this technology could be evaluated.

Table 50 Best Training Technology List

Training Method	Training Technology
Lecture	Oral Description, PowerPoint Slides, Models, Demo Hardware, Workbooks
Demonstration	Oral Description, AR, VR, Live Artifacts
Exhibit	Oral Description, Models, Demo Hardware
Questioning	Oral Description, Models
Seminar	Oral Description
Discussion	Oral Description, Power Point Slides
Performance	AR, VR, Live Artifacts, Demo Hardware
Case Study	Oral Description
Real Life Performance	AR, Live Artifacts, Demo Hardware
Indirect Discourse	Oral Description
Assigned Reading	Workbooks/Handouts
Small Group Method	Oral Description
Student Query	Oral Description, Interactive Computer Courseware, Technical Info

Additionally, Table 44 can be used to help select training technologies according to the estimated utility for the desired attribute. By plotting the columns of Table 44 (attribute utility values) pairwise, technologies that have the highest utility value across multiple categories can be visualized. A simplified feasible trade space is illustrated in the light orange region for the user safety and immersion attributes in Figure 35. It is approximated by the straight red lines connecting the outermost points from the origin

and lines defining the highest utility value for each attribute. Each point on the plot represents a different training technology.

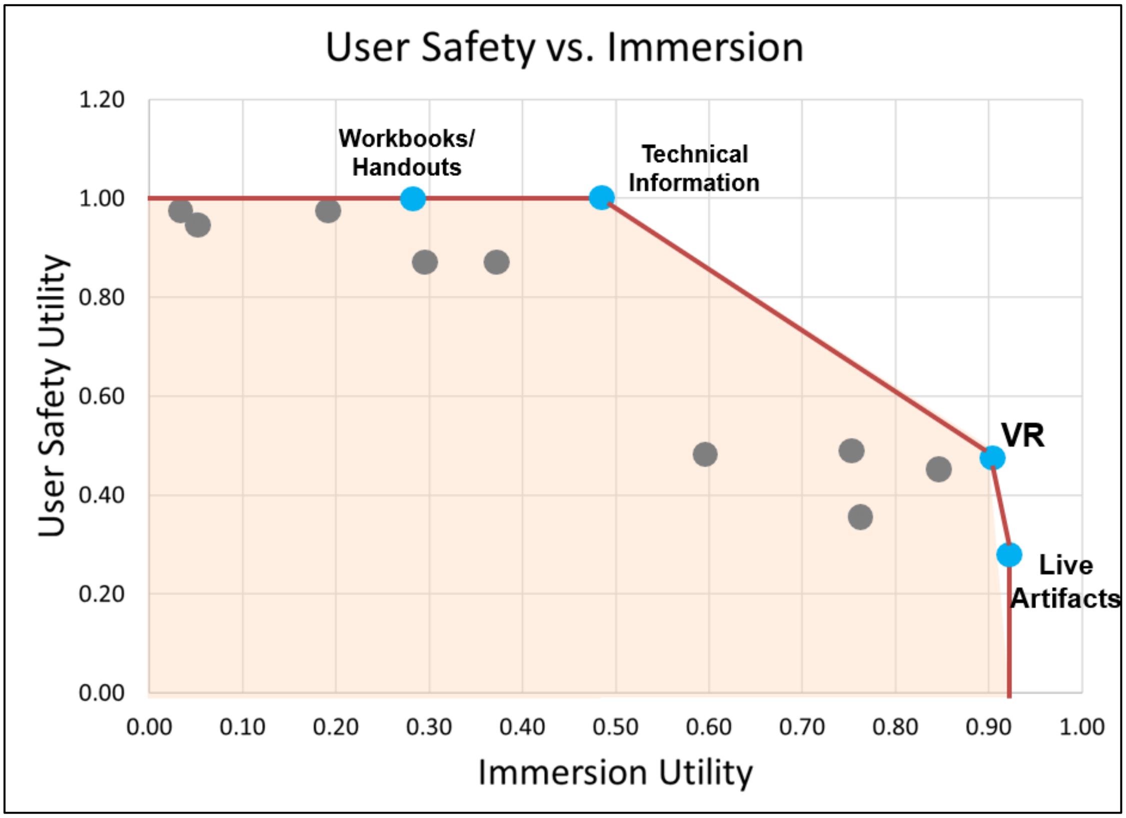


Figure 35 Utility Visualization Example

The technologies represented in blue lie on the perimeter of the feasible technology trade space. Based on the plotted SME utility ratings, the incorporation of VR into a training alternative may enable greater user safety with only a slight decrease in immersion compared to live artifacts, which are used in the original training system. It is apparent from this result that it is important to define the original training system composition to enable the estimation of performance gains in terms of utility.

Since the utility chart for immersion and user safety revealed VR as a training technology that can possibly increase user safety, it might be worthwhile to measure the

cost-effectiveness of an alternative that uses VR. Referencing the objectives created in Step 2 also reveals the need to incorporate virtual guides, which is also supported by VR technology. Since trainees are only learning through observation in the original training system, individual skill level may be increased by virtual performance. Thus, the training alternatives selected for this example are presented in Table 51.

Table 51 Training System Alternatives Example

Original Training System	Training Alternative-VR
Training Method Demonstration **Training Technologies** Live Artifacts, Oral Description, Chalkboard	**Training Method** Performance **Training Technologies** Virtual Reality

The anticipated setup for Training Alternative-VR will only require one trainer onsite to facilitate operation of the training. The VR platform only requires one trainer due to the use of the virtual guide. For this alternative, there will be two VR headsets for simultaneous use in the existing training facility. The live feed of each headset will be projected for the trainees to see when they are waiting for their turn to perform. The trainees using the headset will be able to perform the skill with a virtual replica of the system, with the option of having virtual information lending assistance for performing the task. As for anticipated benefits, the amount of time allotted for the original module is just enough time to let each trainee have a turn attempting a virtual performance. The VR alternative saves the cost of travel, equipment maintenance, and a real trainer, resulting in only a slight cost increase due to development costs. A virtual guide is estimated to be equally as valuable as the trainer who facilitated the demonstration in the original training

system. The instruction team believes that the alternative training system will produce immediate performance gains but believes that the virtual guide will be limited in its adaptability to student learning styles.

Step 7 in the VFT process is to score each training alternative using the total value function in Equation (**2**) and compare the results. This requires finding the utility value of each alternative for all attributes, multiplying it by the corresponding attribute weight of importance, then taking the sum of the results. For this example, there are 10 attributes included in the value model. The equations and inputs to each attribute are listed in Table 52. Note that the equations for availability, trainer access, and immersion vary from the utility equations in Table 47 due to the evaluation measures that were assigned for these attributes.

The resulting utility value outputs and total value scores are shown in Table 53, where training alternative number two is the winner with a total value of 0.735.

Table 52 Attribute Inputs

		Inputs	
Attribute	**Equation**	**Original System**	**Alternative**
Cost	$Cost = 1.0 - \frac{Cost}{Available\ Funds}$	Cost = $40k Available Funds = $50k	Cost = $45k Available Funds = $50k
Immersion	Primary technology attribute rating (%)	92	90
User Safety	$User\ Safety = \begin{cases} 1.0, & Risk = Low \\ 0.67, & Risk = Medium \\ 0.33, & Risk = High \\ 0, & Risk = Extremely\ High \end{cases}$	Medium	Low
Cognitive Competency	$Cognitive = \begin{cases} \frac{Level}{ObjLevel}, & Level \le ObjLevel \\ 1.0 - \frac{(Level - ObjLevel)(1.0 - HiVal)}{HiLevel - ObjLevel}, & Level > ObjLevel \end{cases}$	Level = 2 ObjLevel = 3 HiVal = 0.6 HiLevel = 6	Level = 3 ObjLevel = 3 HiVal = 0.6 HiLevel = 6
Affective Competency	$Affective = \begin{cases} \frac{Level}{ObjLevel}, & Level \le ObjLevel \\ 1.0 - \frac{(Level - ObjLevel)(1.0 - HiVal)}{HiLevel - ObjLevel}, & Level > ObjLevel \end{cases}$	Level = 1 ObjLevel = 2 HiVal = 1 HiLevel = 5	Level = 2 ObjLevel = 2 HiVal = 1 HiLevel = 5
Psychomotor Competency	$Psychomotor = \begin{cases} \frac{Level}{ObjLevel}, & Level \le ObjLevel \\ 1.0 - \frac{(Level - ObjLevel)(1.0 - HiVal)}{HiLevel - ObjLevel}, & Level > ObjLevel \end{cases}$	Level = 2 ObjLevel = 4 HiVal = 0.9 HiLevel = 7	Level = 4 ObjLevel = 4 HiVal = 0.9 HiLevel = 7
Trainer Access	$Trainer\ Access = \begin{cases} 1.0, & 2\ Real\ Trainers \\ 1.0, & 1\ Real\ Trainer + 1\ Virtual\ Guide \\ 0.33, & 1\ Virtual\ Guide \\ 0, & No\ Trainer \end{cases}$	2 Real Trainers	1 Real Trainer + 1 Virtual Guide
Availability	Materiel Availability (% uptime)	70	90
Instructor Attitude	$Value\ by\ Rank = 1.0 - \left(\frac{1.0}{\#\ of\ Alternatives}\right) * (System\ Rank - 1)$	System Rank = 2	System Rank = 1
Adaptability to Learning Styles	$Value\ by\ Rank = 1.0 - \left(\frac{1.0}{\#\ of\ Alternatives}\right) * (System\ Rank - 1)$	System Rank = 1	System Rank = 2

Table 53 Alternative Scoring Example

		Original Training System		Training Alternative-VR	
Attribute	Weight of Importance (w)	Utility Value $U(x)$	$w\ U(x)$	Utility Value $U(x)$	$w\ U(x)$
Cost	0.25	0.20	0.05	0.10	0.025
Immersion	0.05	0.92	0.046	0.90	0.045
User Safety	0.10	0.67	0.067	1.00	0.10
Cognitive Competency	0.15	0.67	0.1005	1.00	0.15
Affective Competency	0.05	0.50	0.025	1.00	0.05
Psychomotor Competency	0.15	0.50	0.075	1.00	0.15
Trainer Access	0.05	1.00	0.05	1.00	0.05
Availability	0.10	0.70	0.07	0.90	0.09
Instructor Attitude	0.05	0.50	0.025	1.00	0.05
Adaptability to Learning Styles	0.05	1.00	0.05	0.50	0.025
Sum (Total Value)	**1.00**		**0.5585**		**0.735**

The alternative scoring table can be used to compare and quantify the difference between training systems within each attribute category. Larger differences between utility values indicate greater disparity in performance with respect to each attribute. In this example, the VR-based alternative is estimated to provide the largest performance gains in instructor attitude, affective competency, and psychomotor competency by a utility value difference of 0.5. Psychomotor competency is also tied for the second-highest attribute weighing, which indicates to the user of the tool that this attribute is a high contributor to the difference in total value between alternatives.

Summary

The interview proceeded through the VFT framework, gathering information to apply across the first seven steps. The results could be used to answer the three primary research questions, which were to find the value attributes that contribute to the total value of training, to determine if value is a function of LOs, and to find the best combination of training method and training technology. Data gathered over the course of five interviews with training SME's were analyzed to formulate decision analysis components that apply directly to steps within the VFT framework and can be represented through MBSE software.

V. Conclusions and Recommendations

Chapter Overview

This chapter presents the conclusions of the research effort with respect to three research questions. These conclusions resulted in the creation of tools that can produce an initial cost-utility function in accordance with the VFT process. The research effort was significant because it provided maintenance organizations with a means to evaluate and select the most effective training solution, supporting readiness of the warfighter through the development of proficient technicians. Both limitations of the research effort and recommendations for future research were also presented.

Conclusions

This research effort sought to answer three primary questions, each of which are reviewed.

RQ1: What value attributes contribute to the total value of training?

A list of training attributes was gathered and presented to SME's who were asked to review any missing attributes, then rate the initial list using a most-least rating method. SME review yielded six additional attributes, including adaptability of the training alternative to learning styles, retention, instructor attitude, fidelity of the training alternative, security, and ease of use. Of these attributes, the researcher would exclude retention and fidelity, as they may have dependencies with other attributes on the initial list. To track additional attributes in the future, a BDD was created in SysML to represent the factors included in the total value function. It is represented in Figure 36.

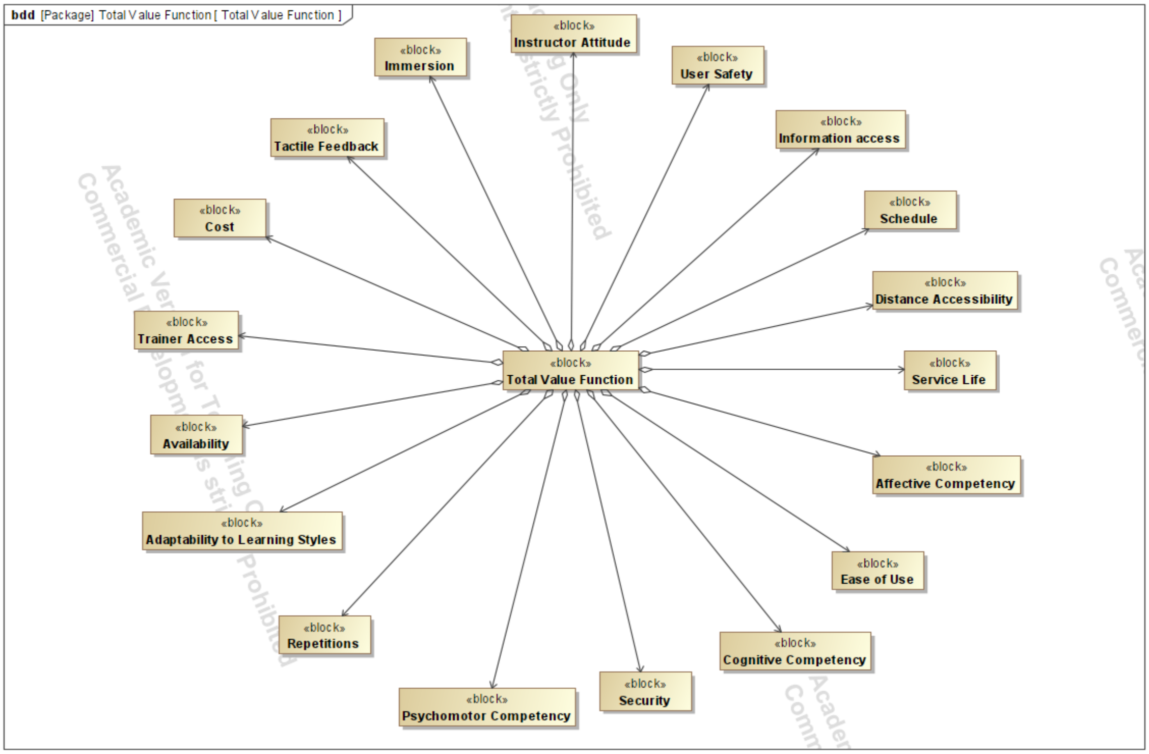


Figure 36 Total Value Function BDD

Each attribute is tied to the total value function by directed aggregation as opposed to directed composition because it implies that there may be additional attributes to consider for the function. The purpose of this diagram is to heighten visibility of the factors that contribute to the total value of a training alternative and to enable its users to make additions as necessary. The weight of importance of each attribute is to be determined by the user.

Additionally, rating the initial attribute list by importance produced an interesting result at an alpha level of 0.1. The Greenhouse-Geisser test led to a failure to reject the null hypothesis that the population means were equal with respect to the attribute rating. This result, in addition to the fact that no confidence intervals contained 0 for any attribute on the list, indicated that there was at least some level of importance for each

attribute. However, groupings of attributes into more important and less important categories could not be determined. Pairwise comparisons yielded no significance, perhaps due to the Bonferroni correction of alpha with 15 variables included in the model. User safety had the highest mean value while distance accessibility had the lowest mean value amongst participants. This result could reflect the in-person nature of maintenance work. Even though all attributes were given some importance, the goals of the module and the opinions of decision-makers within an organization may determine which attributes are ultimately included in a total value function model.

RQ2: Is value a function of the learning objective?

The results qualitatively indicate that value is a function of LOs. Initial analysis also suggests that value is dependent upon the module goal. Qualitative results to answer this research question were obtained by asking SME's to build value functions for the cognitive, affective, and psychomotor domains based on three module goals. Averages of each value function qualitatively indicate a maximum utility value at or near each objective competency level, rather than near the highest competency level for each domain. This suggests that participants based the highest utility value assignment on the module goal, rather than achieving the highest competency in each domain. This result may relate to the training system design best-practice of sequencing, where individual training modules are sequenced according to the objective competency level from lesser to greater skill levels. Thus, training alternatives are scored based on their attainment of the module goal within a module sequence, rather than the end-goal of a sequence. Interview participants expressed the idea of over-teaching once the competency level for

an individual module was exceeded, citing it as confusing for the trainee. Consequently, their value functions for each domain showed a decrease for the next highest competency level immediately after the objective.

Previous research had suggested that the highest competency level should be valued the highest when evaluating LOs for each domain, regardless of the module goal. To demonstrate that this was an incorrect assumption, the highest competency level responses were compared to objective competency level responses using the sign test. This analysis resulted in a rejection of the null hypothesis that the median difference between the two competency levels equals 0 in favor of the alternative. The median difference for the objective level ratings is statistically higher than the median difference for the high competency level ratings. This suggests that the objective competency may have more value than the high competency level, which is supported by the histograms of responses and the frequency with which the objective competency was assigned a higher utility value than the high competency level as shown in Table 37. The value function results point to two different schools of thought for creating value functions for LOs, both deterministic in nature. The first was exhibited by the average value function, which assigned a value of 0 for no competency, then trended upward to a peak utility value at or around the objective competency level, then decreased with higher competency levels. The second was consistently exhibited by Respondent 1, who rated every competency level at a value of 1 from the objective competency through the high competency level. The first may have been motivated by the module goal and the second may have been motivated by the construct of the learning domains, where achievement of higher

competency levels also means achievement of the levels below it. The resulting equation for LO value function creation is compatible with both trends.

RQ3: What are the best combinations of training method and training technology?

Results relating to the third research question were qualitative in nature, producing compatibility matrices with counts of SME responses as to which technologies are possible and which are best to use with each training method. The matrix for possible combinations indicates that oral descriptions and slide shows are the training aids that are most frequently suggested across training methods. The training method that had the largest number of training technologies to select from was demonstration. An interesting trend was visible when comparing the methods of performance and real-life performance. Respondents unanimously selected simulators, live artifacts, and VR as possible technologies to use with performance, while AR obtained four votes. For real-life performances, four SME's selected simulators and live artifacts as possible to use, while only two SME's selected AR and VR. This supports SME preference for physical equipment without virtual interfaces when performing tasks in the real world. Verbal responses across all respondents indicated that the current state of VR technology cannot yet provide the fidelity of immersion or experience that real-life scenarios offer. Not only is this beyond the realm of what is affordable, but of what is currently available.

Analyzing the qualitative results of the counts for the best technology combinations also yielded interesting results for performance and real-life performance methods. Live artifacts, demo hardware, and AR were selected by three SME's as part of a best combination of technologies for performance. However, live artifacts were selected

by four SME's while only one SME selected AR for the best technology to use with real-life performance. This indicates that use of virtuality is not preferred in real-life scenarios by the group of respondents. Additional information on training technologies was gathered by asking the participants to rate the utility of each technology based upon select attributes. AR and VR technologies had a mean normalized utility value 15 counts apart or closer for each attribute category, indicating similar advantages and disadvantages between them. When compared to live artifacts, AR and VR had the advantage of user safety, but had the disadvantage of tactile feedback. In this case, tactile feedback appears to be the primary driver of SME preference for live artifacts during real-life performance. More data needs to be gathered to assist the development of training alternatives.

The example problem at the end of Chapter 4 illustrates how MBSE software can be used to capture problems, objectives, and attributes of training in various diagrams. These diagrams help to visually express information and to trace from one step in the VFT framework to the next, in some instances. The example highlights the differences between the two training system alternatives based on the evaluation measures, the associated utility functions, and attribute weights. Utility values for each training system can reveal differences between alternatives in a single attribute category. The training technology utility values provided by SMEs for several attributes can be plotted pairwise to give a visual indication of performance gains. Initial attribute equations are tailorable to conform to a variety of problem scenarios. The total value function provides a single number that simply conveys the cost-effectiveness of a training system.

Significance of Research

These results are significant because they give maintenance organizations a starting point for evaluating and implementing more effective training programs using the VFT process. By implementing more effective training programs, maintenance organizations can support the warfighter by providing more effective technicians with the required skills. This research effort expanded upon the MPEET model in many ways. First, the results produced a set of tools that is compatible with an established systems engineering framework, the VFT process. Second, because the VFT process starts with problem identification, it directly addresses the user-defined problems and gives traceability to clear objectives related to training, as opposed to focusing solely on training program development. Third, this research produced a tool that allows for a more thorough classification of training effectiveness by incorporating important attributes of training selected through SME input. Fourth, this effort informed a deterministic way of evaluating LOs that places proper emphasis on objective competency levels when evaluating utility value. Fifth, this research evaluated AR and VR as distinct technologies from simulators, quantifying their utility and comparing them with traditional training aids. Finally, by integrating MBSE software to document and perform training alternative evaluations, this research enables digital engineering, providing a centralized location for representing, viewing, and managing training system complexity in real-time.

Steps 1 and 2 of the VFT process are satisfied by the objectives hierarchy created in SysML, where problems can be traced to general solutions. Steps 3 and 7 of the VFT process can be informed by using the total value function BDD to see important attributes, derive evaluation measures, and construct an initial cost-utility equation. To

satisfy step 4 of the VFT process, SME input obtained through interviews can be used to create value functions for LOs in the cognitive, affective, and psychomotor domains according to attainment of the module goal. Step 6 may be informed using the qualitative results of the best combinations of training method and technology. By using these tools, training organizations can evaluate the cost-utility of training solutions which utilize AR, VR, or traditional training aids, thereby selecting the best platform to deliver more effective training, satisfying organizational needs.

Limitations and Future Research

The primary limitations of this research included small sample sizes and lack of specific examples during interviews. Only 5 total participants were interviewed from 2 organizations within the Air Force. More samples from other organizations could add robustness to the results. Another reason for the small sample size was the lack of trainers who fit the profile of the ideal interview candidate who were also willing to participate in a 2-hour interview. During the interviews, most respondents noted that the interview questions could be clarified with specific examples. Because of the breadth of subject matter investigated in the interviews, it was not always possible to provide specific examples.

Future research should focus on repeating this process with a different population, performing a case study of a real example, or improving the user-friendliness of the tool. Specific examples of real training scenarios will be helpful for obtaining reliable results. Additionally, expansion of this model is necessary to provide users with the means to perform the sensitivity analysis portion of the VFT framework. By providing more in-

depth examples of the training analysis process, it will promote digital engineering as a common practice for training evaluation. Utility function equations for some attributes are incomplete or lacking sufficient data for representative formulation. Once reliable equations can be determined for all attributes, these functions can serve as constraints to find an optimal solution. Finally, this tool can be expanded by generating Pareto frontiers for pairwise attributes, which will allow the user to visually determine technologies that should be incorporated into a training alternative based on predicted utility as a function of cost.

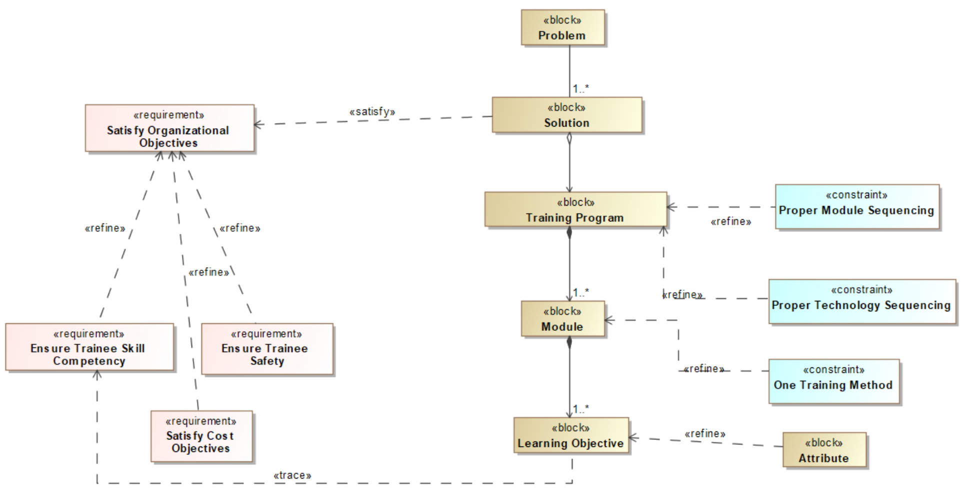
Summary

High rates of attrition in the field of maintenance, particularly 5- and 7-level maintainers, threaten the readiness of the warfighter. To combat this issue, training organizations require a means to implement more effective training solutions. By using the concepts and tools provided by this research, training organizations may analyze the cost-utility of training alternatives using the VFT process and document efforts using MBSE software. Findings of this research effort included the discovery of attributes that contribute to the total value of training, the determination that utility values of the objective and high competency levels are similar, and initial observations on the best training technologies to use with various training methods.

Appendix A: Interview Handout

Interview Handout

Interview Questions:



Module Goal	Training Alternative #1	Training Alternative #2
Assemble outdoor aircraft testing structure	**Training Method:** Lecture **Technologies Used:** -Computer-Based Courseware -Demo Hardware -Workbook	**Training Method:** Small Group Method **Technologies Used:** -Live Artifacts -Technical Information -Workbook

Total Value Function = **Cost**(0.30) + **Schedule**(0.20) + **Training Effectiveness**(0.50)

Attribute	Attribute Weight of Importance	Training Alternative #1 Score	Training Alternative #1 Value	Training Alternative #2 Score	Training Alternative #2 Value
Cost	0.30	0.67	0.20	0.50	0.15
Schedule	0.20	0.50	0.10	1.00	0.20
Training Effectiveness	0.50	0.68	0.34	0.90	0.45
Total Value	**1.00**		**0.64**		**0.80**

-This function is made of three attributes (cost, schedule, training effectiveness)
-The total value function output is on a scale from 0-1, with 0 as the least desirable and 1 as the most desirable
-One goal of this research is to get your opinion on what are the most important pieces of training effectiveness
-Another goal is to find which attributes you consider most important

Problem

1. What problems areas have you encountered which led you to consider revising training programs or implementing new training methods?

Objectives

2. What were the main objectives of revising some aspect of your training program?

Evaluation Measures/Value Function

3. If you were developing a new module and had to decide how to deliver this module for example, **a)** in a classroom versus **b)** hands on with a dedicated piece of training equipment, what would you consider when making this decision?
4. Which attributes would have the most impact on your decision?

5. If you were developing a new training module and you had to decide whether to **a)** deliver this module using a piece of dedicated training equipment or **b)** a virtual reality system, what would you consider while making this decision?
6. Which attributes would have the most impact on your decision?

7. When you made decisions like this in the past, did you use any tool to aid your decision? If so, what was this tool?

8. I have compiled a list of attributes which I think may help to build a total value function. The goal of the total value function is to decide between training alternatives on a module-by-module basis. Only one training method can be used per module. Please review this list, reading each definition. Are there any attributes you can think of that would complete this list?
9. Which attribute do you consider to be the single most important and which do you consider to be the single least important?
10. Rate the following attributes on a scale from 0 to 10 with 0 as the least important and 10 as the most important. Note that multiple attributes may have the same rating. Use integers only.

Attribute	Definition
Cost	Composed of both direct and indirect costs. Subcategories include Personnel, External Training Services, Training Development & Material Prep, Instructional Materials, Equipment, Facilities, Travel, Overhead, and General & Administrative
Immersion	The feeling or perception of one being in a real environment. Immersion scale includes the feeling of being in a real, virtual environment, or irrelevant environment.
User Safety	The health risk level posed to the trainee by the training method. (Low, medium, or high risk where lowest risk corresponds to higher safety and vice versa)
Schedule	Timeline of implementation. Can be measured in days, weeks, months, or years depending upon the alternative systems provided.
Cognitive Competency	The recommended level of cognition for desired trainee competency. This category includes learning objectives related to mental skill (for example, trainee knowledge, application, or evaluation skills)
Psychomotor Competency	The recommended level of psychomotor ability for desired trainee competency. This category includes learning objectives related to physical skills.
Interactive Competency	The recommended interactive skills for desired trainee competency. This category includes learning objectives related to interpersonal interactive skills (for example, the ability to perform the work as a member of a team).
Affective Competency	The recommended level of affectivity for desired trainee competency. This category includes learning objectives related to motivation, attitudes, and values.
Trainer Access	The level of guidance provided by a trainer in real time. This scale is assumed to include no trainer, virtual guide, virtual trainer, and physical trainer.
Information Access	The ability to access various types of data such as diagrams, symbols, information, audiovisuals, etc.
Repetitions	The number of times an activity can be repeated by the same trainee in the same module.
Tactile Feedback	The level of tactile feedback provided to the trainee. For example, hands off, some hands-on, complete hands-on
Service Life	The estimated time until obsolescence of training system alternative. It is assumed lifecycle costs are included in the cost portion of the value function. Typically measured in months or years.
Distance Accessibility	Whether training materials can be administered from a distance or are required to be used on-site or in the classroom.
Availability	The level of certainty that training materials, hardware, or artifacts will be available for use when needed. Could also be classified as reliability or uptime of the system. (ex. Is it a shared system? Does it have connectivity issues? Is it down for Mx?)
Other #1	
Other #2	
Other #3	
Other #4	

Cost

11. Rank each cost type according to order of importance with the most important attribute having a rank of 1, the second most a rank of 2, etc.
12. Are there any cost types that you do not think are important to include (for example, all attributes with a rank of 6 or greater)?

Costs		Examples		Rank
Direct	**Personnel**	Wages and benefits of trainers and trainees		
	External Training Services	Externally produced training materials, delivery costs		
	Training Development and Instructional Material Prep	Costs of training development effort, costs of course content preparation, supplies related to prep		
	Instructional Materials	Instructional materials, books, pencils, paper		
	Equipment	Rented or purchased hardware, hardware maintenance		
	Facilities	Rental of training facilities, facility maintenance		
	Travel	Air fare, housing, per diem		
Indirect	**Overhead**	Materials and labor not directly related to training		
	General & Administrative	Everyday organizational expenses, auditing expenses, legal expenses, internet		

13. You are a decision-maker for a training organization, and you are assigning value to a set of alternatives. Each alternative indicates the total cost to implement a new training program. (Note: The alternatives are for a brand-new training need. They are not replacing an existing training program.) The costs to implement the five different training alternatives are listed below in order of smallest to largest. On a scale from 0-1 (where 0 is the least desirable value, 1 is the most desirable value), what score would you assign each alternative? Round to the nearest .00 (ex. 0.467 = 0.47).

Alternative	Cost ($)	Value (0.00-1.00)
Training Program #1	100,000	
Training Program #2	200,000	
Training Program #3	300,000	
Training Program #4	460,000	
Training Program #5	500,000	

14. Draw the corresponding plot of Value versus Alternative.

This next section contains a video introducing the topic of classifying knowledge, attitude, and behavior of a trainee using specific words. 3 Domains video: https://www.youtube.com/watch?v=N-ZEcFaqcoE

Cognitive, Psychomotor, and Affective Domains

15. Read through the cognitive, affective, and psychomotor domains. These domains are organized from low to high competency levels, having the highest level on top. Then, plot the value function of each cognitive, affective, and psychomotor domain for each corresponding module.

Cognitive Domain

Level	Definition	Common Verbs
Create	Put elements together to form a coherent whole; reorganize into a new pattern or structure.	Generating, Hypothesizing, Planning, Designing, Producing, Constructing
Evaluate	Make judgments based on criteria and standards.	Checking, Coordinating, Detecting, Monitoring, Critiquing, Testing, Judging
Analyze	Break down information into component parts.	Differentiating, Organizing, Distinguishing, Selecting, Structuring Attributing, Parsing
Apply	Carry out or use a procedure in a given situation.	Executing, Carrying Out, Implementing, Using
Understand	Construct meaning from instructional messages, including oral, written, and graphic communication	Interpreting, Exemplifying, Classifying, Summarizing, Inferring, Comparing, Explaining
Remember	Retrieve relevant knowledge from long-term memory	Recognizing, Identifying, Recalling, Retrieving

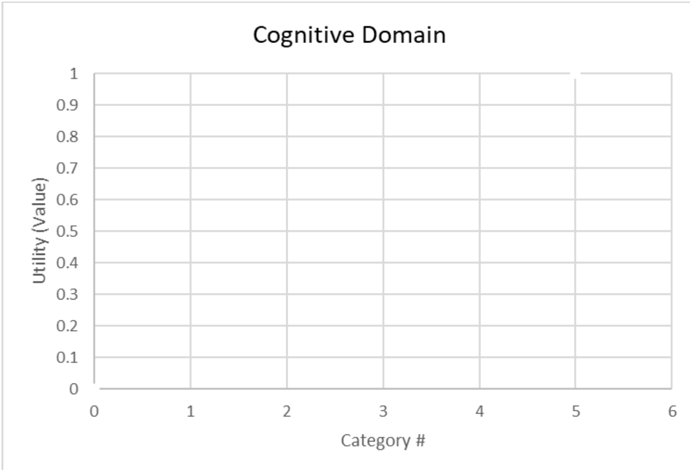
Affective Domain

Level	Definition	Common Verbs
Characterizing	To act consistent with the internalized values	Discriminate, Question, Revise, Change
Organizing	To begin to harmonize internalized values	Arrange, Combine, Compare, Balance, Theorize
Valuing	Being willing to be seen as valuing certain ideas or material	Justify, Propose, Debate, Relinquish, Defend, Initiate
Responding	Committing to the ideas, etc. by responding to them	Answer, Recite, Perform, Report, Select, Follow, Explore, Display
Receiving	Developing awareness of ideas and phenomena	Ask, Follow, Reply, Accept, Prefer

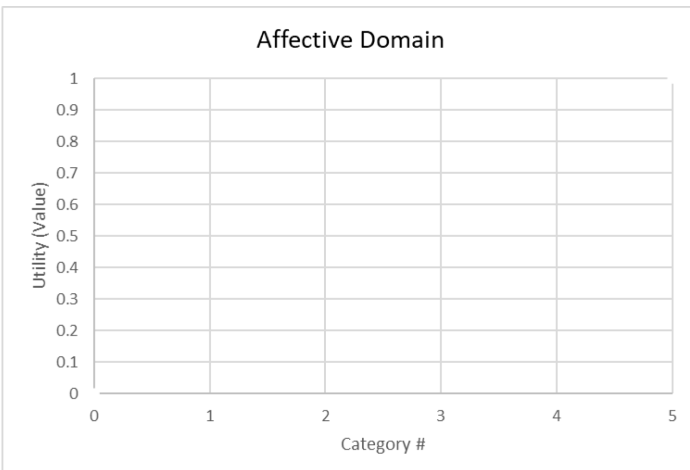
Psychomotor Domain

Level	Definition
Origination	The ability to develop new actions of behavior patterns by adapting already highly developed skills.
Adaptation	The ability to modify actions to meet different or unusual requirements.
Complex Responding/ Mastery	Skillful performance that is fully integrated and automatic. Implies a high level of accuracy and proficiency.
Basic Proficiency	The ability to perform with some confidence and proficiency but without mastery.
Guided Responses	Early stages in skill acquisition involving imitation and trial and error.
Set	Readiness to act. Becoming ready to respond to different situations.
Perception (Awareness)	The ability to use sensory cues to guide motor ability, from sensory stimulation to action.

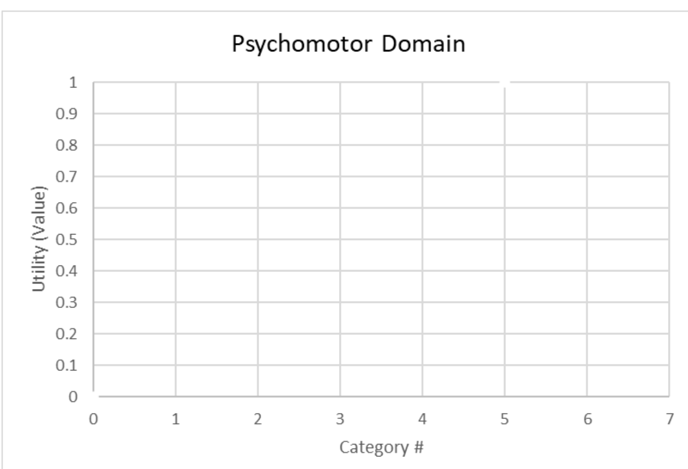
Module	Module Goal	Objective Level of Competency	
1	Identify that a torque wrench is necessary to attach a panel.	Cognitive	Remember
		Affective	Responding
		Psychomotor	Perception (Awareness)
2	Be able to identify and use the proper torque wrench successfully given a verbal prompt by the instructor.	Cognitive	Apply
		Affective	Responding
		Psychomotor	Guided Responses
3	Recognizing improper function of a torque wrench during use and modifying it back to proper functionality through troubleshooting/repair.	Cognitive	Evaluate
		Affective	Characterizing
		Psychomotor	Adaptation



Category Name	Category #
Create	6
Evaluate	5
Analyze	4
Apply	3
Understand	2
Remember	1
None	0



Category Name	Category #
Characterizing	5
Organizing	4
Valuing	3
Responding	2
Receiving	1
None	0



Category Name	Category #
Origination	7
Adaptation	6
Complex Responding/	5
Basic Proficiency	4
Guided Responses	3
Set	2
Perception (Awareness)	1
None	0

16. For each training module, was the highest level of competency always the most desirable? Why or why not?

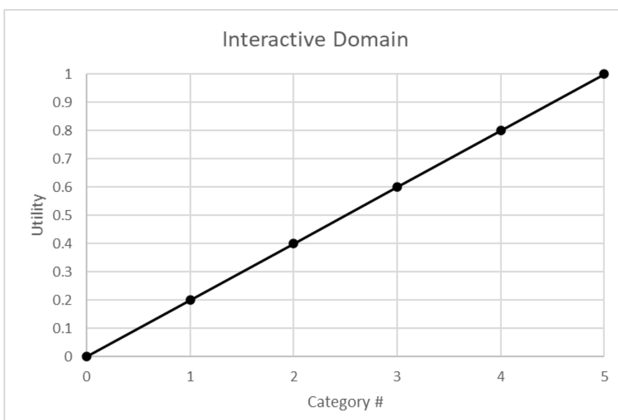
Interactive Domain

17. Read through the Interactive Domain Levels. Do you believe that the Interactive domain levels should be put in a specific order of increasing skill like the cognitive domain, or should each level be evaluated as needed? Support your answer.

Interactive Domain

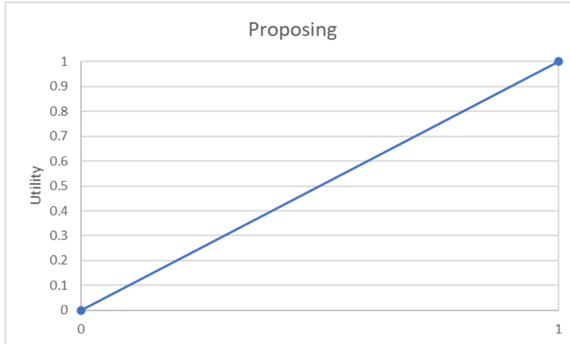
Level	Definition
Seeking/Giving Information	Seeking or offering clarification of facts or opinions to / from individuals.
Proposing	Putting forward a new concept, suggestion or course of action that can be actioned.
Supporting	Conscious and direct declaration of support or agreement with another person or his concepts.
Including	Direct and positive attempt to involve another group member.
Summarizing	Summarizing or restating in a compact form the content of previous discussions or considerations.
Disagreeing	Conscious, direct and reasoned declaration of difference of opinion, or criticism of another person's concepts.

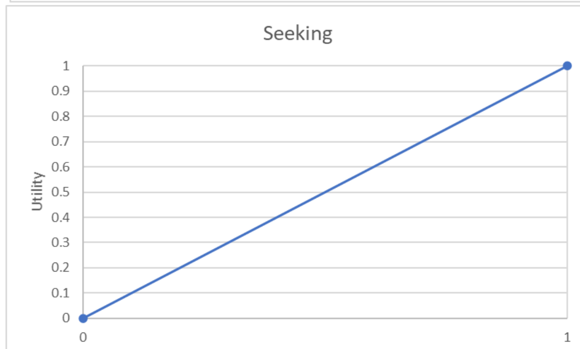
a) Increasing Order of Skill:

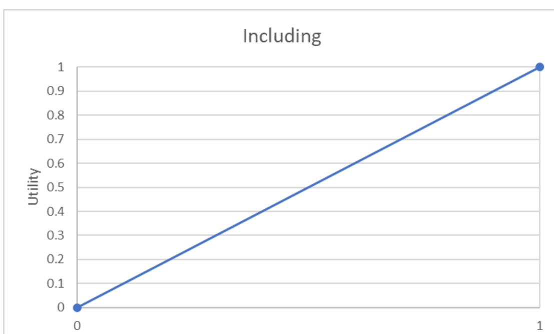


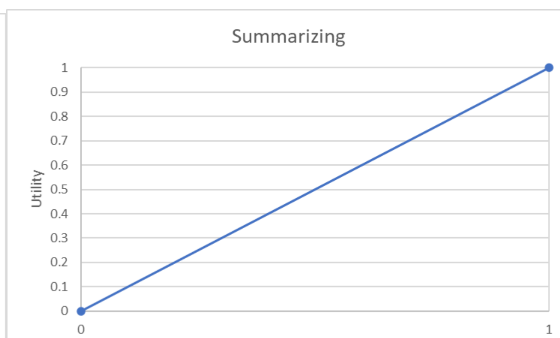
Category Name	Category #	Proficiency (%)
None	0	0
Seeking	1	20
Proposing	2	40
Supporting	3	60
Including	4	80
Summarizing	5	100

b) Evaluated As Needed (0 = No Competency, 1 = Full Competency):









Training Method & Technology

18. Read the definitions of the following training methods and training technologies. Are there any that you see missing from this list?

Training Method	Definition
Lecture	An oral presentation of information by a single individual; facts, concepts, problems, relationships, rules or principles presented orally either directly (as by classroom instructor) or indirectly (as by video).
Demonstration	Presentation or portrayal of a sequence of events to show a procedure, technique, or operation; frequently combines an oral explanation with the operation or handling of systems equipment or material. May be presented directly (as by a classroom instructor) or indirectly (as by video).
Exhibit	A visual display used to present information; for example, actual equipment, models, mockups, graphic materials, displays, chalkboard, or projected images in which the procedure is demonstrated but the trainee does not interact with the system
Questioning	An instructor and/or courseware controlled interactive process used to emphasize a point, stimulate thinking, keep students alert, check understanding, or review material. Questioning may be direct, as by a classroom instructor, or may be designed into a film or television presentation.
Seminar	A peer-controlled group interactive process in which task- or objective related information and experience are evoked from the students. Questions may be used to evoke student contributions, but the seminar is distinguished from questioning.
Discussion	An instructor-controlled interactive process of sharing information and experiences related to achieving a training objective.
Performance	A student interaction with things, data, or persons, as is necessary to attain training objectives; includes all forms of simulation (for example, games and interaction with hardware simulators) and interaction with actual equipment or job materials (for example, forms). Performance may be supervised by classroom instructor, tutor, coach, or peer to provide needed feedback.
Case Study	A carefully designed description of a problem situation, written specifically to provoke systematic analysis and discussion.
Real life performance	The trainee performs the task in the real environment without feedback
Indirect Discourse	Verbal interaction among two or more individuals which is heard by the student; may be a dramatization, such as role playing, or a dialogue between panel members, or a teaching interview (a question-and-answer session between instructor and visiting expert).
Assigned Reading	Printed materials such as books, periodicals, manuals, or handouts. Readings may be course-assigned or self-assigned.
Small Group Method	A means of delivering instruction which places the responsibility for learning on the student through participation in small groups led by a leader who serves as a role model throughout the activity. The small group method uses small group processes, methods, and techniques to stimulate learning. The leader is an instructor who facilitates role modeling, counseling, coaching, learning, and team building in the small group. Under the small group method, brainstorming, buzz session, role playing, and committee problem-solving techniques may be applied.
Student Query	The provision by which students are given the opportunity to search for information, as by questioning a classroom instructor, tutor, coach, or an appropriately programmed computer.

Training Technologies	Description
Chalkboard/ Whiteboard	A panel on which words, illustrations, plots, and symbols are drawn by an instructor.
Oral Description	Communication tool used for presenting in the form of statements, descriptions, questions, answers, and corrections
Augmented Reality	A student interaction with things, data, or persons, with the ability to access illustrations specific to the task at hand in real time as is necessary to attain training objectives. Includes hardware simulators and interaction with systems which provide step by step instructions to guide this interaction, usually provided through a tablet or head mounted display. Performance may be supervised by classroom instructor, tutor, coach, artificial coach, or peer to provide needed feedback.
Virtual Reality	A student interaction with things, data, or persons, as is necessary to attain training objectives within a virtual environment. Performance may be supervised by an artificial coach within the environment and final critiques of performance may be provided by an instructor. It is typically administered by head mounted display.
PowerPoint Slides	An instructor-controlled delivery of projected words, sounds, images, videos, graphs, and diagrams
Computer-based lessons	An organized set of information administered by computer that is self-paced, containing words, sounds, images, videos, graphs, and diagrams
Interactive computer courseware	A courseware-controlled training delivery method requiring specific software design to emphasize training objectives based upon trainee activity and interaction with words, sounds, images, videos, graphs, test questions, and diagrams.
Simulator	Hardware that imitates the operation of a real-world behavior, process, scenario, or system which is used to attain training objectives.
Models	Scaled replicas or components of physical artifacts that do not possess full system functionality
Live artifacts	System that provides full functionality of use equivalent to systems in a real-life scenario
Demo Hardware	Components or replicas of physical artifacts that do not possess full functionality of a live artifact
Technical Information	Documents which provide technical information such as diagrams, procedures, and manuals regarding the performance of a task or use of tools or hardware
Workbooks/ Handouts	Physical documents which accompany a training course that provide information, case studies, forms, problems, and questionnaires

19. Indicate whether the training technology is appropriate for the training method by entering "1" for yes. Then, indicate the best combination of technologies for each training method in the right column (ex. AEM).

	Training Technology													
	A	B	C	D	E	F	G	H	I	J	K	L	M	
Training Method	Chalk board / White board	Oral Desc	AR	VR	PPT Slides	Computer-based lessons	Interactive computer courseware	Simulator	Models	Live artifacts	Demo Hardware	Tech Info	Workbooks/ Handouts	Best Combination
Lecture														
Demonstration														
Exhibit														
Questioning														
Seminar														
Discussion														
Performance														
Case Study														
Real Life Performance														
Indirect Discourse														
Assigned Reading														
Small Group Method														
Student Query														

20. Provide the best estimate of the level of utility (from 0-100%) provided by each training technology for the given attributes.

		Attributes								
		Immersion	User Safety	Repetitions	Tactile Feedback	Service Life	Distance Accessibility	Trainer Access	Information Access	Availability
Training Technology	Chalkboard/ Whiteboard									
	Oral Description									
	Augmented Reality									
	Virtual Reality									
	PowerPoint Slides									
	Computer-based lessons									
	Interactive computer courseware									
	Simulator									
	Models									
	Live artifacts									
	Demo Hardware									
	Technical Information									
	Workbooks/ Handouts									

Bibliography

Bartley, S. J., & Golek, J. H. (2004). Evaluating the cost effectiveness of online and face-to-face instruction. *Journal of Educational Technology & Society, 7(4)*, 167-175.

Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction, 21(6)*, 715-730.

Brown, C. C. (2014). A methodology for the evaluation of training effectiveness during early phase defense acquisition. (Doctoral dissertation, Georgia Institute of Technology).

Buede, D. M., & Miller, W. D. (2016). *The Engineering Design of Systems: Models and Methods.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Cameron, B., Crawley, E., & Selva, D. (2016). Systems architecture. Strategy and product development for complex systems. Pearson Education.

Clark, D. R. (1999, June 5). Retrieved December 10, 2021, from Learning strategies or instructional strategies: http://www.nwlink.com/~donclark/hrd/strategy.html

DAF. (1998, October). *CFETP 2A3X3B.* Retrieved from Career One Stop: https://www.careeronestop.org/competencymodel//modelFiles/F16%20Aircraft%20Main%20Spec.pdf

DAF. (2019, April 01). Air Force Instruction (AFI) 90-802. *Risk Management*. Air Force Safety Center. Retrieved from http://static.e-publishing.af.mil/production/1/af_se/publication/afi90-802/afi90-802.pdf

Day, N. (2020). *Cost of virtual reality training: Full VR*. Retrieved from Roundtable Learning: https://roundtablelearning.com/cost-of-virtual-reality-training-full-vr-2020/

Decker, C. A., & Campbell, C. P. (1996, May 20). Determining the cost-effectiveness of training (A self-contained instructional model). The University of Tennessee.

Fade, L. (2019, August 27). *VR & AR Application Development Costs Explained*. Retrieved February 2022, from VR Vision: https://vrvisiongroup.com/vr-ar-application-development-costs-explained/

GAO. (2019). *Military Personnel: Strategy Needed to Improve Retention of Experienced Air Force Aircraft Maintainers.* Washington, DC: United States Government Accountability Office.

Garzon, J., Pavon, J., & Baldiris, S. (2019). Systematic review and meta-analysis of augmented reality in educational settings. *Virtual Reality, 23(4)*, 447-459.

Gavish, N., Gutierrez, T., Webel, S., Rodriguez, J., Peveri, M., Bockholt, U., & Tecchia, F. (2015). Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. *Interactive Learning Environments, 23(6)*, 778-798.

Goldstein, P. (2020, April 13). *Air Force Turns to VR, AR for Training and Maintenance*. Retrieved from FedTech: https://fedtechmagazine.com/article/2020/04/air-force-turns-vr-ar-training-and-maintenance

Grossman, R., Oglesby, J., & Salas, E. (2015). The training process: Using the science each step of the way.

Hermans, L. M., & Cunningham, S. W. (2018). Actor and strategy models: Practical applications and step-wise approaches. 84. John Wiley & Sons.

International Civil Aviation Organization. (2016, December). Taxonomy to Assist in the Identification of Instructional Methods (E-learning, Classroom, and Blended Training). Retrieved October 7, 2021, from https://www.icao.int/training/Documents/GAT%20Training%20Taxonomy%202016.pdf

Iowa State University. (2012). A Model of Learning Objectives based on A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Iowa State University Center for Excellence in Learning and Teaching. Retrieved October 2021, from https://www.celt.iastate.edu/wp-content/uploads/2015/09/RevisedBloomsHandout-1.pdf

Kaplan, A. D., Cruit, J., Endsley, M., Beers, S. M., Sawyer, B. D., & Hancock, P. A. (2021). The Effects of Virtual Reality, Augmented Reality, and Mixed Reality as Training Enhancement Methods: A Meta-Analysis. *Human Factors, 63(4)*, 706-726.

Keeney, R. L. (2009). Value-focused thinking: A path to creative decisionmaking. Harvard University Press.

Keesling, R. B. (2019, March 21). Exploratory Analysis of the Potential Use of Augmented Reality in Aircraft Maintenance. Wright-Patterson Air Force Base, OH, USA: Air Force Institute of Technology.

Kraiger, K. (2008). Third-Generation Instructional Models: More About Guiding Development and Design Than Selecting Training Methods. *Industrial and Organizational Psychology, 1(4)*, 501-507.

Lackey, S. J., Salcedo, J. N., Szalma, J. L., & Hancock, P. A. (2016). The stress and workload of virtual reality training: the effects of presence, immersion and flow. *Ergonomics, 59(8)*, 1060-1072.

Lyndall III, F. (2020, July 22). *The effectiveness of virtual simulation as a training tool*. Retrieved October 26, 2021, from Army University Press: https://www.armyupress.army.mil/Journals/NCO-Journal/Archives/2020/July/The-Effectiveness-of-Virtual-Simulation-as-a-Training-Tool/

Manacapilli, T., Bailey, A., Beighley, C., Bower, A., & Bennett, B. (2007). Finding the balance between schoolhouse and on-the-job training. *(Vol. 555)*. Rand Corporation.

McCarty, J. A., & Shrum, L. J. (2000). The measurement of personal values in survey research: A test of alternative rating procedures. *Public Opinion Quarterly, 64(3)*, 271-298.

Merriam, S. B. (2002). Introduction to qualitative research. *Qualitative research in practice: Examples for discussion and analysis, 1(1)*, 1-17.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems, 77(12)*, 1321-1329.

Mulders, M., Buchner, J., & Kerres, M. (2020). A framework for the use of immersive virtual reality in learning environments. *International Journal of Emerging Technologies in Learning (iJET), 15(24)*, 208-224.

Murty, K. G., Djang, P., Butler, W., & Laferriere, R. (1995). The army training mix model. *Journal of the Operational Research Society, 46(3)*, 294-303.

National Commission on Military Aviation Safety. (2020). 224 Lives, $11.6 Billion, 186 Aircraft: Military Aviation Losses FY2013-2020. United States.

O'Neill, G., & Murphy, F. (2010, January 20). Guide to Taxonomies of Learning. University College Dublin. Retrieved October 2021, from https://www.ucd.ie/t4cms/ucdtla0034.pdf

Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning, 37(1)*, 226-241.

Patten, M. L. (2009). Understanding research methods: An overview of the essentials. Pyrczak.

Rupp, D. E., Gibbons, A. M., & Snyder, L. A. (2008). The role of technology in enabling third-generation training and development. *1(4)*, 496-500. Industrial and Organizational Psychology.

Shoviak, M. J. (2001). Decision analysis methodology to evaluate integrated solid waste management alternatives for a remote Alaskan air station.

Simpson, E. (1971). Educational objectives in the psychomotor domain. *Behavioral objectives in curriculum development: Selected readings and bibliography, 60(2)*, 1-35.

Simpson, H. (1995). Cost-Effectiveness Analysis of Training in the Department of Defense. Monterey, CA: Defense Manpower Data Center.

Svensson, E. (2001). Guidelines to statistical evaluation of data from rating scales and questionnaires. *Journal of Rehabilitation Medicine, 33(1)*, 47-48.

United States Department of Defense. (2001, August 31). Department of Defense Handbook Instructional Systems Development/Systems Approach to Training and Education (Part 2 of 5 Parts). (2001, Ed.) Lakehurst, NJ: Naval Air Warfare Center Aircraft Division.

VRTL Academy. (2018). *Stereoscopic vision and virtual reality*. Retrieved from VRTL: https://courses.vrtl.academy/lessons/whats-up-with-stereoscopic-and-virtual-reality/

Webel, S., Bockholt, U., Engelke, T., Gavish, N., Olbrich, M., & Preusche, C. (2013). An augmented reality training platform for assembly and maintenance skills. *Robotics and autonomous systems, 61(4)*, 398-403.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
24-03-2022	Master's Thesis	March 2021 – March 2022

TITLE AND SUBTITLE	
Exploring the value of training system alternatives through methods, technologies, and relevant attributes	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Novitsky, Brian R.	**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/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-7765	AFIT-ENV-MS-22-M-243
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
711th Human Performance Wing 2620 Q Street, Wright Patterson, AFB, OH 45433 (602)418-9513, winston.bennett@us.af.mil ATTN: Dr. Winston "Wink" Bennett	711th HPW/RHW
	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
Skilled maintainers are being lost at a higher rate than average of all enlisted career fields within the Air Force. Although incentive programs can provide some retention, attrition rates may continue to vary, placing the readiness of the warfighter in jeopardy. Additionally, the technologies used to train maintainers have evolved substantially, including augmented reality (AR) and virtual reality (VR) solutions, among others. These technologies have recently gained notoriety as novel ways to increase training effectiveness, providing avenues for maintainers to achieve greater skill levels at a faster rate than traditional platforms. Through semi-structured subject matter expert (SME) interviews, this qualitative research investigated the relationships between training technologies, methods, and student learning objectives to build a tool for organizations deciding between training alternatives. The results provided important attributes of training system alternatives to use in a total value function for cost-utility estimation and information on common problems requiring a training-based solution. Key findings also included SME input about the best training technology combinations for each training method and data demonstrating that value is a function of learning objectives. Together, these findings give maintenance organizations an initial roadmap on how to decide between training system alternatives using the value-focused thinking (VFT) framework.

15. SUBJECT TERMS
Value-focused thinking, training alternatives, value attributes, objective competency level

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Michael E. Miller, AFIT/ENV
U	U	U	UU	159	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-6565, ext 4651 michael.miller@afit.edu

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