Effect of Topography on Learning Military Tactics – Integration of Generalized Intelligent Framework for Tutoring (GIFT) and Augmented REality Sandtable (ARES)

Technical Report · September 2016

DOI: 10.13140/RG.2.2.32090.75208

7 authors, including:

Michael Boyce
Yale University
59 PUBLICATIONS **627** CITATIONS
SEE PROFILE

Benjamin Scott Goldberg
Army Research Laboratory
74 PUBLICATIONS **910** CITATIONS
SEE PROFILE

Robert Sottilare
Soar Technology, Inc.
219 PUBLICATIONS **2,229** CITATIONS
SEE PROFILE

ARL-TR-7792 ● SEP 2016

US Army Research Laboratory

Effect of Topography on Learning Military Tactics – Integration of Generalized Intelligent Framework for Tutoring (GIFT) and Augmented REality Sandtable (ARES)

by Michael W Boyce, Ramsamooj J Reyes, Deeja E Cruz, Charles R Amburn, Benjamin Goldberg, Jason D Moss, and Robert A Sottilare

Approved for public release; distribution is unlimited.

NOTICES

Disclaimers

ARL-TR-7792 ● SEP 2016

US Army Research Laboratory

Effect of Topography on Learning Military Tactics – Integration of Generalized Intelligent Framework for Tutoring (GIFT) and Augmented REality Sandtable (ARES)

by Michael W Boyce
Oak Ridge Associated Universities, Oak Ridge, TN

LT COL Ramsamooj J Reyes
Air Force Institute of Technology, Wright-Patterson AFB, OH

Deeja E Cruz
Institute for Simulation and Training, Orlando, FL

Charles R Amburn, Benjamin Goldberg, Jason D Moss, and Robert A Sottilare
Human Research and Engineering Directorate, ARL

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE ***(DD-MM-YYYY)***	**2. REPORT TYPE**	**3. DATES COVERED (From - To)**
September 2016	Technical Report	February 2015–February 2016

4. TITLE AND SUBTITLE

Effect of Topography on Learning Military Tactics – Integration of Generalized Intelligent Framework for Tutoring (GIFT) and Augmented REality Sandtable (ARES)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Michael W Boyce, Ramsamooj J Reyes, Deeja E Cruz, Charles R Amburn, Benjamin Goldberg, Jason D Moss, and Robert A Sottilare

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

US Army Research Laboratory / Oak Ridge Associated Universities
ATTN: RDRL HRT T
Aberdeen Proving Ground, MD 21005

8. PERFORMING ORGANIZATION REPORT NUMBER

ARL-TR-7792

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

10. SPONSOR/MONITOR'S ACRONYM(S)

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

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report discusses the findings of a research project that integrated the Generalized Intelligent Framework for Tutoring and the Augmented REality Sandtable (ARES) for the assessment of military tactics. An experiment involving 19 Reserve Officer's Training Corps cadets from a large southeastern university assessed performance, physiological, and experiential data in a between-subjects design. The conditions consisted of a 2-D map displayed on either a flat or contoured surface, both of which leveraged the projection technology of ARES. Results of the study did not indicate significant differences between time on task, accuracy, or electrodermal activity, but a larger sample size is needed to verify findings. Preference between conditions was more prevalent in support of ARES; however, individuals that preferred the flat condition discussed issues with the ARES condition being more difficult to precisely measure, consistent with the findings in the literature. This study is being refined for a follow-on experiment to be conducted at the United States Military Academy at West Point.

15. SUBJECT TERMS

military tactics, displays, intelligent tutoring systems, military sand tables, ARES, GIFT

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Michael W Boyce
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER (Include area code)**
Unclassified	Unclassified	Unclassified	UU	78	407-384-3910

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

Contents

List of Figures

List of Tables

Acknowledgments

The authors would like to recognize the support of Mr Nathan Vey, the Augmented REality Sandtable project engineer, for all of his assistance in making this project possible. While he may not appear on the byline of this work, without him the project would have not succeeded. We would also like to thank our subject matter experts from the SFC Paul Ray Smith Simulation and Training Technology Center, the United States Military Academy at West Point, the Florida National Guard at Camp Blanding, and the faculty of the University of Central Florida Reserve Officer Training Corps program.

Research was sponsored by the US Army Research Laboratory (ARL) and was accomplished under Cooperative Agreement Number W911NF-12-2-0019. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of ARL or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

INTENTIONALLY LEFT BLANK.

1. Introduction

The US Army Research Laboratory (ARL) Adaptive Training program of research for Domain Modeling aims to expand the use of adaptive training systems and educational tools into militarily relevant learning domains of increasing complexity and where the domain parameters may be ill-defined (Sottilare et al. 2015). This report supports the Adaptive Training program of research through the development of an integrated system of the Generalized Intelligent Framework for Tutoring (GIFT) (Sottilare et al. 2012) and the Augmented REality Sandtable (ARES) (Amburn et al. 2015) to assess junior-level military tactics knowledge. This preliminary experiment aims to establish a baseline of comparing different display types for use in future experiments.

GIFT, an Intelligent Tutoring System (ITS), is an open-source, modular architecture to support developing intelligent tutors by reducing the cost to author content, increasing the automated delivery of content, and providing a user-friendly means to collect and manage learner data (Sottilare et al. 2015). GIFT supports research in several military-relevant domains including combat medic training (Baker et al. 2015), marksmanship (Goldberg et al. 2014), and military team training (Walton et al. 2015). However, the domain of military tactics has not yet been explored, providing an opportunity to contribute to the existing literature. Furthermore, GIFT has interoperability functionality, which allows it the ability to connect with external systems like ARES (Hoffman and Ragusa 2015).

ARES is a research project that combines a traditional military sand table with a commercial off-the-shelf (COTS) laptop, monitor, projector, and Microsoft Kinect sensor to create a projection of terrain on the sand and provides mission briefing and display capabilities through a drag-and-drop user interface (Amburn et al. 2015). Although potential interface designs supporting military tactics training using ARES exist in the literature (Boyce et al. 2015), this research helps to investigate the specifics on how the display capabilities of ARES can impact cadet performance. Figure 1 is a photo of a cadet interacting with the ARES/GIFT system, with ARES projecting the map onto the sand and GIFT presenting questions via the liquid-crystal display.

Fig. 1 Participant using the ARES/GIFT system

The combination of ARES and GIFT presents many different use cases to facilitate adaptive training research, both within and outside of the military. The domain of military tactics was chosen due to its complexity and relevance to the cadet population. Alternatively, the combination of these 2 technologies could be incorporated into geology classes to explain how water runoff changes the possible long-term damage to a shoreline. It could also be used to assist in teaching young children how to successfully write in cursive by drawing patterns into the sand. It could even assist a triathlete in understanding the course that is in front of him.

In this experiment, the ARES projection displays a 2-D tactical map onto either a flat surface or a contoured surface (Fig. 2) in which the sand has been manually shaped by subject matter experts. GIFT asks questions associated with each map (Appendix A), and any pre- and postsurveys.

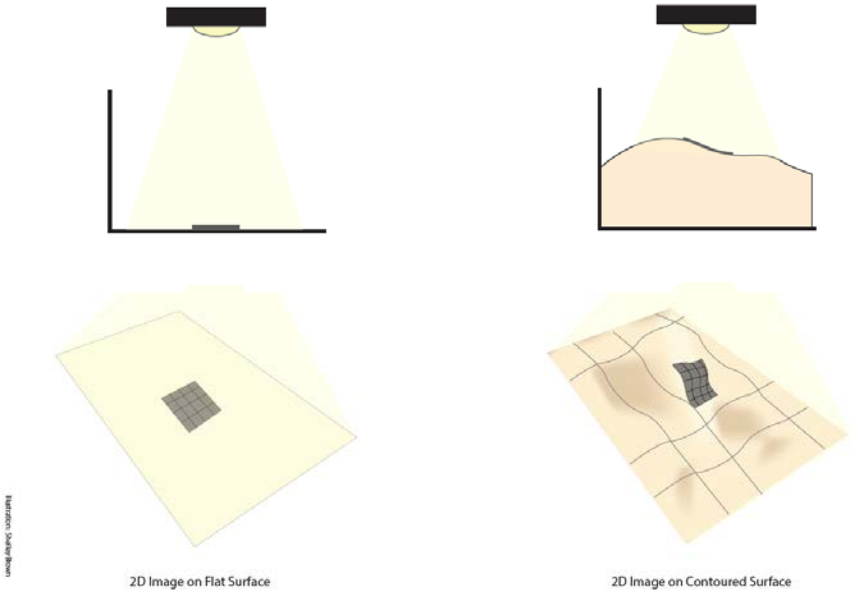


Fig. 2 Experimental conditions, flat and contour

Following is a listing of the hypotheses associated with this research study. They are developed through a desire to examine workload, time on task, accuracy, and engagement:

- The ARES condition will demonstrate an increase in cognitive workload for cadets in comparison to the flat display condition, as assessed by the global workload score of the National Aeronautics and Space Administration Task Load Index (NASA-TLX) (Hart and Staveland 1988), since ARES is a novel interface compared to the control condition.
- ARES will produce higher levels of accuracy on the assessment questions than the flat condition due to the additional information provided by the contoured display.
- Individuals in the ARES condition will take more time on task, due to increased information presented in the display, than individuals in the flat display. There is a limit on the amount of information that an individual can comprehend, and at a certain level of complexity it can become very difficult to grasp (Ware 2012).
- Use of ARES will show a greater overall increase in electrodermal activity than the flat condition due to an increase in arousal. It is expected that there will be a novelty effect supporting ARES that will lead to a higher level of arousal.

2. Definition of the Problem

The implementation of tactical decision exercises to support the development of the Soldier has been listed as "instrumental for BOLC [Basic Officer Leadership Course] cadre in developing leaders with critical thinking, problem solving, and decision making skills" according to the US Army Training and Doctrine Command's regulation 350-36 (2014). Each exercise has entities that provide the context for making a decision. With the increased use of technology to support these exercises, research has begun investigating how the use of different types of displays for tactics information, such as the Follow Me after action review tool (Jensen et al. 2014; Jensen et al. 2013) and the RAPTOR (Representation Aiding Portrayal of Tactical Operations Resources) interface, are supporting military command and control (Hall et al. 2012). However, there is a research gap in examining how displays on sand table type surfaces impact performance in small unit tactics training. This research aims to fill that gap by examining performance on tactics questions from 2 different types of displays.

The integration of GIFT (an engine to provide content) with ARES (a platform capable of displaying tactics content) can expand both of these respective systems and help improve the training of military tactics. The output of the system needs to be representative of how current classroom performance is measured. This research serves as the beginning to a continued effort to improve the quality of authoring and presenting content and placing it in an applied context that would be beneficial for military learners.

3. Background Research

In developing the hypotheses for this experiment, several areas of research were considered and are summarized here. The discussion begins with how the military makes considerations of terrain. Specifically it ties into the mission variables that we used when designing the experiment questions. There is also a discussion about the perception of depth cues as it relates to terrain. This is less focused on classifying displays and more focused on the cues that the brain uses to perceive an environment. This is relevant because it had influence on the metrics used in the experiment. Following this is the discussion on depth cues as applied to a learning domain supported by displays, specifically focused on augmented reality.

The next sections focus on 2 of the hypotheses: workload and physiological responses. In terms of workload, it discusses the specific impact of how a user's perception of information presented to them can impact performance. It terms of the physiological, this has to do with the responses to stimuli that are exhibited by a participant's body, even if the participant themselves is not aware. Together this literature served as the foundation for the experiment.

3.1 Military Considerations of Terrain

During tactical decision making, situations may be changing across several variables at once; therefore, training is often a balance of basic principles and understanding how those variables are represented (Jensen et al. 2014). The Army Field Manual for Offense and Defense (FM 3-90 2013) describes terrain in the context of tactical analysis variables: "The tactical situation is defined as the mission variables of mission, enemy, terrain and weather, troops and support available, time available, and civil considerations (METT-TC)." According to the Operations Process Manual (ADRP 5-0 2012), the mission variables provide an understanding of the potential impacts to the operation, which allows a commander to refine their battle plan according to the specific context. When analyzing terrain, commanders use a process known as OAKOC, which stands for observation and fields of fire, avenues of approach, key and decisive terrain, obstacles, cover and concealment (for details the reader is referred to ATTP 3-34.80, Geospatial Engineering).

- Observation and fields of fire refers to the area that a Soldier can perceive, with or without the assistance of technology. Fields of fire includes being able to understand appropriate ranges for weapons (both friendly and enemy) and determining if unit positioning provide adequate spacing for minimizing friendly fire (Department of the Army 2014; Oskamp 2015).
- Avenues of approach are ground routes that allow the squad or platoon to reach their objective safely (Department of the Army 2014; Oskamp 2015).
- Key terrain is an area that provides a tactical advantage to the force that controls it. Decisive terrain is terrain that through maintaining and controlling a particular area the force has the capability to control the outcome of a battle (Department of the Army 2014).
- Obstacles are natural or artificial features that are used to impede enemy movement to a position. These could include barriers, waterways, and buildings, etc. It is also used to direct an enemy toward a particular location (Oskamp 2015, Department of the Army 2014).
- Cover refers to the analysis of terrain that will protect the Soldier from bullets, explosions, and hazards. Concealment refers to the analysis of terrain that will hide the Soldier from enemy detection (Department of the Army 2014; Oskamp 2015).

These considerations helped develop the questions supporting the experiment. The questions were validated by multiple subject matter experts.

3.2 Depth Cues in Analyzing Terrain

This experiment is evaluating the impact of depth cues, as monocular depth cues can be used to easily distinguish differences in "relative size, occlusion, shading, spatial frequency of textures, motion parallax . . ." (McIntire et al. 2014). Experiments have shown that depth cues provide assistance in making relative judgments and comparisons between objects (St. John et al. 2001), use shading to increase the perception of terrain (Willet et al. 2015), and perspective, which can assist in representing either vertical or horizontal space for pilots (Lif 2012).

The curvature of the sand can also provide information, as the monocular cue of elevation can provide height (Barfield and Rosenberg 1995), which is a third point of reference beyond static positioning. Depth cues can be further divided into pictorial depth cues and non-pictorial depth cues (Van Beurden 2013), where pictorial depth cues are ones that can be captured via photographs or still images while nonpictorial require the use of motion. The cues that this experiment focuses on include the following:

- Relative size: being able to see different elevations and features and to make comparisons
- Occlusion: the use of terrain features to support cover and concealment or rationalization of decision making
- Perspective: According to Ware (2012), "the position of each feature…is determined by extending a ray from the viewpoint to that feature in the environment." Ware cites the work of Kubovy (1986) for the robustness of linear perspective. This means that the human brain is capable of not attending to the distortion in viewing location after looking at an image for a few minutes. This is the reason for using the mental rotation test, which will be discussed later. Therefore, perspective applied to the experiment refers to the aspects of individual terrain features which, based on how individuals view them, promote use or avoidance (i.e., the appearance of a steep hill that may be difficult to ascend).

3.3 Theoretical Foundations for Using Augmented Reality for Teaching and Learning

Augmented reality (AR) is defined as "overlaying virtual objects onto the real world" (Dünser et al. 2006, p. 1). It has also been defined as a type of virtual environment where objects are added to a real environment as opposed to virtual reality, which replaces the environment (Amburn et al. 2015; Azuma 1997; Milgram and Kishino 1994). AR has been shown to be an effective tool not only to treat psychological disorders such as phobias (Botella et al. 2010) but to enhance learning and performance in neurosurgical training (Luciano et al. 2011), mathematics (Bujak et al. 2013), spatial ability (Dünser et al. 2006), and many other areas of cognition. Thus, it would be appropriate to discuss how AR could be used to enhance learning, performance, retention, and transfer in academic settings.

According to Dunleavy and Dede (2014), "the assertion that AR could provide enhanced learning experiences is grounded in 2 interdependent theoretical frameworks: 1. Situated learning theory; and 2. constructivist learning theory" (p. 5). In terms of situated learning Brown et al. (1989) note that knowledge is "inextricably situated in the physical and social context of its acquisition and use" (p. 1). Teachers then use the embedding conditions of knowledge to help learners to ascribe meaning and construct understanding.

At its core, constructivism assumes that no single reality or truth exists, thus, all knowledge is subjective and personal (Schunk 1996). Constructivist theories of learning assume that: a) people are active learners that construct their own knowledge, and b) teachers should structure instruction around situations that allow learners to become involved with the content and other learners through social interaction. Moreover, according to Dunleavy and Dede (2014, p. 7), constructivist learning theory outlines 5 conditions most likely to enhance learning: 1) embed learning within relevant environments, 2) make social negotiation integral to the learning experience, 3) provide multiple perspectives and multiple modes of representation, 4) provide self-directed and active learning opportunities, and 5) support and facilitate metacognitive strategies within the experience (Bruner 1966; Cunningham 1992; Driscoll 2000; Piaget 1969; Vygotsky 1978).

One of the goals of the experiment is to construct a platform that can assist in fostering self-directed learning through different information visualizations and provide multiple different customizations for learners. Although the system does not currently extend to active learning, it is something that can serve as a target for future development and experimentation.

3.4 Display and Workload

Understanding how the display could potentially impose workload is essential. If appropriate interface elements exist to help the user understand what is going on with the system, also known as transparency, then workload can be reduced (Chen and Barnes 2014; Mercado et al. 2016). However, there needs to be an appropriate balance of information. Too much information could lead to an increase in workload due to an increase in information processing, which will make the system less usable (Bevan and Macleod 1994; Mercado et al. 2016). Too little information means that the users may not have the necessary information to use the system. Therefore, careful consideration for providing information while not overloading the user is important for optimal display design.

Past research has indicated that several design characteristics can assist in reducing workload (Boyce 2014). The device should have intuitiveness of use, which means that learning comprehension should be possible irrespective of the expertise of the user. This is not to say that it will be the same across users, rather that even novice users should be able to achieve functionality for basic procedures. This is accomplished by streamlining and simplifying the operation of the system so it maps with the mental model of the user. Example activities include eliminating unnecessary complexity, having consistency, highlighting important content, and effectively using feedback. Finally, the design needs to be able to handle users making errors. Errors are inevitable, so providing fail-safes, warnings, and clear descriptions of the consequences of actions will be essential.

3.5 Physiological Measurement Supporting Tactics and Learning

Physiological activity occurs from the responses of the sympathetic nervous system, which responds to external stimuli by increasing metabolic output (Poh et al. 2010). Electrodermal activity (EDA) is a method to measure arousal according to an increase in skin conductance (Boucsein 2012; Taylor et al. 2015). Arousal is relevant to this experiment because of its potential to produce a novelty effect. A novelty effect is when a participant exhibits a response based on the novelty of the situation, rather than on the intervention itself (Ary et al. 2013). Novelty effect can be accounted for through repeated presentations of the same stimuli, as the effect diminishes after initial interactions, as demonstrated in a study measuring affect in children interacting with robots (Leite et al. 2013).

EDA is influenced by the parts of the brain that deal with the management of emotion and as such is looked to for understanding changes in emotional/affective state, like mood indicators for students (Sano et al. 2015), at risk mothers (Rajan et al. 2012), and individuals with disabilities (Noordzij et al. 2012). The measurement of physiological data can aid in the understanding of arousal, stress, and other responses during tactical decision making (Hernandez et al. 2011). The standard process for measuring EDA involves the use of silver/silver chloride (Ag/AgCl) electrodes, which are typically placed on the second or third fingers. In this study the EDA is collected via a watch band-type sensor (Affectiva Q-Sensor) that can help in situations where participants need to move around freely (Boucsein 2012; Poh et al. 2010).

Research demonstrates that an increase in motivation or systems can trigger certain types of emotional states such as increased arousal and valence that can impact learning in a positive way (D'Mello and Kory 2015). It could also be that it is not just the type of display but an interaction between the display and the particular time during which a participant is interacting with the system that delineates a difference. Since the ARES/GIFT integration is a novel system using a nontraditional display, knowing whether this display leads to higher levels of interest and arousal can help justify integrating the classroom. The current research examines the effect of combining an ITS with contour and flat display surfaces while evaluating changes in EDA.

4. Methodology

4.1 Research Objective

The research objective for this work is to examine the effects of a 2-D map projected on a flat (flat condition) or contoured surface (ARES condition) integrated with an ITS on the ability to answer military tactics questions. This is assessed through participant answers to questions. Performance measures (e.g., accuracy and time on task), a spatial ability measure (i.e., mental rotation), a physiological measure (i.e., electrodermal activity), and video recording are used to provide results. The study was reviewed by the ARL Institutional Review Board and approved as protocol ARL-15-062 “Effect of Topography on Learning Military Tactics”.

4.2 Participant Recruitment

A total of 19 participants participated in the study. The participants were Reserve Officers’ Training Corps (ROTC) students at the University of Central Florida (UCF). The reason for the selection of these is to help support basic military tactics instruction exercises. The criteria for participation in the study are the following: 1) participants were older than 18 years of age, 2) have completed the Military Science 300 (MS-300) level course in ROTC, and 3) have 20/40 corrected- (self-reported) to-normal vision. Participants were recruited with the assistance of UCF ROTC faculty, with participants receiving 1.5 volunteer credit hours in exchange for their participation in the study.

4.3. Environment

The experiment took place at the UCF ROTC facility. In the facility, there is laboratory space (called the Battle Lab) where cadets typically participate in military simulation exercises. The Battle Lab has overhead halogen lights, and is access controlled so that only individuals with proper access codes can enter. The room has no windows and is air conditioned.

4.4 Procedure

Prior to arrival, participants are randomly assigned into 1 of 2 conditions, which is balanced for as close to equal numbers as possible. The 2 conditions, both using a 2-D map image, are either a contoured display (ARES) or a flat display. Upon arrival, participants receive a brief overview of the study and asked to fill out a paper informed consent form. As a part of the consent form, participants are asked if they also consent to video and audio recording. If the participant gives consent, a video camera is turned on and they are given a lavalier microphone to record audio. Next, GIFT administers a demographics survey, the Vandenberg and Kuse Mental Rotation Test, and the Self-Assessment Manikin Test. Participants are also fitted with the Affectiva Q-Sensor, which is an electrodermal activity sensor that is worn like a wrist watch.

Next, the participant reads a short introduction via GIFT explaining the scenario and the concepts to be covered in the lesson. The participant also takes a pretest to assess their knowledge of basic military tactical maneuvers. It takes approximately 45 min to perform all pre-experiment activities. They are then placed in the experimental scenario, which lasts approximately 20 min and consists of multiple prompts (Appendix A). Participants are asked to think aloud as they were answering the questions so the researchers can understand their reasoning processes.

Following the completion of the experimental scenario the participants are given a posttest, which is the same as the pretest. This is followed by an administration of the NASA-TLX, the Self-Assessment Manikin Test, and a system usability survey. The participant is debriefed on the study and free to leave. The total study is completed in less than 90 min per participant.

4.5 Equipment

4.5.1 Demographics Questionnaire

A demographics questionnaire is used to assess participant background and experience. It is attached in Appendix B for reference.

4.5.2 Mental Rotation Test

A 24-item mental rotation questionnaire is used to assess the spatial ability of the participant. Each item has a target figure followed by 2 reproductions of the target and 2 distractors. The participant has to select which 2 of the 4 figures are rotated representations of the target (Vandenberg and Kuse 1978). The questionnaire is attached in Appendix C for reference.

4.5.3 Generalized Intelligent Framework for Tutoring (GIFT)

GIFT is the ITS managing the experience during the study. It also allows for course content creation, pedagogical strategy implementation, and student assessment across a variety of domains (Sottilare et al. 2012). By having a better understanding of individual learners, through the collection of data, GIFT can provide in-depth knowledge about a learner’s state.

4.5.4 Affectiva Q-Sensor

The Affectiva Q-Sensor is a wearable, wireless biosensor that measures emotional arousal via skin conductance (Poh et al. 2010). It is integrated with GIFT for time-synced logging of EDA data with GIFT domain session information. The Q-Sensor collected data at a rate of 4 Hz/s.

4.5.5 Augmented REality Sandtable (ARES)

ARES is a proof-of-concept research test bed using a COTS projector, monitor, laptop, and Microsoft Kinect. The interested reader is referred to Amburn et al. (2015) for more detailed information on ARES.

4.5.6 NASA-Task Load Index (NASA-TLX)

The NASA-TLX uses a 6-dimension scale to assess subjective perception of workload. The following are the 6 dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. Participants were asked to rate each factor on a scale from low to high following completion of a task, which was followed by a series of pairwise comparisons to compare how individual dimensions were viewed in relation to one another (Hart and Staveland 1988). The scale can assist in accounting for variance in performance scores. For reference, the NASA-TLX is included in Appendix D.

4.5.7 Self-Assessment Manikin (SAM)

The Self-Assessment Manikin (SAM) is a picture-oriented scale to assess the affect dimensions of valence, arousal, and dominance. The SAM is composed of 3 sets of 5 figures (manikins), which stand for the 3 major affective dimensions. It is administered pre- and post-task (Bradley and Lang 1994). For reference, the SAM is included in Appendix E.

4.6 Study Design

Participants are assigned to 1 of 2 conditions using a 2-D map image, with the flat condition serving as the control while the contour (ARES) condition served as the experimental condition. The experiment is a between-subjects design with the display type serving as the main independent variable.

4.7 Data Analysis Techniques

4.7.1 Performance Data

For the performance data, there were several dependent variables, including time on task, tactics scores, NASA TLX scores, mental rotation scores, and the scores from the subjective surveys.

4.7.2 Physiological Data

For the physiological data, dependent variables included the number of skin conductance responses and the presence or absence of a skin conductance response. A visual inspection analysis process is used in determining the presence of a skin conductance response.

The analysis consists of the following steps:

1) After verifying that the data were represented at 4 Hz, and that analysis started after the first full second, the first derivative is taken to establish change in amplitude.
2) Any change in amplitude exceeding 3 μS is removed as an outlier.
3) The amplitude criterion was set at 0.05 μS, which was due to using visual inspection as a part of the analysis process.
4) The average (non-zero) standard deviation of the amplitude is taken, and this number is added to the threshold for the participant to accommodate for individual differences.
5) The subset consisting of EDA data from the start of the first question through the end of the eighth question is extracted and examined to determine locations where the data exceeded the criterion for at least 1 s.
6) The point at which the derivative crosses the x-axis is used to locate the peak of each potential skin conductance response (SCR).
7) Once the peak is located, the data points surrounding the peak are examined for exponential decay.

4.7.3 Qualitative Data

The qualitative data analysis used is an Interpretative Phenomenological Analysis (IPA). IPA is a process through which the participant experience is analyzed through detailed descriptions provided by the participant as to what caused them to choose the actions they did. Then those descriptions are categorized for comparison. The process for analyzing the data followed guidance described in Smith et al. (2009). The responses are from the cadets reasoning of the tactics questions listed in Appendix A, the same tactics questions that are discussed with the performance data. Before beginning the analysis, participant verbalizations need to be transcribed as close to verbatim as possible. Once transcribed, steps for the process are as follows:

1) Reading and rereading participant transcripts. This involves a close reading of the dialogue and verbalization of each participant. The purpose of this step is to ensure that the participant is the focus of the analysis.

2) Initial coding of each transcript. The aim of this step is to produce a set of comprehensive, detailed notes on the data. Emphasis is placed on things that matter to the participants and what the meanings of those things are in relation to what the participant is trying to accomplish.

3) The development of emergent themes from the data. An emergent theme is a concise statement of something that is important and attached to various comments and notes within a transcript. The goal of emergent themes is to reflect an understanding at a higher level, while at the same time, remaining true to the verbiage of the participant.

4) Developing connections across emergent themes. This is a process of clustering emergent themes so that those with similar meanings are placed together.

5) The process is repeated across each participant to determine if new themes arose between participants' data.

6) Once all of the transcripts have been analyzed, the final step is to extract superordinate themes that represent an aggregation of the emergent themes. These themes are those which represent patterns across the entire participant pool and can be identified using several different examples from both within and between the transcripts of participants.

5. Results

5.1 Performance Data Results

Table 1 shows means and standard deviations associated with the 2 conditions. It is intended to provide a high-level overview of the variables of interest. More detailed analysis will be provided.

Table 1 Summary of means and standard deviations

Condition	Flat		ARES	
	Mean	Std. deviation	Mean	Std. deviation
Age	21.33	0.50	22.30	3.02
Number of hours of sleep	5.56	1.67	6.20	1.32
Energy level	74.00	20.17	73.40	21.46
Confidence computer use	78.56	18.24	75.80	22.18
Knowledge military tactics	73.11	21.66	69.40	14.52
Pretest score	0.67	0.09	0.74	0.11
Tactics score	5.11	1.90	5.40	1.51
Tactics total time	355.07	132.01	308.84	116.08
Posttest score	0.79	0.12	0.83	0.12
Differential between pre/post	0.12	0.08	0.09	0.13
Global workload	52.26	12.03	60.14	10.94
Promotor score	8.00	1.66	8.30	2.06

Age

An independent samples T test did not indicate a significant difference for age between ARES condition (*M* = 22.30, *SD* = 3.02) and the flat condition (*M* = 21.33, *SD* = 0.50); *t*(17)= –0.946, *p* = 0.357.

Gender

Out of the 19 participants, 17 were male and 2 were female. Both females participated in the flat condition. For future studies, the ratio of women in each condition will need to be equalized.

Class Year

All participants had completed MS-300 at the time of the experiment. With the exception of one graduate student, all other participants were seniors.

College Majors

Out of 19 participants, 4 individuals majored in political science, 6 in criminal justice, 2 in mechanical engineering, and the rest were a variety of different majors (e.g., finance, history, sports, and exercise science)

Number of Hours of Sleep

An independent samples T test did not indicate a significant difference for number of hours of sleep by participants in the ARES condition (*M* = 6.2, *SD* = 1.32) and the flat condition (M = 5.56, SD = 1.67); t(17)= –0.940, p = 0.360.

Energy Level

An independent samples T test did not indicate a significant difference for energy level by participants in the ARES condition (*M* = 73.40, *SD* = 21.46) and the flat condition (*M* = 74.00, *SD* = 20.17); *t*(17)= 0.063, *p* = 0.951.

Vision Problems

Out of the 19 participants, 15 reported normal vision, 7 in the ARES condition, and 8 in flat condition. Two individuals, both in the ARES condition, reported corrected-to-normal using glasses, and one participant in each condition reported corrected-to-normal using contact lenses.

Confidence in Computer Use

An independent samples T test did not indicate a significant difference for confidence in computer use by participants in the ARES condition (*M* = 75.80, *SD* = 22.18) and the flat condition (*M* = 78.56, *SD* = 18.24); *t*(17)= 0.294, *p* = 0.773.

Video Game Frequency

An independent samples T test did not indicate a significant difference for video game frequency by participants in the ARES condition (*M* = 3.70, *SD* = 1.89) and in the flat condition (*M* = 3.11, *SD* = 1.97); *t*(17)= –0.666, *p* = 0.514 (Scale of 1– 7, 1 = never and 7 = often).

Knowledge of Military Tactics

An independent samples T test did not indicate a significant difference for knowledge of military tactics by participants in the ARES condition (*M* = 69.40, *SD* = 14.52) and in the flat condition (*M* = 73.11, *SD* = 21.66); *t*(17)= 0.443, *p* = 0.663.

Caffeine Consumption in the Last 2 Hours

Only participants 11, 12, and 15 had caffeine within the last 2 h.

Score on Mental Rotation Test

An independent samples T test did not indicate a significant difference for score on mental rotation test by participants in the ARES condition (*M* = 0.41, *SD* = 0.18) and in the flat condition (*M* = 0.44, *SD* = 0.22); *t*(17)= 0.321, *p* = 0.752.

Self-Assessment Manikin

A paired samples T test examining self-assessment manikin scores pre- and postexperiment split between conditions indicated a significant difference for participants in the ARES condition on the arousal (M_{pre}= 2.40, SD_{pre} = 0.97; M_{post}= 2.80, SD_{post}= 1.14; *t*(9) = –2.45, *p* = 0.037) and dominance scales (M_{pre}= 3.50, SD_{pre} = 0.53; M_{post} = 3.90, SD_{post} = 0.57; *t*(9) = -2.45, *p* = 0.37). A paired samples T test did not indicate a significant difference for participants in the ARES condition for pleasure (M_{pre}= 3.90, SD_{pre} = 0.74; M_{post}= 4.10, SD_{post}= 0.74; *t*(9) = –0.802, *p* = 0.443).

A paired samples T test examining self-assessment manikin scores pre- and postexperiment split between conditions was also done for the flat condition. The following paired samples T tests did not indicate significant differences for all scales: pleasure (M_{pre}= 3.67, SD_{pre} = 0.87; M_{post}= 4.00, SD_{post}= 0.71; t(8) = –1.41, p = .20), arousal (M_{pre}= 2.89, SD_{pre} = 0.928; M_{post}= 3.00, SD_{post}= 0.50; t(8) = –0.426, p = 0.68) and dominance scales (M_{pre}= 4.00, SD_{pre} = 1.00; M_{post} = 4.11, SD_{post} = 0.928; t(8) = –0.56, p = 0.59).

Pre- and Postconceptual Test Scores

A paired samples T test examining military conceptual test scores pre- and postexperiment split between conditions indicated a trend toward significance for participants in the ARES condition (M_{pre}= 0.74, SD_{pre} = .11; M_{post}= 0.83, SD_{post}= 0.12; t(9) = –2.20, p = 0.055) and a significant difference for participants in the flat condition (M_{pre}= 0.67, SD_{pre} = 0.09; M_{post}= 0.79, SD_{post}= 0.12; t(8) = –4.47, p = 0.002).

Tactics Scores

An independent samples T test did not indicate a significant difference for score on the 8 tactics questions by participants in the ARES condition (M = 5.40, SD = 1.50) and in the flat condition (M = 5.11, SD = 1.90); t(17) = –0.369, p = 0.716 (Table 2).

Table 2 Scores by question

Condition		N	Mean	Std. deviation
Q1	Flat	9	0.56	0.53
	ARES	10	0.80	0.42
Q2	Flat	9	0.78	0.44
	ARES	10	0.40	0.52
Q3	Flat	9	0.67	0.50
	ARES	10	0.80	0.42
Q4	Flat	9	0.78	0.44
	ARES	10	0.90	0.32
Q5	Flat	9	0.44	0.53
	ARES	10	0.40	0.52
Q6	Flat	9	0.44	0.53
	ARES	10	0.60	0.52
Q7	Flat	9	0.67	0.50
	ARES	10	0.60	0.52
Q8	Flat	9	0.78	0.44
	ARES	10	0.90	0.32

N = number of subjects

Tactics Time

An independent samples T test did not indicate a significant difference for time to answer the 8 tactics questions by participants in the ARES condition (*M* = 308.84, *SD* = 116.08) and in the flat condition (*M* = 355.07, *SD* = 132.01); *t*(17) = 0.813, *p* = 0.428.

Specific times by question are listed in Table 3. For all questions with the exception of the first, participants in the ARES condition answered more quickly.

Table 3 Times according to condition

Condition		N	Mean	Std. deviation
Q1_Time	Flat	9	44.13	17.88
	ARES	10	55.49	30.71
Q2_Time	Flat	9	44.52	28.91
	ARES	10	35.04	19.97
Q3_Time	Flat	9	40.93	31.65
	ARES	10	33.30	17.06
Q4_Time	Flat	9	36.94	19.45
	ARES	10	31.25	14.83
Q5_Time	Flat	9	41.48	19.80
	ARES	10	39.30	18.25
Q6_Time	Flat	9	51.15	34.34
	ARES	10	47.91	26.92
Q7_Time	Flat	9	69.48	34.34
	ARES	10	45.51	14.30
Q8_Time	Flat	9	26.43	11.26
	ARES	10	21.25	10.95

N = number of subjects

NASA-TLX

An independent samples T test did not indicate a significant difference for global workload according to the NASA-TLX for participants in the ARES condition (*M* = 60.14, *SD* = 10.94) and in the flat condition (*M* = 52.26, *SD* = 12.03); *t*(17) = –1.496, *p* = 0.153.

Means and standard deviations according to individual workload scales are listed in Table 4.

Table 4 NASA-TLX workload scores

Scale		N	Mean	Std. deviation
Mental demand	Flat	9	59.78	23.68
	ARES	10	76.20	22.85
Physical demand	Flat	9	9.33	11.12
	ARES	10	5.30	8.65
Temporal demand	Flat	9	43.56	25.17
	ARES	10	38.70	25.82
Performance	Flat	9	61.78	25.96
	ARES	10	78.60	15.71
Effort	Flat	9	65.44	20.26
	ARES	10	62.50	18.40
Frustration	Flat	9	40.33	36.66
	ARES	10	24.80	25.38

N = number of subjects

Promoter Score

When asked after participating in the experiment "How likely is it that you would recommend this system to a friend or colleague?" Individuals scored very closely with ARES (*M* = 8.30, *SD* = 2.06) and the flat condition (*M* = 8.00, *SD* = 1.66).

This question is what is known as the net promoter score (NPS). The score of 8 indicates individuals who are passives. "They are satisfied but unenthusiastic customers who are vulnerable to competitive offerings" (NPS). In terms of ARES this indicates that while it achieved their desired goals, they may not be as motivated to use it with the current capabilities.

Whether Participant Felt the System Helped to Learn Tactics

When asked after participating in the experiment whether they felt the system helped to learn tactics, individuals in both conditions scored very closely with one another: ARES (*M* = 3.80, *SD* = 0.79) and the flat condition (*M* = 3.56, *SD* = 0.73). This indicates that the participants are between the neither and agree points on the scale. This is consistent with the promoter score discussed earlier (Scale of 1–5, Strongly Disagree, Disagree, Neither, Agree, Strongly Agree).

Whether Participant Felt They Would Use System to Practice in Spare Time

When asked after participating in the experiment whether they felt they would use the system to practice in their spare time, individuals in both conditions again scored closely with one another: ARES (*M* = 4.20, *SD* = 0.42) and the flat condition

(*M* = 3.78, *SD* = 0.83). However, for this question, participants in the ARES condition did move into the agree range of the scale (Scale of 1–5, Strongly Disagree, Disagree, Neither, Agree, Strongly Agree).

Whether Participant Preferred the Condition They Were in or the Other Condition

At the end of the experiment, each participant was shown the opposite condition from the one they participated in. Following are the comments from each participant. The majority of the participants favor the ARES condition.

> P1 (experimental condition – ARES): preferred ARES
> "I would have preferred the one I was in [ARES]…I mean, it's just…the flat is fine but it doesn't add any depth…I think the sand table, as simple as it is, allows you to kind of more easily see and more easily connect to the terrain in the end."
>
> P3 (experimental condition – ARES): preferred ARES
> "I think having the table makes it a little bit easier in terms of line of sight and actually seeing distance, because with this [map on flat surface] it seems more straight, but obviously if you're going over hills you can see on the sand table it makes it longer or that you're more hidden."
>
> P4 (experimental condition – ARES): preferred ARES
> "I would have preferred the other one [referring to the condition he was in] because if you have basic map reading skills you know that these are areas of elevation and everything, locations of rivers and roads and stuff, but it helps to engage whomever you're telling the plan to more. They can see the mountain, and it just offers a better view for the situation."
>
> P5 (experimental condition – ARES): preferred ARES
> "Definitely the sand table, the one I was in. Because just looking at this right now, you can just tell. It is way more direct when you see it, and you can judge how it's laid out better. Part of your mind is thinking about what these lines mean, whereas if you're looking at the sand table, you're not even worried about that, you're looking at what the environment actually is. You don't have to try and decipher anything."
>
> P6 (experimental condition – ARES): preferred ARES
> "The one I was in. It's easier to identify all of the terrain features rather than looking at the flat map."
>
> P7 (experimental condition – ARES): preferred ARES

"I would definitely prefer the sand table. We're all used to using these maps but it's different when you can actually see the contour lines and how the enemy may use the terrain because it's just like if you were actually on the ground looking at the terrain; that's always better than looking at a map. So looking at something that is a mixture of terrain and a map is always nice. I thought that I didn't have to focus on what the topographical map looked like, I could look at the sand table and I could see based off of that and that's what I based a lot of my judgment off of was the sand table. I think it would be fantastic to use in teaching classes because it was… I just imagined the whole time if I was here teaching with it and talking about why I would do this or that I thought it would be good for that and learning tactics and avenues of approach."

P8 (experimental condition – ARES): preferred ARES
"I think it's pretty obvious the one with the sand table. I would prefer that better. Just looking at this flat surface it goes into your ability to understand map reading and symbols and what things look like but with it being flat, understanding that that's a hilltop, being able to see the hilltop and how it flows around it makes it a lot easier to make decisions because it gives you a 3-D view actually of the ground. Easier to make decisions with the 3-D."

P9 (experimental condition – ARES): preferred ARES
"I would prefer the condition I was in because you can see what it actually looks like. If you're not familiar with how to read a map, it could be a little bit more difficult, even if you are familiar with reading a map, it still gives you an actual better perception of what you're actually looking at so it just gives you a better idea of everything because this is nice to have, the flat map, but definitely the 3-D map helps a lot better in making any of the decisions. I probably would have had different decisions, I wouldn't really say different decisions, but I would have had to think a lot harder on my decisions with this map [on a flat surface] as opposed to the other map that I was using."

P10 (experimental condition – flat): preferred ARES
"This one [ARES]. Because not only do I get to play in the sand, but I just think it gives you more of what I'm actually looking at. Like yes you need to know the lines and what they mean, but this you have the actual picture. Say I'm carrying tanks with me, they can only go slowly down the hill, this is more of a picture I could be looking at, and it makes sense for not only me but I mean, my Soldiers are going to be fresh out of high school, so pictures are more knowledgeable for Soldiers I think."

P11 (experimental condition – flat): preferred ARES
“This one [ARES] for sure, without a doubt, makes it more 3-D…I mean I saw the elevation just off the contour lines and all that, but this puts it…you have to focus less on map reading, and you can focus more on like strategy."

P12 (experimental condition – flat): preferred flat
"Ahhh [when viewing ARES]…oh yeah, then I definitely would have picked B [referring to tactics question 1]…yeah that's terrible. Definitely this condition, I mean you just get a better view of everything. Like how it looks and the way the surrounding area is. Like I picked that over there for my ORP (objective rally point), definitely over here B. Yeah because you’re behind a hill and they can't really...if you walk, I mean you could walk this way and they couldn't really see you because even at this point like maybe like right there but they're still...yea no they're way too far yea definitely B.”

P13 (experimental condition – flat): preferred ARES
"I would prefer this one [ARES] because you actually get more 3-D visual of the terrain that you're dealing with…usually the other one is more of a flat surface, and yea it has the contour lines and shows the elevation, but this gives you more of a realistic view.”

P14 (experimental condition – flat): preferred ARES
"Oh it is the same…I would prefer this one [ARES] because you can actually see the elevation…like it's projected out… so when you explain situations like to another person they can see what you're talking about…like this…obviously this is higher elevation than somewhere over here...you can see it...clear."

P15 (experimental condition – flat): preferred flat
“Well I kind of like…this one [ARES] is obviously easier…but you know…given…I mean it's a pretty simple map to read so being able to visualize…you know…I guess the texture of the land is a pretty easy assessment with this map, but if it were…I guess if it were done with just like a...one of these satellite images, I think it might be a little bit more difficult, because yea you have like an actual bird's eye view of everything but it’s not going to spell out…you know inclines/declines and all these different measurements and numbers so yea this is easier to read, but this use of map is easier to read than a satellite image to me, but I don't know that having a three dimensional like actual sand table is a lot harder to do so I would stick with these regular like flat maps...it's cool though."

P16 (experimental condition – flat): preferred ARES
"This one [ARES] because it gives you a better view, instead of just showing you that there's a hill here you look and you see there's a hill here as opposed to just like - alright on this map there's a hill - you look down here and you can visually see the hill with the topographic symbols on it."

P17 (experimental condition – flat): preferred ARES
"This [ARES]… this is a lot better you can see like actually 3-D because like you were saying if you don't…like planar or like a linear surface like…that…you can see a lot more through this, like the hills, the…yea you can just see a lot more through it this way so it gives you a lot better...like for avenues of approach where you want to have your guys at...so it's a lot better in my opinion, prefer this a lot more."

P18 (experimental condition – flat): preferred ARES
"I was actually thinking about that…it would be better if you could project it on a sandtable. I prefer this one [ARES] just because it's a lot more...natural, I would say, to look at it because now you are in a situation like a cloud or a plane, you're looking down on the objective, with being able to see, oh yea this one right here is slightly higher than this one (pointing between 2 hills), which you wouldn't be able to directly notice on a map right away, so it makes things a lot more clear and faster."

P19 (experimental condition – ARES): preferred ARES
"Definitely the condition I was in. We've used these [flat maps] for four years now, and I'm sure that the army will keep using them, but I like the other one because you can actually see the draws and everything and actually look at the elevation. On this one you can't really tell the difference between that peak and this peak, or that peak and this objective. You can't really see that you're going to be able to see here if you're over here."

P20 (experimental condition – ARES): preferred ARES
"I like the other one, the sand one. It gave a realistic view of the contouring and B while I am trying to conduct the operation. It's better than just the contour lines, you can't know exactly how high up it is or the structure of that mountain. I think it makes it a lot more realistic at least planning wise so you could come up with your best way to successfully complete the operation."

Comments Related to Things They Did Not Like About System

"No compass to assess direction." P15

"No direct manipulation." P18

Comments on Potential Improvements for the System

> "Ability to move element icons in accordance with tactical plan." P4
>
> "I would like to see it expand more and be able to display missions occurring and all the moving pieces." P14
>
> "Labels for the overlays that were used." P14
>
> "Add a way to see physical movement beyond the use of arrows." P15
>
> "More notable distances shown." P16
>
> "More hands-on approach…" P17

Reason They Felt System Helped To Learn Tactics

> "Good application of real tactics and land navigation." P1
>
> "Helped to reinforce ideas. Made it easier to visualize." P3
>
> "Provides a better view of the overall positions of elements/objectives." P4
>
> "Allows you to look at how the terrain can enhance an operation." P5
>
> "It gave a good aerial view of how the battlefield looks in respect to the mission." P12
>
> "I thought more critically about different scenarios." P15

Reason Why They Would Be Willing to Use it to Practice

> "Easy to use." P1
>
> "I thought it was fun to use." P3
>
> "Offers a more sophisticated view of tactical situations." P4
>
> "Map was useful." P11
>
> "It gets you to actually think more about tactics regarding positions where you would place your elements." P13

5.2 Physiological Data Results

Out of the 19 participants, 2 had data that were incomplete. Therefore, 10 participants from the ARES condition and 7 participants from the flat condition are included in the following analysis. Out of the 17 remaining participants, 4 participants from the ARES condition, and 5 participants from the flat condition had SCRs according to the criteria outlined previously. SCRs according to condition and individual question are broken down in the Table 5.

Table 5 Skin conductance response by condition

ARES	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Totals
P4	1	0	0	0	0	0	1	0	2
P5	1	0	0	0	1	0	0	0	2
P7	0	1	1	1	1	1	0	0	5
P8	0	0	0	1	0	1	1	0	3
Total	2	1	1	2	2	1	2	0	12
Flat									
P10	0	0	0	0	1	0	0	0	1
P11	0	1	1	0	0	0	0	1	3
P13	0	1	1	0	0	0	0	0	2
P15	1	1	1	1	0	0	1	1	6
P16	1	1	1	0	1	1	0	1	6
Total	2	4	4	1	2	1	1	3	18

As can be seen from the table, the results are in the opposite direction with respect to H_1 with more SCRs occurring in the flat condition. An independent samples t-test was conducted to compare the number of SCRs (for participants who had at least one SCR) and the experimental condition. An independent samples T test did not indicate a significant difference for the number of SCRs for the ARES condition ($M = 3.60$; $SD = 2.30$) and the flat condition ($M = 3.00$; $SD = 1.41$); t (7) = 0.454, $p = 0.66$.

Additional Exploratory Analysis on Participant Data

To get a better explanation of the results, descriptive data from each participant were examined in terms of their responses and movement around the table as they answered:

ARES.

P4 (Male, Average Temperature 94 °F): P4 had SCRs on the first and seventh questions. In comparing his performance data, on the first question he got the answer correct and took 58.50 s, and on the seventh question he got the answer correct and took 65.65 s (total score 5/8, average time 62.33 s).

In terms of body language and movement around the table, he began the questions facing the screen at the long side of the table. For the first question he rested his index fingers on the ledge and leaned forward onto them. He then mouthed the question, looking at the screen, followed by moving 90° right to the short side of the table and rested his hands on the ledge, again leaning on them. For the seventh question he demonstrated similar movement as the first question. He moved from long side 90° right to the short side of table, while resting his fingers on the ledge, and looking at the sand table.

P5 (Male, Average Temperature 94 °F): P5 had SCRs on the first and sixth questions. On the first question he got the answer correct and took 27.55 s, and on the sixth question he got the answer correct and took 24.78 s (total score 6/8, average time 25.15 s). In terms of body language and movement around the table, on both questions he rested his hands on the table ledge but didn't change position.

P7 (Male, Average Temperature 92 °F): P7 had SCRs on questions 2–6. He got all the answers correct with times of 20.23, 14.43, 14.86, 30.42, and 25.63 s, respectively (total score 8/8, average time 21.10 s). In terms of body language and movement around the table, for Q2 he moves right to the long side of the table, facing the screen and with his fingers resting on the edge. He remains in this position throughout the other SCRs.

P8 (Male, Average Temperature 95 °F): P8 had SCRs on the fourth, sixth, and seventh questions. On the fourth question he got the answer right and took 36.60 s. On the sixth question he got the answer wrong and took 47.08 s, and on the seventh question he got the answer right and took 67.72 s (total score 4/8, average time 48.39 s). In terms of body language and movement around the table, prior to Q7 he moves along the long side of the table, shifts his weight and taps his fingers against the sides of the table.

Flat

P10 (Female, Average Temperature 95 °F): P10 had an SCR on the fifth question, which she got wrong and took 60.80 s (total score 2/8, average time 49.90 s). In terms of body language and movement around the table, she moved her position from looking at the map from the short side of the table, to looking at the map from the long side of the table.

P11 (Male, Average Temperature 93 °F): P11 had 3 SCRs on questions 2, 3, and 8. For question 2, he got the answer correct and took 44.17 s. On question 3, he got the answer correct and took 30.82 s, while on question 8 he got the question correct and took 21.33 s (total score = 7/8, average time 28.69 s). In terms of body language and movement around the table, for Q2 he moves 180° to the other side of the table, moves with his hands in his pockets, and then moves another 90° to the left long side of the table. For Q3 he also moves around the table from the starting short side to the right long side with hands in his pockets.

P13 (Male, Average Temperature 95 °F): P13 had 2 SCRs during questions 2 and 3. For question 2 he got the answer correct and took 54.72 s, and for question 3 he got the answer correct and took 47.94 s (total score = 6/8, average time 48.40 s). In terms of body language and movement around the table, for Q2 he moves down the side of the table and clasps his hands. He did not change position while answering Q3.

P15 (Male, Average Temperature 93 °F): P15 had 6 SCRs on questions 1–4 and 7 and 8. His scores and times are as follows: Q1, wrong, 54.50 s; Q2, correct, 35.29 s; Q3, wrong, 25.23 s; Q4, correct, 71.94 s; Q7, wrong, 99.84 s; and Q8, correct, 36.85 s (total score 3/8, average time 61.10). In terms of body language and movement around the table, he moved to the long side of the table during Q1 but did not change position once there.

P16 (Male, Average Temperature 91 °F): P16 had 6 SCRs on questions 1–3 and 5, 6, and 8. His scores and times are as follows: Q1, wrong, 33.78 s; Q2, correct, 21.47 s; Q3, correct, 21.47 s; Q5, wrong, 29.12 s; Q6, wrong, 24.95 s; and Q8, wrong, 24.08 s (total score 4/8, average time 24.96 s). In terms of body language and movement around the table, he stayed on the long side of the table. During Q1 he tilts his head to the right, during Q5 he chews on his lip, and during Q8 he tilts his head to the right.

5.3 Qualitative Data Results

5.3.1 Step 1: Reading and Re-reading

As a first step, a list was developed to understand what biases might be present. The list is as follows:

1) The cadets will vocalize everything that is important to them. However, it is possible that they process information in a way that skips steps along the process, thereby missing important decisions.

2) The displays presented are viewed as novel by the cadets. Being that most people have not seen a digitally augmented sand table, the assumption is made that this is new and they will not leverage from prior experience.

3) The cadets are not performing at an expert level and will use the system to help facilitate their learning. It could be the case that, although they are cadets, they have achieved enough experience and proficiency to produce at an expert level.

4) The questions are difficult enough to have a need for decision-making. If the questions are too easy on the other hand, the responses may be simplified and not as detailed.

Transcripts are as close to verbatim as possible, reviewing video clips multiple times to verify phrasing by participants.

5.3.2 Step 2: Initial Coding

Two independent coders reviewed each of the transcripts and highlighted specific segments that were coded into 48 categories. Table 6 is a short excerpt from P4 with the associated codes used by one of the coders.

Table 6 Example coding from single participant

Content	Coding
If your ORP is here, it looks like you could maneuver back behind a little bit, come down here, and then use this area of elevation as some covering and concealment, to maneuver toward the objective.	Refers back to common practices
Whereas here you would be out in the open, coming in between these 2 areas. They would be able to see you.	Desire to minimize visibility
Also these 2 would take way too long to go all the way over here and try to hit them from behind	Reference to time

The coding for each participant was done using the existing codes that were already created. Once a code was discovered it was then added to the list and after finishing that individual transcript, the previously coded transcripts were revisited to see where they might be applicable.

The new codes were also communicated daily between each rater, with a brief definition of what that code means (without giving a specific example in the text).

Once the initial codes were created, the 2 coders examined their individual codes for discrepancies. There were a total of 126 points of contention across the 19 participants with an average of 7 points of contention per participant. However, this number is inflated because several of the points occurred more than once (most often due to a difference in interpretation of the meanings of the codes).

Once the issues with the codes were rectified, a second pass through the coding process was done by each coder. The emphasis this time was to be able to group various codes together to get to a second reduced set of codes.

The reduced set of codes is as follows (Table 7):

Table 7 Reduced set of codes

Code	Description	Code	Description
1	Use of the phrase "mission dependency(ies)"	9	Unfamiliarity (use of any term that indicates lack of knowledge)
2	Avenue of retreat	10	Mention of the word "road"
3	Debating between only 2 options	11	Minimization of risk
4	Desire to minimize distance	12	Prediction of enemy approach
5	Desire to minimize visibility	13	Reference to time
6	Identification of terrain feature	14	Relevance of waterways
7	Elevation discrepancy favors one force	15	Mention of fatigue
8	Forms of maneuver	16	Use of the phrase "line of sight"

Once the initial and secondary codes were created, the next step was to verify our codes by examining the video recordings. Both researchers watched all of the video recordings together.

5.3.3 Step 3: Developing Emergent Themes

What was noticed as the team was reviewing the video footage was several common threads which continue to arise within the participants. From these the following emergent themes were created:

1) Minimizing Uncertainty: This theme is defined as the cadet making the selection to choose one option over another based on their knowledge of the enemy situation related to one of the positions.

2) Minimizing Visibility: This theme refers to the cadet choosing an option because it allows them to stay hidden, avoiding detection by the enemy.

3) Distance with Respect to Spacing: This theme refers to the cadet adjusting his or her planning due to being too close or too far away to make that selection an ideal choice.

4) Distance with Respect to Fatigue: This theme refers to the cadet's consideration for the fatigue of the squad that they are moving with. Based on the concern of fatigue, they choose one option over another.

5) Distance with Respect to Time: This theme refers to the cadet making their decision based on how long it will take to cover a particular swath of terrain.

6) Relevance of Waterways: This theme refers to the consideration of waterways when the cadets are making their decisions. This could include both the use and avoidance of waterways.

7) Elevation with Respect to Height: This theme reflects the cadet using elevation heights to determine which of the options they were going to select.

8) Elevation with Respect to Advantage/Disadvantage: This theme represented the cadet's awareness on how they positioned themselves in relationship to the enemy's position and how that provides a tactical impact for one side or the other.

Together these themes helped to establish higher level meaning while at the same time holding true to the essence of what was in the transcripts.

5.3.4 Step 4: Searching for Connections across Emergent Themes

This next step consists of looking across the emergent themes and fitting them together into a coherent story. To accomplish this, a process called abstraction was used. Abstraction looks for identifiable patterns and then uses those patterns to come up with "super-ordinate" themes (Smith et al. 2009). Based on the 8 emergent themes discussed earlier, 3 super-ordinate themes were created: Elevation Discrepancies; Distance; and Cover and Concealment.

1) Elevation Discrepancies: Since elevation can indicate multiple relationships related to height differential (high/low, friendly/enemy, positive/negative), the emergent themes of elevation with respect to height and elevation with respect to advantage/disadvantage were collapsed into elevation discrepancies.

2) Distance: Distance was represented in multiple ways. Cadets appeared to use distance as a part of their decision-making but prioritized aspects of distance differently. Some individuals made their decision using distance as a measure of time. Others used distance in terms of the level of fatigue that it was going to cause their squad. Further, there were also situations where distance was a matter of spacing between themselves and the opposing forces. Therefore the 3 distance emergent themes were collapsed into the superordinate theme of distance.

3) Cover and Concealment: This was a combination that also had multiple aspects associated with it. Some individuals used cover and concealment as a means to minimize uncertainty, while others didn't specifically mention uncertainty but were concerned about visibility. Therefore, those 2 categories (Minimization of Uncertainty and Risk) were collapsed into the superordinate theme of cover and concealment.

6. Discussion

6.1 Research Hypotheses Summary

The overall finding for the study was that there were not significant differences between conditions. It is possible that the major factor behind this result is the lack of complexity of the scenario content. Future iterations of the study will increase the complexity to help further delineate how ARES/GIFT performs under more demanding task load. Further, the design will be modified to allow for individual differences analysis between participants.

6.1.1 ARES Increase in Workload

ARES will demonstrate an increase in cognitive workload for cadets in comparison to the flat display condition, as assessed by the global workload score of the NASA-TLX (Hart and Staveland 1988), since ARES is a novel interface compared to the control condition.

An independent samples T test did not indicate a significant difference between the 2 conditions, but the ARES condition did show an increase in workload (M = 60.14, SD = 10.94) in comparison to the flat condition (M = 52.26, SD = 12.03) (Fig. 3).

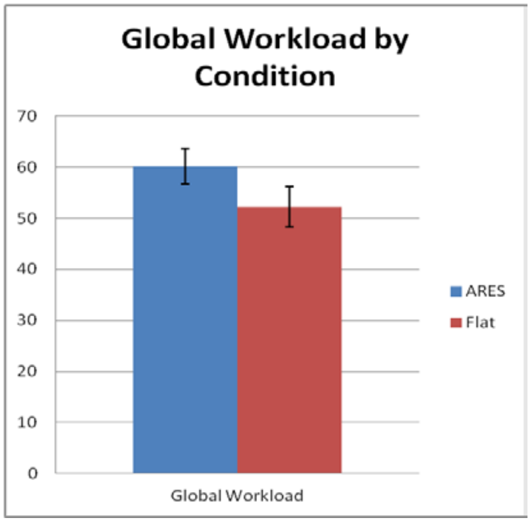


Fig. 3 Global workload (error bars indicate standard error of the mEAn)

Upon further examination, the mental workload was higher, but not significant, for the ARES condition than for the flat condition which could possibly be due to the novelty effect of ARES (Fig. 4).

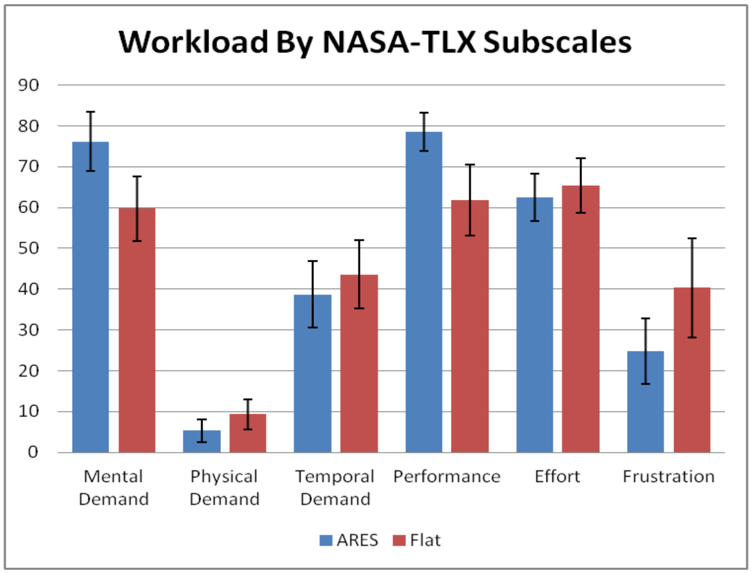


Fig. 4 Subscale workload between conditions (error bars indicate standard error of the mean)

6.1.2 ARES Higher Levels of Accuracy

ARES will produce higher levels of accuracy on the assessment questions than the flat condition due to the additional information provided by the contoured display.

An independent samples T test did not indicate a significant difference for score on the 8 tactics questions by participants in the ARES condition (M = 5.40, SD = 1.50) and score on the 8 tactics questions by participants in the flat condition (M = 5.11, SD = 1.90) (Fig.5).

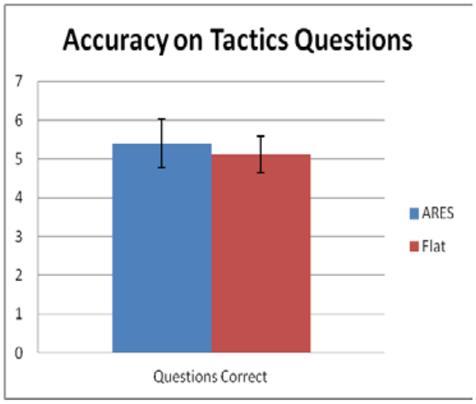


Fig. 5 Accuracy between conditions (error bars indicate standard error of the mean)

6.1.3 ARES More Time on Task

Individuals in the ARES condition will take more time on task, due to increased information presented in the display, than individuals in the flat display.

An independent samples T test did not indicate a significant difference for time to answer the 8 tactics questions by participants in the ARES condition (M = 308.84, SD = 116.08) and time to answer the 8 tactics questions by participants in the flat condition (M = 355.07, SD = 132.01) (Fig. 6).

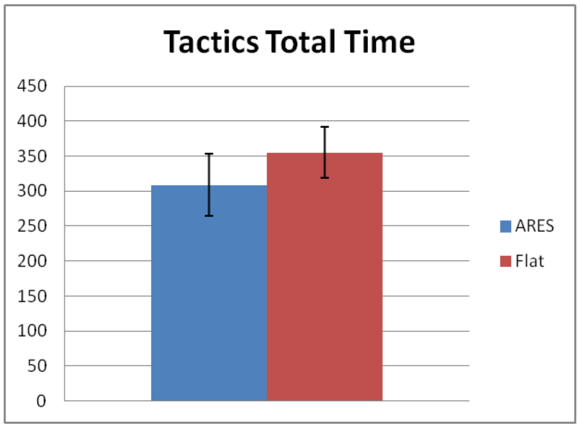


Fig. 6 Total time answering questions between conditions (error bars indicate standard error of the mean)

6.1.4 ARES Higher Electrodermal Activity

Use of ARES will show a greater overall increase in electrodermal activity than the flat condition due to an increase in arousal.

The data actually showed a difference in the opposite direction than expected, with more electrodermal activity being represented in the flat condition. It is possible that underlying this increased activity could be an increase in frustration by the participants in the flat condition (nonsignificant). However, more research needs to be done to verify this.

6.2 Study Limitations

This study had limitations that can provide insight but need to be addressed before the research can be applied on a larger scale.

- Potentially, the sample size was not large enough to generate enough power in the context of a between-subjects design. To accommodate this, the study will be expanded in collaboration with the United States Military Academy (USMA) at West Point starting in the late summer/early fall of 2016.
- The exposure to only one map format. Participants in this study were only presented one type of map surface (flat or contour, not both), which did not allow for the comparisons within an individual and their interactions

between the 2 displays. It is possible that the format that the participant did not experience may have caused a difference in performance or in their physiological output, which could not be detected due to the study design. To accommodate this limitation, a within-subjects design will be used in the expanded study mentioned in limitation one.

- The qualitative data analysis and the mapping between body movements and participants had to be done manually. This naturally leads to bias and potential error in interpretation. However, there is not a clear objective way to use a technique like IPA. The follow-on study will have to more closely address any process that was accomplished manually and determine if there is a more objective methodology.
- Although small, there was a gender discrepancy between the 2 conditions with both female participants being in the flat condition. If the results are judged by an Army population, it may not be possible to get equal gender balance, but it will be important in future studies to ensure that the ratios are the same.

6.3 Directions for Future Work

One of the challenges with using research products such as GIFT and ARES is that the technology is not fully developed. First, future efforts should continue to strengthen the interoperability between these 2 systems and provide relevant data that can support military adaptive training. Besides the use cases mentioned in the introduction, it can also be used as a way to demonstrate signal interference due to terrain, emergency management planning (specifically crowd and flow modeling of evacuation routes), and can be used by human-robot teams learning how to navigate extraterrestrial landscapes.

Second, an expansion into different populations (or different subsets) of military learners would help to better understand the difference between a military cadet and an operational Soldier. One of the goals of the ARES program is to be able to demonstrate quantitative value in terms of training using ARES in an operational context. To that end, a table has been placed with the 3rd Infantry Division at Fort Stewart in Hinesville, Georgia, to support a study on Soldier performance using ARES in fiscal year 2016. With the proper collaborators who can provide the relevant content needed for their learners, it is possible that we can analyze across experimental studies and multiple populations of learners.

Finally, a goal of the GIFT program is to be able to provide easy-to-use adaptive training for inexperienced users. Currently a project is underway called GIFT-Wrap to assist in building training across various simulation platforms. In terms of tactics, it would be interesting to explore, quantitatively, how fast an instructional lesson can be created by a novice user. Can a USMA instructor create a tactical decision exercise including questions associated with it and then test their cadets? All of these ideas and others lead to a promising future for the ARES/GIFT integration project.

7. Conclusion

This research effort provide qualitative and quantitative data on how tactics learning can be supported using different types of projection surfaces. The data show promise in terms of arousal as well as examining across different types of workload. The results from this study demonstrate the feasibility of the integration between ARES and GIFT. Lessons learned will be applied to the full experimental study set to begin in fall 2016 at USMA.

8. References

Amburn CR, Vey NL, Boyce MW, Mize JR. The augmented reality sandtable (ARES). Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2015 Oct. Report No.: ARL-SR-0340.

Ary D, Jacobs L, Sorensen C, Walker D. Introduction to research in education. Independence (KY): Cengage Learning; 2013.

Azuma RT. A survey of augmented reality. Presence. 1997;6(4):355–385.

Baker RS, DeFalco JA, Ocumpaugh J, Paquette L. Toward detection of engagement and affect in a simulation-based combat medic training environment. Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym2); 2015 Jun 12–13; Pittsburgh, PA. Orlando (FL): Army Research Laboratory (US); c2015. p.111–121.

Barfield W, Rosenberg C. Judgments of azimuth and elevation as a function of monoscopic and binocular depth cues using a perspective display. Human Factors: The Journal of the Human Factors and Ergonomics Society. 1995;37(1):173–181.

Bevan N, Macleod M. Usability measurement in context. Behaviour & Information Technology. 1994;13(1–2):132–145.

Boyce MW. Assessment of instructional presentation for emergency evacuation assistive technology [dissertation]. [Orlando (FL)]: University of Central Florida; 2014.

Boyce MW, Sottilare RA, Goldberg BS, Amburn CR. Work domain analysis for ecological interface design of tangible interfaces. Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2015 Nov 30–Dec 4; Orlando, FL. http://www.iitsecdocs.com/.

Boucsein W. Electrodermal activity. New York (NY): Springer Science & Business Media, 2012.

Botella C, Bretón-López J, Quero S, Baños R, García-Palacios A. Treating cockroach phobia with augmented reality. Behavior Therapy. 2010;41(3):401–413.

Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry 1994;25(1):49–59.

Brown JS, Collins A, Duguid P. Situated cognition and the culture of learning. Educational Researcher. 1989;18(1):32–42.

Bruner JS. Toward a theory of instruction. Cambridge (MA): Harvard University Press; 1966.

Bujak KR, Radu I, Catrambone R, Macintyre B, Zheng R, Golubski G. A psychological perspective on augmented reality in the mathematics classroom. Computers & Education. 2013;68:536–544.

Chen JY, Barnes MJ. Human–agent teaming for multirobot control: a review of human factors issues. Human-Machine Systems, IEEE Transactions On. 2014;44(1):13–29.

Cunningham DJ. Beyond educational psychology: Steps toward an educational semiotic. Educational psychology review. 1992;4(2):165–194.

Headquarters, Department of the Army. The operations process. Washington (DC): Headquarters, Department of the Army; 2012 May. ADRP 5-0.

Headquarters, Department of the Army. Geospatial engineering. Washington (DC): Headquarters, Department of the Army; 2014 Jun. ATTP 3-34.80.

Headquarters, Department of the Army. Offense and defense. Washington (DC): Headquarters, Department of the Army; 2013 Mar. FM 3-90.

D'mello SK, Kory J. A review and meta-analysis of multimodal affect detection systems. ACM Computing Surveys (CSUR). 2015;47(3):43.

Driscoll M. Psychology of learning for instruction. Needham Heights (MA): Allyn and Bacon; 2000.

Dunleavy M, Dede C. Augmented reality teaching and learning. Handbook of research on educational communications and technology. New York (NY): Springer; 2014. p. 735–745.

Dünser A, Steinbügl K, Kaufmann H, Glück J. Virtual and augmented reality as spatial ability training tools. Proceedings of the 7th ACM SIGCHI New Zealand chapter’s international conference on Computer-human interaction: design centered HCI. 2006 Jul 6–7; Christchurch, New Zealand. New York (NY): Association for Computer Machinery; c2006. p. 125–132.

Goldberg B, Amburn C, Brawner K, Westphal M. Developing models of expert performance for support in an adaptive marksmanship trainer. Paper presented at: Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC); 2014 Dec 1–4; Orlando, FL. http://www.iitsecdocs.com/.

Hall DS, Shattuck LG, Bennett KB. Evaluation of an ecological interface design for military command and control. Journal of Cognitive Engineering and Decision Making. 2012; 6(2):165–193.

Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Advances in Psychology. 1988;52:139–183.

Hernandez J, Morris RR, Picard RW. Call center stress recognition with person-specific models. Affective computing and intelligent interaction. Berlin (Germany): Springer; 2011. p. 125–134.

Hoffman M, Ragusa C. Unwrapping GIFT: a primer on authoring tools for the generalized intelligent framework for tutoring. Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym2); 2015 Jun 12–13; Pittsburgh, PA. Orlando (FL): Army Research Laboratory (US); c2015. p. 11–24.

Jensen R, Presnell B, Cobb MG, Lunsford J, Kidd D. Generalizing automated assessment of small unit tactical decision making. DTIC document; Paper presented at: Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC); 2013 Dec 2–5; Orlando, FL.

Jensen R, Presnell B, Lunsford J, Cobb MG. Automated intelligent training with a tactical decision making serious game. San Mateo (CA): Stottler Henke Associates Inc.; 2014.

Kubovy M. The psychology of perspective and renaissance art. Cambridge (UK): Cambridge University Press; 1986.

Leite I, Henriques R, Martinho C, Paiva A. Sensors in the wild: Exploring electrodermal activity in child-robot interaction. Paper presented at the Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction; 2013 Mar 3–6; Tokyo, Japan. Piscataway (NJ): IEEE Press; 2013. p. 41–48.

Lif P. The use of 2D and 3D displays in military settings [doctoral thesis]. [Uppsala (Sweden)] Uppsala University; 2012.

Luciano CJ, Banerjee PP, Bellotte B, Lemole GM, Jr, Oh M, Charbel FT, Roitberg B. Learning retention of thoracic pedicle screw placement using a high-resolution augmented reality simulator with haptic feedback. Neurosurgery. 2011:69(Suppl OPERATIVE):ons14–ons.

McIntire JP, Havig PR, Geiselman EE. Stereoscopic 3D displays and human performance: a comprehensive review. Displays. 2014;35(1):18–26.

Mercado J, Rupp M, Chen J, Barnes M, Barber D, Procci K. Intelligent agent transparency in human-agent teaming for multi-UxV management. Human factors; The Journal of the Human Factors and Ergonomics Society. 2016;58(3):401–415.

Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Transactions on Information and Systems. 1994;77(12):1321–1329.

Noordzij ML, Scholten P, Laroy-Noordzij ME. Measuring electrodermal activity of both individuals with severe mental disabilities and their caretakers during episodes of challenging behavior. In: Spink AJ, Grieco F, Krips OE, Loijens LWS, Noldus LPJJ, Zimmerman PH, editors. Proceedings of Measuring Behavior 2012: 8th International Conference on Methods and Techniques in Behavioral Research; 2012 Aug 28–31; Utrecht, The Netherlands; Wageningen (Netherlands): Noldus Information Technology; c2012. p. 201–205.

Oskamp M. TerraGuide: A multi-surface environment for visibility analysis [master’s thesis]. [Kingston (Ontario)]: Queen’s University; 2015.

Piaget J. Science of education and the psychology of the child. New York (NY): Viking; 1969.

Poh M-Z, Swenson NC, Picard RW. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. Biomedical Engineering, IEEE Transactions on. 2010;57(5):1243–1252.

Rajan S, Leonard N, Fletcher R, Casarjian B, Casarjian R, Cisse C, Gwadz M. Ambulatory autonomic activity monitoring among at-risk adolescent mothers. Journal of Mobile Technology in Medicine. 2012;1(3):25.

Sano A, Phillips AJ, Yu AZ, McHill AW, Taylor S, Jaques N, Czeisler CA, Klerman EB, Picard RW. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. Proceedings of the 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN); 2015 Jun 9–12; Cambridge, MA. New York (NY): IEEE; c2015. p. 1–6.

Schunk D. Learning Theories, Englewood Cliffs (NJ): Merrill. Prentice Hall; 1996.

Smith J, Flowers P, Larkin M. Interpretative phoneomological analysis: theory, method and research. London (UK): SAGE Publications; 2009.

Sottilare R, Sinatra A, Boyce M, Graesser A. Domain modeling for adaptive training and education in support of the US army learning model-research outline. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2015 Jun. Report No.: ARL-SR-0325.

Sottilare RA, Brawner KW, Goldberg BS, Holden HK. The generalized intelligent framework for tutoring (GIFT). Orlando (FL): Army Research Laboratory (US); 2012. doi: 10.13140/2.1.1629.6003.

St. John M, Cowen MB, Smallman HS, Oonk HM. (2001). The use of 2D and 3D displays for shape-understanding versus relative-position tasks. Human Factors: The Journal of the Human Factors and Ergonomics Society. 2001;43(1):79-98. doi: 10.1518/001872001775992534.

Taylor S, Jaques N, Chen W, Fedor S, Sano A, Picard R. Automatic Identification of Artifacts in Electrodermal Activity Data. Proc. International Conference of the IEEE Engineering in Medicine and Biology Society; 2015 Aug 25–29; Milan, Italy. New York (NY): IEEE; c2015. p. 1934–1937.

US Army Training and Doctrine Command. Basic officer leader training policies and administration. Fort Eustis (VA): Headquarters, US Army Training and Doctrine Command; 2014 Jan. TRADOC Regulation 350-36.

Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. Perceptual and motor skills. 1978;47(2):599–604.

Van Beurden M. 2013: Interaction in depth. Interaction group. [thesis]. [Eindhoven (The Netherlands)]: Eindhoven University of Technology; 2013.

Vygotsky LS. Mind and society: The development of higher mental processes. Cambridge (MA): Harvard University Press; 1978.

Walton J, Dorneich MC, Gilbert S, Bonner D, Winer E, Ray C. (2015). Modality and Timing of Team Feedback: Implications for GIFT. Paper presented at the Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym2).

Ware C. Information visualization: perception for design. Amsterdam (The Netherlands): Elsevier; 2012.

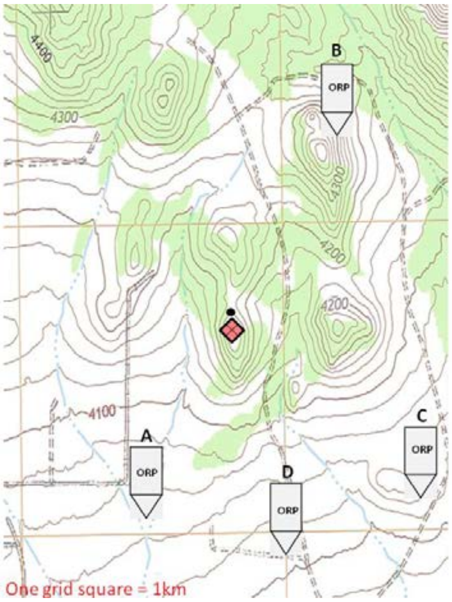
Willett W, Jenny B, Isenberg T, Dragicevic P. Lightweight relief shearing for enhanced terrain perception on interactive maps. Proceedings of the 33rd ACM Conference on Human Factors in Computing Systems (CHI 2015); 2015 Apr 18–23; Seoul, Korea. New York (NY): ACM.

INTENTIONALLY LEFT BLANK.

Appendix A. Tactics Assessment Questions

This appendix appears in its original form, without editorial change.

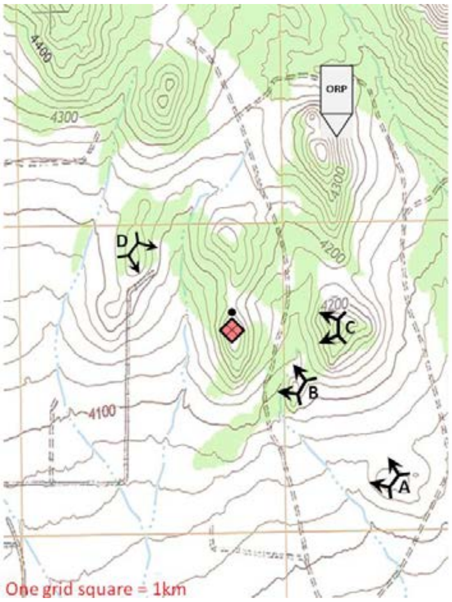
1.



Assumptions: All green areas represent transitable, forest-type vegetation. One grid square = 1km.

As a platoon leader, from an offensive standpoint, given these specific terrain conditions, what is the best objective rally point (ORP)?

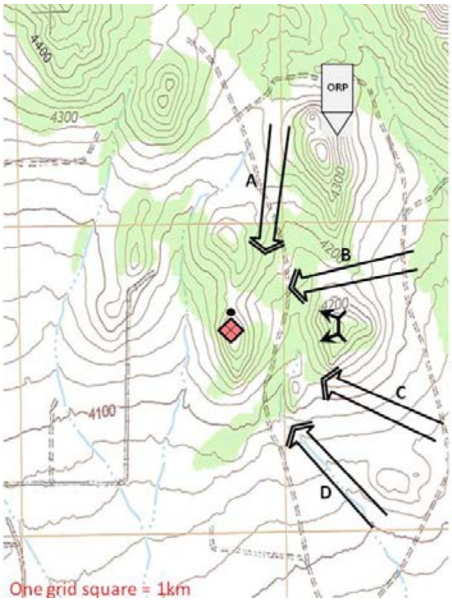
2.



Assumptions: All green areas represent transitable, forest-type vegetation. One grid square = 1km.

From an offensive standpoint, what is the best support by fire position?

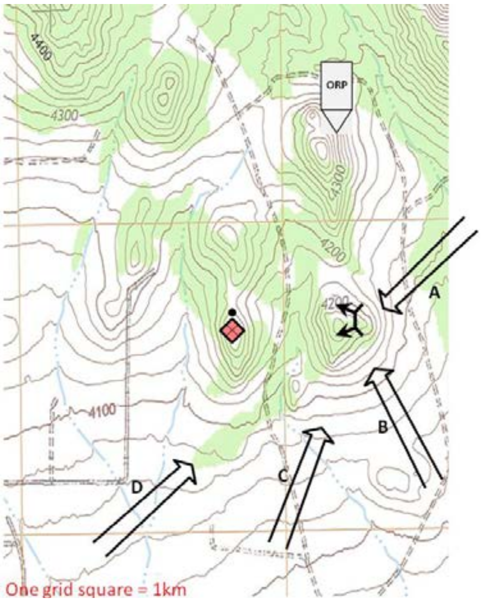
3.



Assumptions: All green areas represent transitable, forest-type vegetation. One grid square = 1km.

What is the best axis of advance for the main attack?

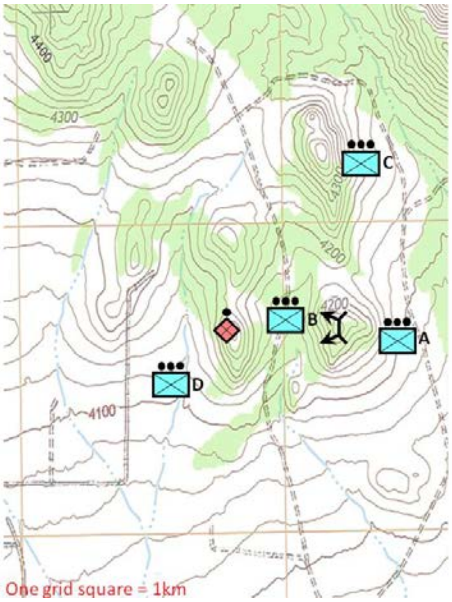
4.



Assumptions: All green areas represent transitable, forest-type vegetation. One grid square = 1km.

What is the best movement route from ORP for the supporting attack?

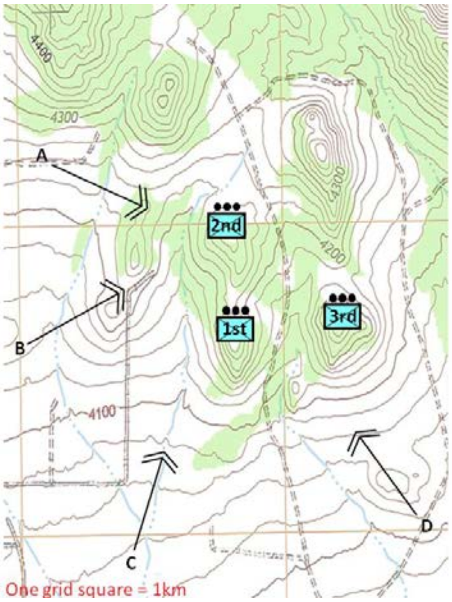
5.



Assumptions: All green areas represent transitable, forest-type vegetation. One grid square = 1km.

After raiding the enemy and according to the terrain given, where would you conduct reorganization to prepare for follow-on attack mission?

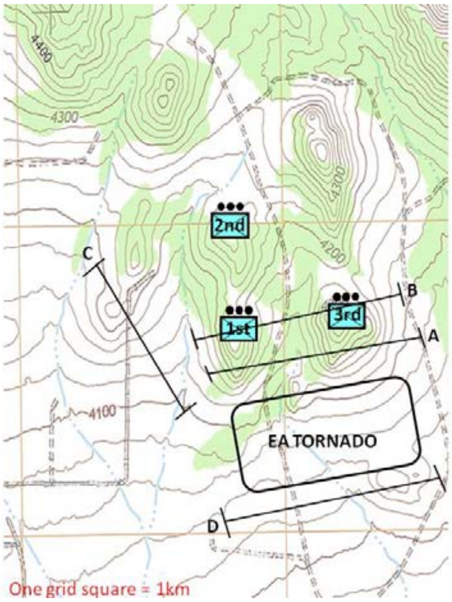
6.



Assumptions: All green areas represent transitable, forest-type vegetation. One grid square = 1km.

As 2nd platoon leader, from a defensive standpoint, given these specific terrain conditions, what is the enemy likely avenue of approach?

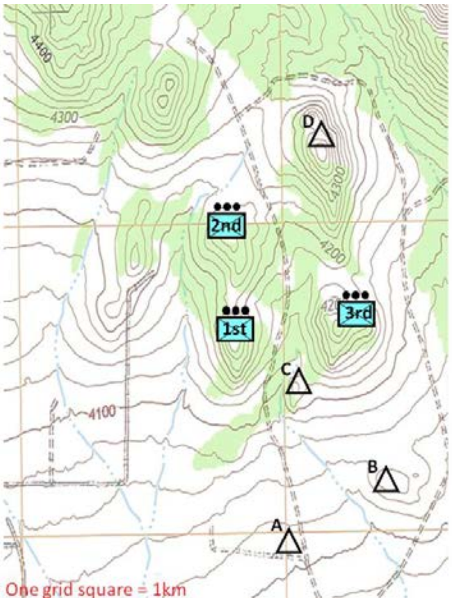
7.



Assumptions: All green areas represent transitable, forest-type vegetation. One grid square = 1km.

As 1st and 3rd platoon leader, from a defensive standpoint, given a primary engagement area, where would you place your coordinated final protective fires?

8.



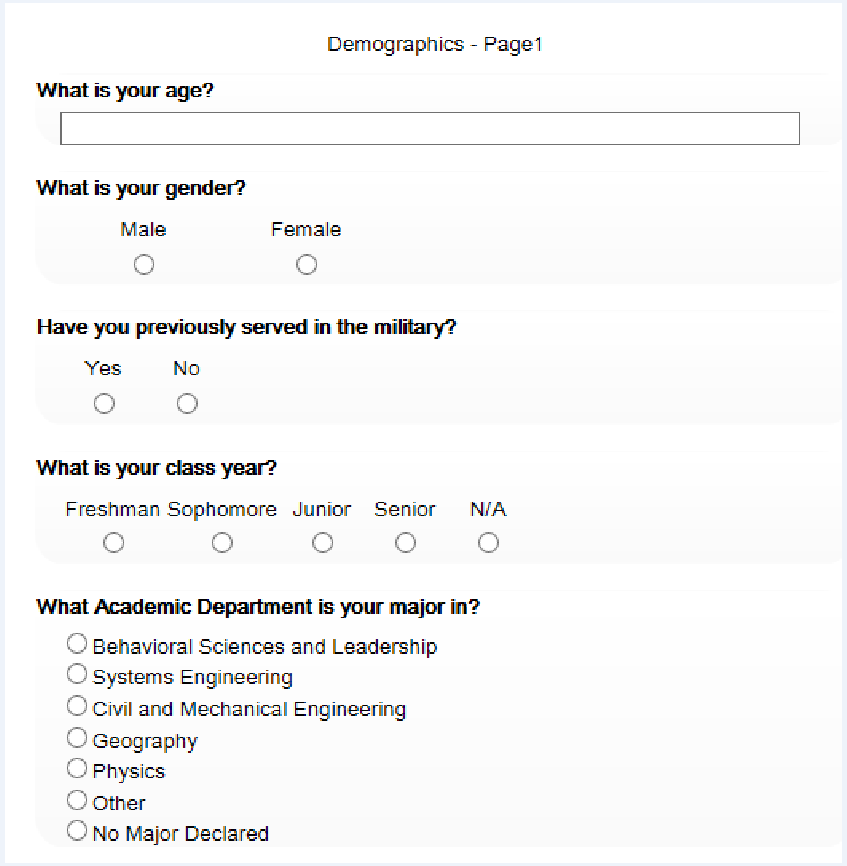
Assumptions: All green areas represent transitable, forest-type vegetation. One grid square = 1km.

As a platoon leader, from a defensive standpoint, given these specific terrain conditions, what is the enemy likely observation post?

INTENTIONALLY LEFT BLANK.

Appendix B. Demographics Questionnaire

This appendix appears in its original form, without editorial change.



Demographics - Page1

What is your age?

☐

What is your gender?

Male ○ Female ○

Have you previously served in the military?

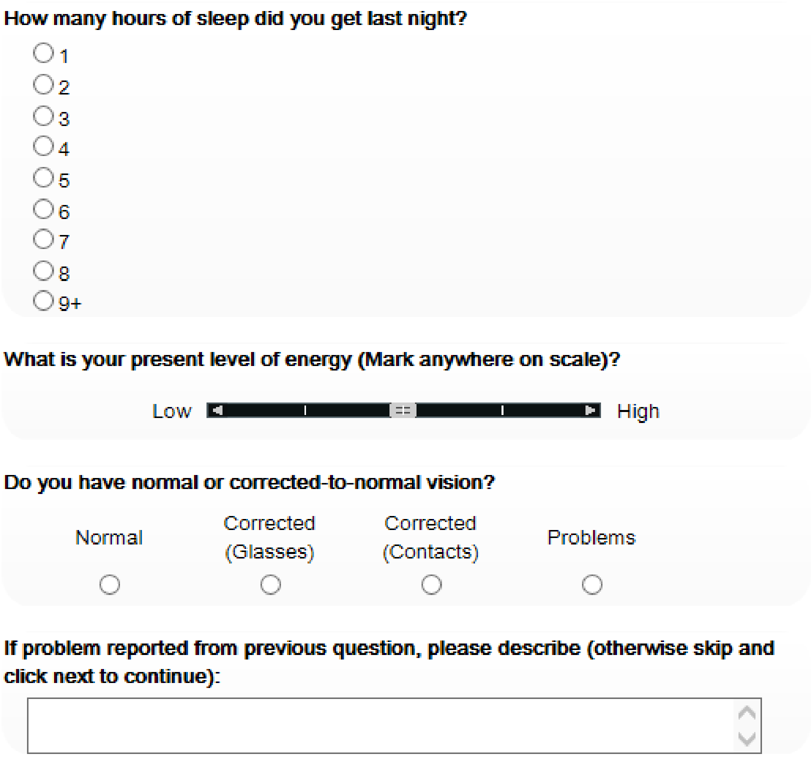
Yes ○ No ○

What is your class year?

Freshman ○ Sophomore ○ Junior ○ Senior ○ N/A ○

What Academic Department is your major in?

- ○ Behavioral Sciences and Leadership
- ○ Systems Engineering
- ○ Civil and Mechanical Engineering
- ○ Geography
- ○ Physics
- ○ Other
- ○ No Major Declared



How many hours of sleep did you get last night?

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8
- ○ 9+

What is your present level of energy (Mark anywhere on scale)?

Low High

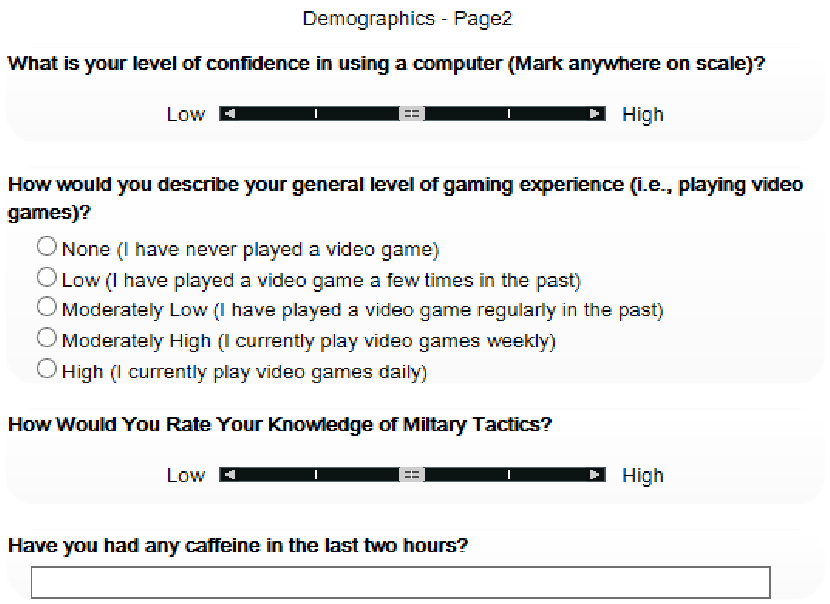
Do you have normal or corrected-to-normal vision?

Normal ○ Corrected (Glasses) ○ Corrected (Contacts) ○ Problems ○

If problem reported from previous question, please describe (otherwise skip and click next to continue):

☐

Approved for public release; distribution is unlimited.



Demographics - Page2

What is your level of confidence in using a computer (Mark anywhere on scale)?

How would you describe your general level of gaming experience (i.e., playing video games)?

○ None (I have never played a video game)
○ Low (I have played a video game a few times in the past)
○ Moderately Low (I have played a video game regularly in the past)
○ Moderately High (I currently play video games weekly)
○ High (I currently play video games daily)

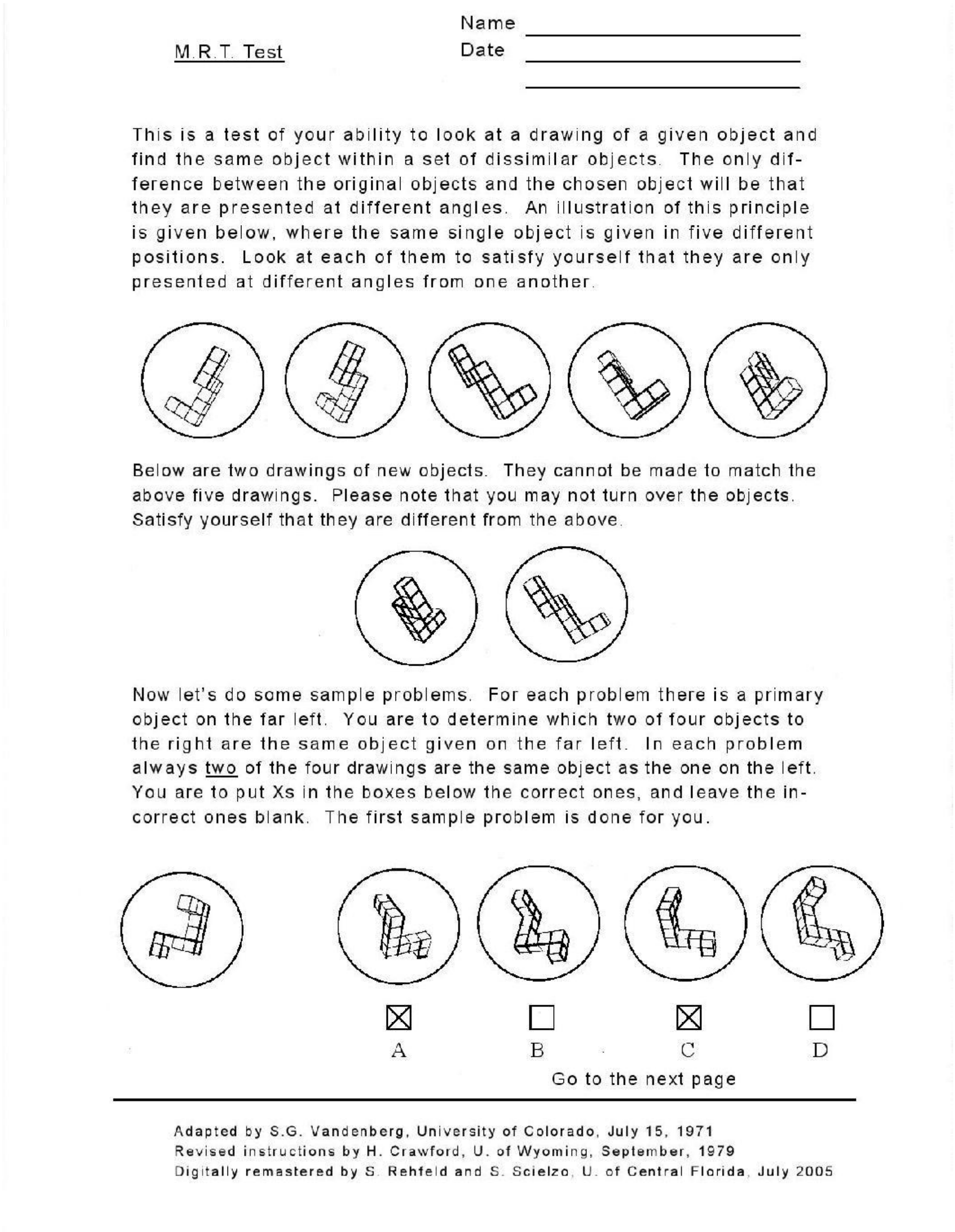
How Would You Rate Your Knowledge of Miltary Tactics?

Have you had any caffeine in the last two hours?

INTENTIONALLY LEFT BLANK.

Appendix C. Mental Rotation Test

This appendix appears in its original form, without editorial change.



M.R.T. Test

Name ______________________

Date ______________________

____

This is a test of your ability to look at a drawing of a given object and find the same object within a set of dissimilar objects. The only difference between the original objects and the chosen object will be that they are presented at different angles. An illustration of this principle is given below, where the same single object is given in five different positions. Look at each of them to satisfy yourself that they are only presented at different angles from one another.

Below are two drawings of new objects. They cannot be made to match the above five drawings. Please note that you may not turn over the objects. Satisfy yourself that they are different from the above.

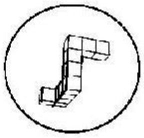
Now let's do some sample problems. For each problem there is a primary object on the far left. You are to determine which two of four objects to the right are the same object given on the far left. In each problem always two of the four drawings are the same object as the one on the left. You are to put Xs in the boxes below the correct ones, and leave the incorrect ones blank. The first sample problem is done for you.

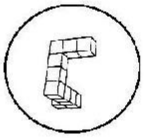
☒ A ☐ B ☒ C ☐ D

Go to the next page

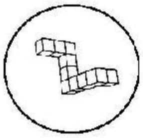
Adapted by S.G. Vandenberg, University of Colorado, July 15, 1971
Revised instructions by H. Crawford, U. of Wyoming, September, 1979
Digitally remastered by S. Rehfeld and S. Scielzo, U. of Central Florida, July 2005

Do the rest of the sample problems yourself. Which two drawings of the four on the right show the same objects as the one on the left? There are always two and only two correct answers for each problem. Put an X under the two correct drawings.

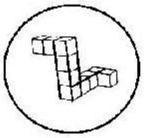




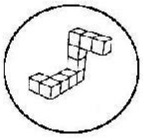
☐ A



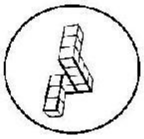
☐ B

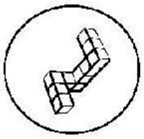


☐ C

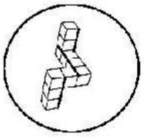


☐ D

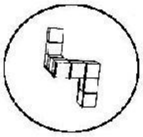




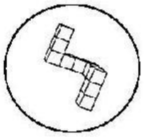
☐ A



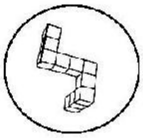
☐ B

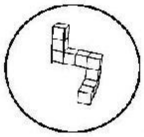


☐ C

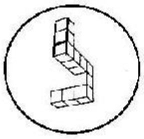


☐ D

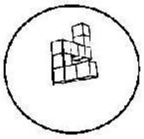




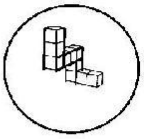
☐ A



☐ B



☐ C



☐ D

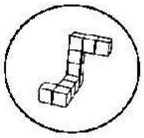
Answer: (1) first and second drawing are correct
(2) first and third drawing are correct
(3) second and third drawing are correct

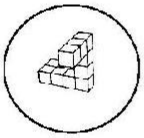
This test has two parts. You will have 3 minutes for each of the two parts. Each part has two pages. When you have finished Part I, STOP. Please do not go on to Part 2 until you are asked to do so. Remember: There are always two and only two correct answers for each item.

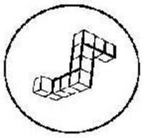
Work as quickly as you can without sacrificing accuracy. Your score on this test will reflect both the correct and incorrect responses. Therefore, it will not be to your advantage to guess unless you have some idea which choice is correct.

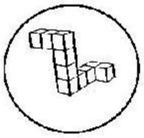
DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO

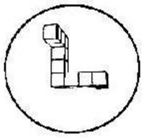
1.



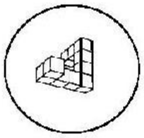

☐ A

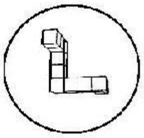

☐ B

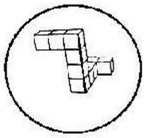

☐ C

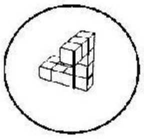

☐ D

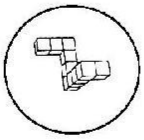
2.



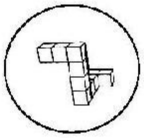

☐ A

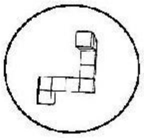

☐ B

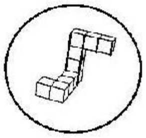

☐ C

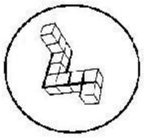

☐ D

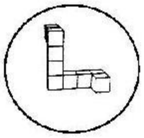
3.



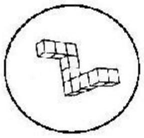

☐ A

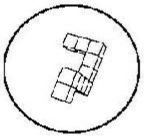

☐ B

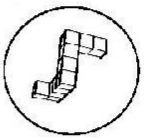

☐ C

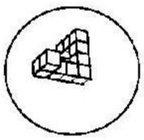

☐ D

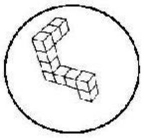
4.



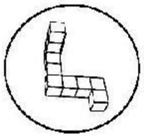

☐ A

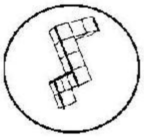

☐ B

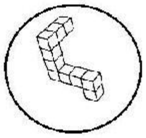

☐ C

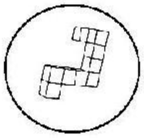

☐ D

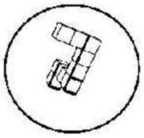
5.




☐ A

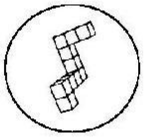

☐ B

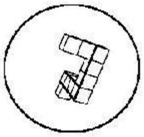

☐ C


☐ D

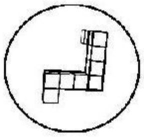
GO ON TO THE NEXT PAGE

6.

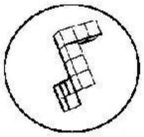




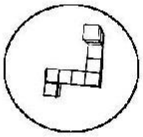
☐ A



☐ B

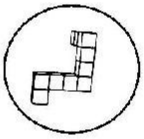


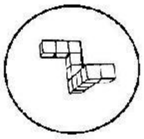
☐ C



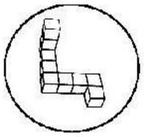
☐ D

7.

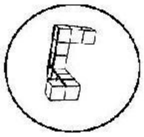




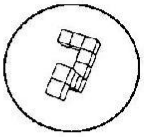
☐ A



☐ B

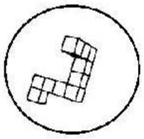


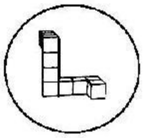
☐ C



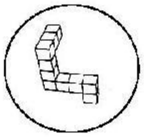
☐ D

8.

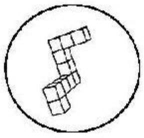




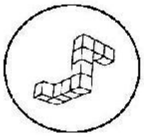
☐ A



☐ B

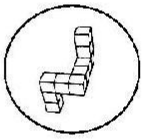


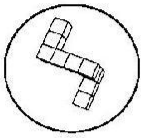
☐ C



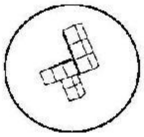
☐ D

9.

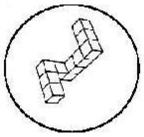




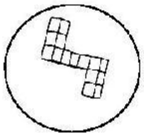
☐ A



☐ B

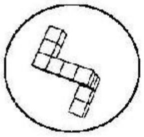


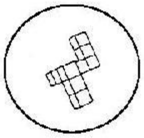
☐ C



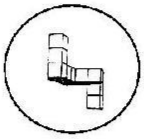
☐ D

10.

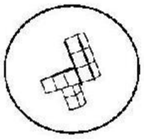




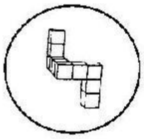
☐ A



☐ B



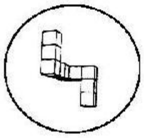
☐ C

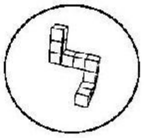


☐ D

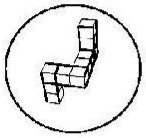
DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO STOP

11.

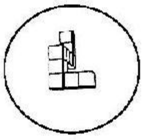




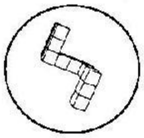
☐ A



☐ B

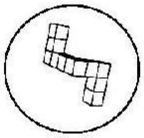


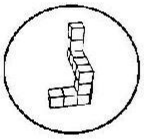
☐ C



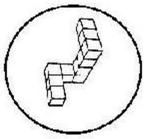
☐ D

12.

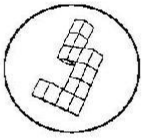




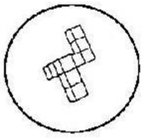
☐ A



☐ B

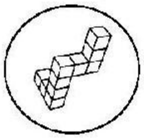


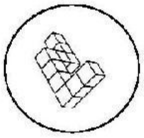
☐ C



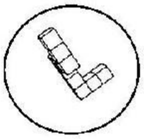
☐ D

13.

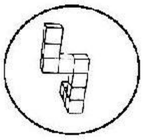




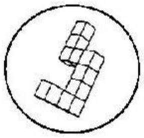
☐ A



☐ B

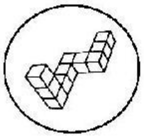


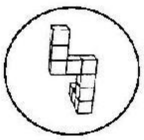
☐ C



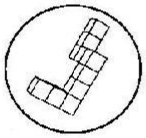
☐ D

14.

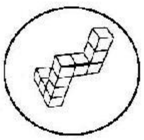




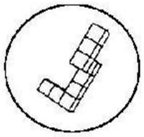
☐ A



☐ B

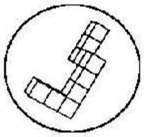


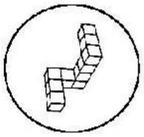
☐ C



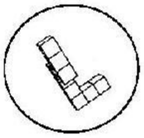
☐ D

15.

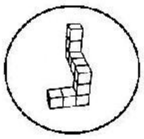




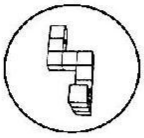
☐ A



☐ B



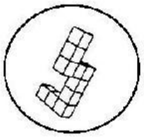
☐ C

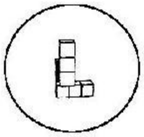


☐ D

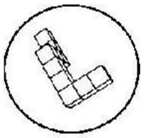
GO ON TO THE NEXT PAGE

16.

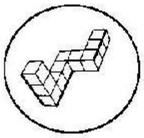




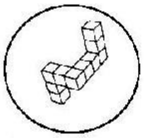
☐ A



☐ B

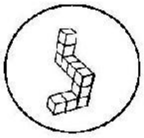


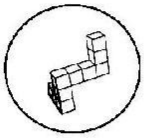
☐ C



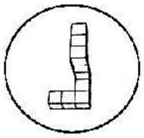
☐ D

17.

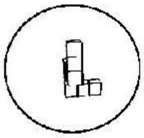




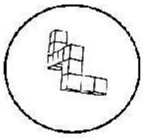
☐ A



☐ B

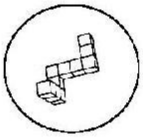


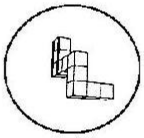
☐ C



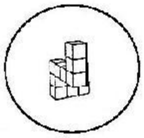
☐ D

18.

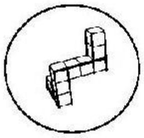




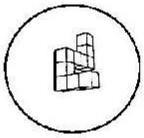
☐ A



☐ B

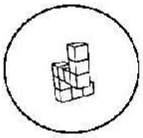


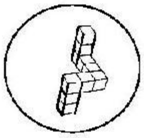
☐ C



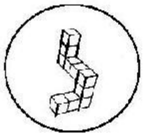
☐ D

19.

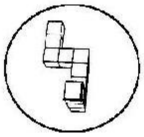




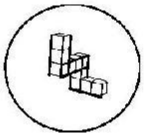
☐ A



☐ B

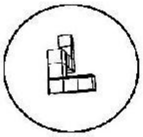


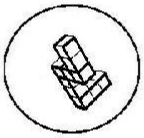
☐ C



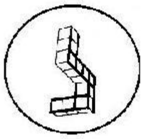
☐ D

20.

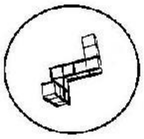




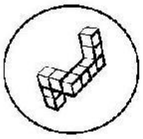
☐ A



☐ B



☐ C



☐ D

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO

STOP

INTENTIONALLY LEFT BLANK.

Appendix D. NASA Task Load Index (NASA-TLX)

This appendix appears in its original form, without editorial change.

NASA-TLX Questionnaire

Please rate your overall impression of demands imposed on you during the exercise.

1. Mental Demand: How much mental and perceptual activity was required (e.g., thinking, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

LOW |---|---|---|---|---|---|---|---|---|---| HIGH
1 2 3 4 5 6 7 8 9 10

2. Physical Demand: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

LOW |---|---|---|---|---|---|---|---|---|---| HIGH
1 2 3 4 5 6 7 8 9 10

3. Temporal Demand: How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

LOW |---|---|---|---|---|---|---|---|---|---| HIGH
1 2 3 4 5 6 7 8 9 10

4. Level of Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

LOW |---|---|---|---|---|---|---|---|---|---| HIGH
1 2 3 4 5 6 7 8 9 10

5. Level of Frustration: How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

LOW |---|---|---|---|---|---|---|---|---|---| HIGH
1 2 3 4 5 6 7 8 9 10

6. Performance: How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

LOW |---|---|---|---|---|---|---|---|---|---| HIGH
1 2 3 4 5 6 7 8 9 10

Pair-wise Comparison of Factors

Select the member of each pair that provided the most significant source of workload variation in these tasks.

Physical Demand vs. Mental Demand

Temporal Demand vs. Mental Demand

Performance vs. Mental Demand

Frustration vs. Mental Demand

Effort vs. Mental Demand

Temporal Demand vs. Physical Demand

Performance vs. Physical Demand

Frustration vs. Physical Demand

Effort vs. Physical Demand

Temporal Demand vs. Performance

Temporal Demand vs. Frustration

Temporal Demand vs. Effort

Performance vs. Frustration

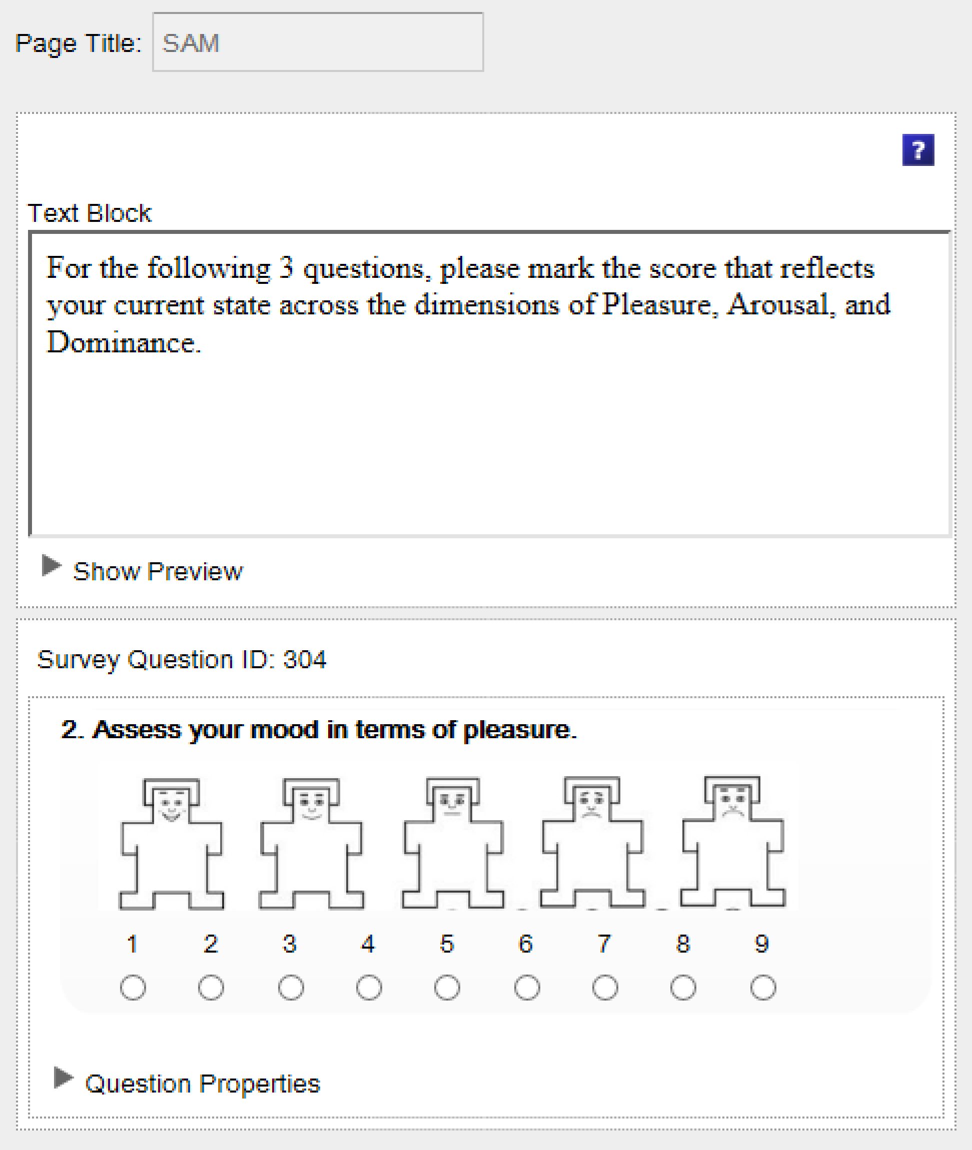
Performance vs. Effort

Effort vs. Frustration

INTENTIONALLY LEFT BLANK.

Appendix E. Self-Assessment Manikin Test

This appendix appears in its original form, without editorial change.



Page Title: SAM

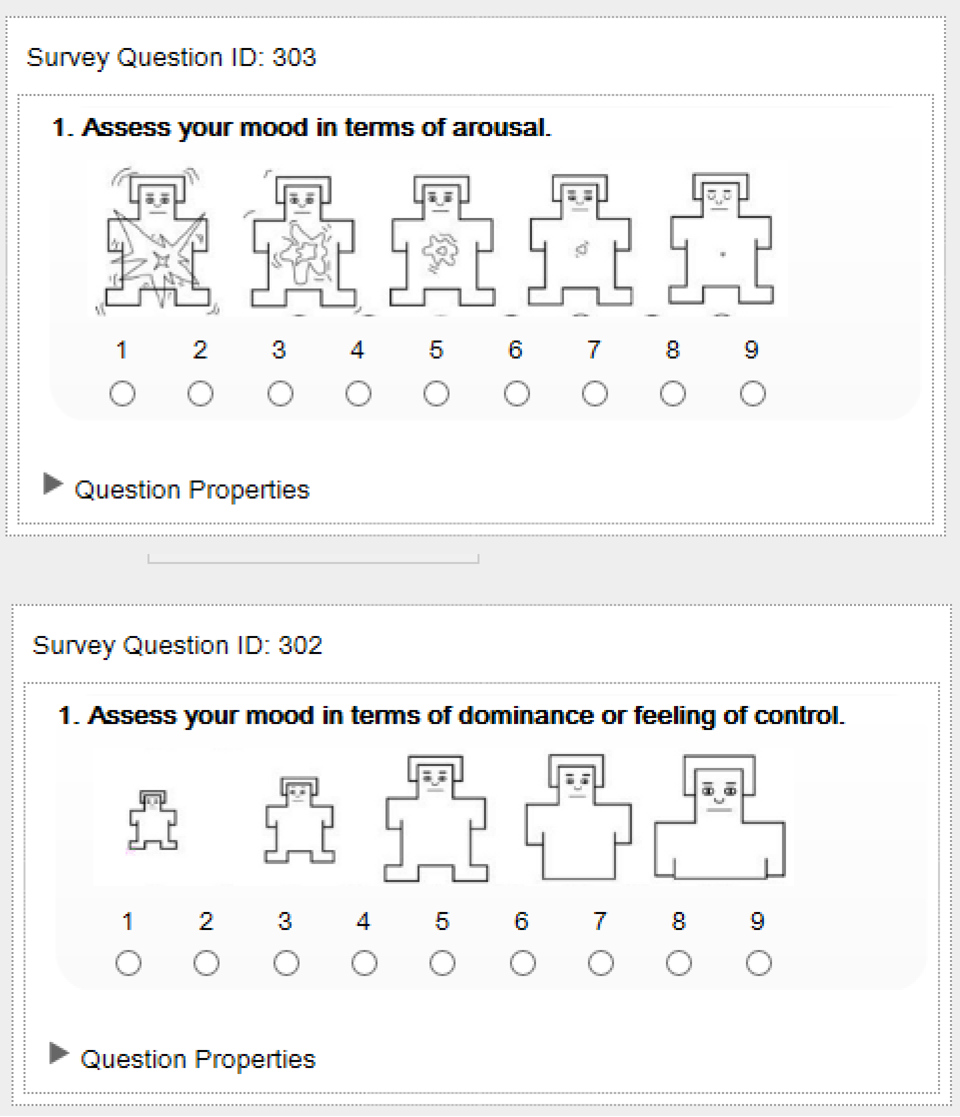
Text Block

For the following 3 questions, please mark the score that reflects your current state across the dimensions of Pleasure, Arousal, and Dominance.

Survey Question ID: 304

2. Assess your mood in terms of pleasure.

1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○



Survey Question ID: 303

1. Assess your mood in terms of arousal.

1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○

Survey Question ID: 302

1. Assess your mood in terms of dominance or feeling of control.

1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○

INTENTIONALLY LEFT BLANK.

List of Symbols, Abbreviations, and Acronyms

2-D	2-dimensional
3-D	3-dimensional
Ag/AgCl	silver/silver chloride
AR	augmented reality
ARES	Augmented REality Sandtable
ARL	US Army Research Laboratory
COTS	commercial off-the-shelf
EDA	electrodermal activity
GIFT	Generalized Intelligent Framework for Tutoring
IPA	Interpretative Phenomenological Analysis
ITS	Intelligent Tutoring System
MS	military science
NASA-TLX	National Aeronautics and Space Administration Task Load Index
NPS	net promoter score
ORP	objective rally point
ROTC	Reserve Officer Training Corps
SAM	Self-Assessment Manikin
SCR	skin conductance response
UCF	University of Central Florida
USMA	United States Military Academy

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
2 (PDF)	DIRECTOR US ARMY RESEARCH LAB RDRL CIO L IMAL HRA MAIL & RECORDS MGMT
1 (PDF)	GOVT PRINTG OFC A MALHOTRA
1 (PDF)	ARL/ORAUDA J KIM
1 (PDF)	UNIV OF CENTRAL FLORIDA T GORIS
13 (PDF)	DIR USARL RDRL HR M BOYCE K AMAYA R SOTTILARE S OSOSKY RDRL HRA A C AMBURN C METEVIER R LONG RDRL HRA B B GOLDBERG G GOODWIN J MOSS J JOHNSTON K BRAWNER J HART