The Impacts of Robots and Leadership on Cadets' Workload in Military Training

David T. Feltner

Department of Psychology, North Carolina State University

PSY 880: Quantitative Methods in Psychology

Dr. Shevaun Neupert

December 6, 2024

Author Note

Data collection and preliminary analysis were sponsored by the United States Military Academy and North Carolina State University. There are no conflicts to disclose at this time.

Correspondence concerning this article should be addressed to David T. Feltner, LACE Lab in the Department of Psychology, North Carolina State University, Poe Hall, 2310 Stinson Drive, Raleigh, NC 27607. Email: dtfeltne@ncsu.edu.

Abstract

This study, conducted at the United States Military Academy, explores the impact of incorporating robotic teammates, specifically the semi-autonomous robotic dog Spot, into military training scenarios. Placing emphasis on Platoon-level operations and leadership responsibilities, the research utilizes Hart & Staveland's NASA-TLX to measure subjective workload perceptions among Cadets. Results did not indicate that leaders experience greater cognitive demands when integrating robotic assistance, likely due to the confounding factors surrounding Cadets unique field training environment. The study underscores the need for careful consideration in training protocols to enhance mission effectiveness while minimizing cognitive burdens on military leaders. The findings contribute to understanding the complex dynamics of human-robot interaction in military settings, informing future technology integration.

Keywords: workload, robotic teammate, military robot, military training, United States Military Academy, West Point

The Impacts of Robots and Leadership on Cadets' Workload in Military Training

Robots & Leaders on the Battlefield

In the 1990s, ground robots became a key component of high-tech military arsenals, ranging from the U.S. and Europe to Asia and the Middle East. Since then, robots have continued to play key roles for US military forces deployed in many countries throughout the Global War on Terror. As the Army prepares for Large Scale Combat Operations against a near-peer threat (i.e., China or Russia), it continues to modernize its ground combat forces by fielding robots to Soldiers. These robots, or robotic teammates, are incredibly powerful, as they have the ability to increase both Soldier capabilities and individual Soldier survivability. For the leaders on the front lines, that means that the incorporation of robots is a foregone conclusion – robots are here to stay. Now, leaders need to onboard them and train with them in preparation for battle.

Robots are an incredible tool, but they levy a cognitive tax against the leader that directs the robots' actions. The Army has fielded semi-autonomous robots that are capable of receiving tactical tasks and moving across the battlefield. These robots are enablers that bring unique capabilities, but they cannot execute tasks on their own. Often, these robots are novel, and leaders have to take extra effort to best incorporate them into the formation. Leaders at echelon have to understand the capabilities, assess their usage on the battlefield, and continually monitor the robot during the mission on top off all their other inherent requirements.

While all people have limited mental resources, leaders on a kinetic battlefield are taxed in unique ways. At all levels, these leaders are immersed in a dynamic and lethal environment; simultaneously, often tired and mentally degraded, they are balancing the lives of their subordinates with the success of the mission. As people rise within the Army, they are responsible for larger organizations, more technological capabilities, and more complex problems to solve. Compared to individual Soldiers, who fight and are only responsible for executing their tasks, leaders have to plan, resource, and execute their own tasks. Squad Leaders

are in charge of 9-12 Soldiers, planning specific parts of missions, and rehearsing before the mission starts. Platoon Leaders are in charge of all 40 Soldiers, responsible for all four squads, coordinating with higher headquarters, and planning all tactical operations.

Training Tactical Tasks

Soldiers prepare for combat operations by executing offensive and defensive situational training exercises (STX) in peacetime. Offensive operations, such as a raid or a movement to contact, are incredibly difficult – they require planning, moving under load across varied terrain, and often end with friendly forces coming under direct enemy contact. Additionally, if conditions change, leaders have to re-orient themselves on the situation, make decisions, and re-task subordinates while on the move. Offensive operations are split into two main categories – deliberate and hasty. A deliberate attack is planned ahead of time by leadership, rehearsed by the leaders and Soldiers, and then executed by the formation. A raid is the most complex operation, requiring significant planning and decision making by Platoon level leaders. Whereas a hasty attack, has no deliberate plan and only requires the reaction to any unplanned enemy contact.

Mental Resources

Humans have limited mental resources, and because of this fact, mental workload can be defined as the difference between the amount of available mental processing resources and cognitive task demands (Hart & Staveland, 1988). For instance, a routine task may require only 10% of a person's available resources; whereas, a very difficult task might require 90% of the same person's mental resources. Mental overload occurs when there are too few resources available to allocate to required tasks, increasing stress and errors. Hart and Staveland's (1988) standard NASA-TLX definitions were used in the experiment to measure workload with the following dimensions: physical, mental, effort, and frustration. In line with Hart's (2006) later investigation into the NASA-TLX, the six dimensions of the TLX can be used without the initial pairwise comparisons while maintaining the scale's same reliability and sensitivity.

Purpose of the Present Study

In the present study, the researchers tried to determine the impact of providing novel robotic teammates into military training. The novice military members – Cadets at USMA – would simultaneously navigate into leadership positions and difficult training scenarios. In the midst of this complex environment, some Cadets would work through the option of how to use a new and impressive semi-autonomous capability. While conducting this field research, we sought to answer a singular but multi-faceted question: how does the incorporation of a robot impact Cadets as they execute combat operations?

Hypotheses

Based on the novelty of using a robot during a field training exercise, it was expected that leaders using a robotic teammate during the attack lane would impose a greater cognitive workload on users. Alternatively, the non-robot condition was expected to generate the lowest cognitive workloads on both lanes. These general research hypotheses are translated here in terms of specific DVs and IVs recorded and manipulated during the experiment:

- Hypothesis 1: The robot condition will generate a higher cognitive workload than the non-robot condition.
- Hypothesis 2: Leaders (Squad Leaders and Platoon Leaders) will experience a higher cognitive workload than non-leaders (Members of Squad).
- Hypothesis 3: Cadets, regardless of position, with a robotic teammate in the deliberate attack lane will generate the highest workload.
- Hypothesis 4: Leaders with a robotic teammate will have the highest workload.

Method

Experiment Overview

In line with required military training from the United States Military Academy (USMA) at West Point, Cadets organize into standard military formations for summer training. Within

these formations, Cadets are assigned leadership positions to facilitate collective training. After being aligned within squads, platoons, and companies, the Cadets began preparations to complete seven situational training exercise (STX) lanes across 14 days, meaning that every Cadet would complete every lane. Within the first two Cadet Companies to complete STX training, half of the Platoons were randomly assigned to the Robot Condition, while the other half of Platoons were assigned to the Non-Robot Condition. The researchers monitored leadership positions but were not able to assign Cadets to specific leadership positions.

Procedure

During the STX preparation, every Cadet platoon received the following: a block of instruction on the capabilities of a robotic dog named Spot, a demonstration of Spot's capabilities, and introductory surveys.

In line with their required military training, every Cadet completed seven different STX lanes, but researchers only walked with Cadets on Lane 1 (Hasty Attack lane) and Lane 5 (Raid lane). Half of the Platoons going through these lanes were assigned to the Robot Condition - completing the attack or defense lane with Spot, the robotic dog. Whereas, the other half were assigned to the Non-Robot Condition, without the aid of Spot. As the Cadets started their 12-hour hasty planning cycle before starting the lane, Cadets planned their next day's operation and were able to use (or not use) Spot within their plan. During the conduct of the 6-hour lane, Cadet leaders could modify pre-existing tasks or give new tasks to subordinates as they executed their mission. At the conclusion of Lane 1 and Lane 5, all cadets completed the NASA-TLX to gauge their level of workload during that lane.

Based on the austere conditions of the military training, packets with the appropriate questionnaires were pre-printed and filled out with pen and paper by the Cadets.

Participants

Ninety-four Cadets from West Point, 81 male and 13 female, were recruited for the study during Cadet military summer training. All participants were Juniors or Seniors at USMA with an average age of 20.91 years (range 19 – 25, *SD*: 2.08).

Independent Variables

Hasty Attack and Raid Lanes

Every Cadet completed STX Lanes 1-7 sequentially with the same planning and execution timelines. Lane 1 was a hasty attack lane where the dismounted Cadet platoon moved four kilometers through severely restrictive terrain against an unknown enemy. During the movement through the woods, the enemy force of 3-5 personnel would attack the Cadet formation at an advantageous moment. Conversely, Lane 5 was a deliberate attack lane where the Cadets conducted a raid on a complex compound. Their raid consisted of a 3-kilometer movement, deliberate reconnaissance of the objective, an attack on an enemy stronghold, and a planned withdrawal off the stronghold. See Appendix A for a full description of the Cadets' training as well as a more detailed Concept of Operation for both Lane 1 and Lane 5.

Robot and Non-Robot Conditions

Cadet formations were randomly assigned a robotic dog, Spot, seen below in Figure 1, to facilitate operations on their lane. Spot was able to move across the wooded terrain, respond to naturalistic communication, and did not require any additional assets from the Cadets. During the Platoon's planning cycle, Spot was available for any rehearsals at the Squad or Platoon level. Although Spot was available for use, Cadet leadership was not required to use the robot. Cadet Platoons without Spot completed the STX lane without any additional aids or external enablers.

Figure 1

Spot the Robot Conducting Tactical Operations

Note. Researchers controlled Spot to simulate the semi-autonomous capabilities.

Leadership Level

Every STX mission was a Platoon-level operation, with a Platoon Leader and four Squad Leaders in the formation set by the USMA cadre. Based on USMA requirements, there was one set of leaders responsible for planning and a different set responsible for the execution of the operation. The roles were not randomly assigned by the researchers but were instead assigned by the USMA cadre. When Cadets assumed the positions of Squad Leader or Platoon Leader, they were assigned as Leaders, whereas when not in these positions they were rated as Members of Squad. These Cadets assigned as Members of Squad still conducted the operation but were only responsible for their individual tasks.

Dependent Variable: NASA-TLX

At the conclusion of each lane, every Cadet completed Hart & Staveland's (1988) NASA-TLX to provide a subjective measure of determining how much workload that person "felt." Subjective workload measurements can provide insight into how people perceive demand but must occur after completing a task or subtask. Although the measure required substantial

time to complete, it has been shown to be highly accurate (Miller, 2001; Hill et al., 1992). The NASA-TLX for the Cadets was completed at the conclusion of STX Lane 1 and Lane 5, as shown in Appendix B. To calculate the overall workload felt by the Cadet, the individual dimension scale weights were averaged to create a final score.

Prior to the application of an independent T-test or ANOVA, a reliability analysis on the NASA-TLX's individual dimensions was conducted. An initial reliability revealed a Cronbach's Alpha of .29; however, upon further investigation, it appeared that Cadets failed to understand the reverse wording of the NASA-TLX's performance dimension. When Performance was removed from the analysis, Cronbach's Alpha revealed an updated reliability of .787, which aligns with the NASA-TLX's historical reliability (Hart, 2006; Braarud, 2021).

Results

At the conclusion of each STX lane, Cadets were given instructions to rate their subjective workload based off the dimensions of the NASA-TLX. Consequently, the aggregated workload for the Cadets is broken down for each independent variable in Table 1 below.

Table 1

Workload Descriptive Statistics

	N	Workload	
		Mean	SD
Robot Condition			
Robot Condition	65	5.94	1.66
Non-Robot Condition	29	6.62	0.89
Total	94	6.15	1.50
STX Lane			
Hasty Attack (Lane 1)	57	6.26	1.37
Raid (Lane 5)	37	5.97	1.68
Total	94	6.15	1.50
Leadership			
Leaders	49	6.12	1.50
Members of Squad	45	6.18	1.51
Total	94	5.94	1.50

Robot Condition

Levene's test for equality of variances was performed and the data was found to violate the assumption of equal variance. An independent t-test revealed that cadets in the Non-Robot Condition (M = 6.62, SD = 0.89) had a significantly higher workload than cadets in the Robot Condition (M = 5.94, SD = 1.66), $t(88.81) = -2.58$, $p = .012$, $\eta^2 = .05$ which is a small effect as the robot condition accounts for 5.0% of the variance in the cadets' workload.

Contrary to Hypothesis 1, the data demonstrated that cadets in the Non-Robot condition had a significantly higher workload than cadets executing STX lanes in the Robot condition.

Leadership

Levene's test for equality of variances was performed to check the validity of equal variance. An independent t-test revealed that there was no significant difference between cadets in leadership positions (M = 6.11, SD = 1.50) and cadets serving as members of squad (M = 6.18, SD = 1.51), $t(92) = 0.20$, $p = .844$, $\eta^2 = .01$ which is a negligible effect on cadets' workload.

Contrary to Hypothesis 2, the data demonstrated that there was no significant difference in cadets' workload if the cadets were in leadership positions or were members of squad.

Robot Condition Conducting a Raid

Levene's test showed that the variances of the groups were not equal, $F(2, 91) = 6.68$, $p = .002$. A two-way ANOVA was conducted to compare the main effects of the Robot Condition and STX Lane as well as their interaction effects on workload; the analysis revealed no significant main or interactive effects on workload. The main effect of the robot condition yielded an effect size of .04, indicating that 4% of the variance in the workload was explained by if they had a robot $F(1, 91) = 3.48$, $p = .065$, $\eta^2 = .04$. The main effect of the STX lane yielded an effect size of .001, indicating that a mere .1% of the variance in the workload was explained by whether they conducted a hasty attack or a raid $F(1, 91) = 0.06$, $p = .815$, $\eta^2 = .001$. The

interaction effect was not significant, indicating there was no combined effect for Robot Condition and STX Lane on overall reported workload.

Contrary to Hypothesis 3, Cadets in the Non-Robot Condition conducting the hasty attack lane felt the greatest workload as seen below in Table 2.

Table 2

Cadets' Workload in STX Lanes for both the Robot and Non-Robot Conditions

	N	Workload	
		Mean	SD
Robot Condition			
Hasty Attack (Lane 1)	28	5.89	1.67
Raid (Lane 5)	37	5.97	1.68
Non-Robot Condition			
Hasty Attack (Lane 1)	29	6.62	0.89
Raid (Lane 5)	0		

Leaders in the Robot Condition

Levene's test showed that the variances of the group were not equal $F(3, 90) = 4.79$, $p = .004$. A two-way ANOVA was conducted to compare the main effects of the Robot Condition and Leadership position as well as their interaction effects on workload; the analysis revealed that only the main effect of the Robot Condition had a significant effect on the cadets' workload. The main effect of the robot revealed an effect size of .05, indicating that 5% of the variance in the workload was explained by if they had a robot $F(1, 90) = 4.20$, $p = .043$, $\eta^2 = .05$. The main effect of the leadership position revealed an effect size of less than .001, indicating that a negligible amount of the workload's variance is explained by the cadets' leadership position $F(1, 90) = 0.04$, $p = ..84$, $\eta^2 < .001$. The interaction effect was not significant ($F(1,90) = 0.60$, $p = .441$), indicating there was no combined effect of Robot Condition and Leadership position on overall reported workload.

Contrary to Hypothesis 4, Leaders in the Robot condition reported the lowest overall workload among all conditions as seen below in Table 3.

Table 3

Leaders' Workload in the Robot and Non-Robot Conditions

	N	Workload	
		Mean	SD
Robot Condition			
Leadership Position	35	5.85	1.62
Member of Squad	30	6.04	1.73
Non-Robot Condition			
Leadership Position	14	6.79	0.86
Member of Squad	15	6.46	0.92

Discussion

Conclusion

The present study investigated the impact of incorporating robotic teammates, specifically the semi-autonomous robotic dog Spot, into military training scenarios conducted by Cadets at the United States Military Academy (USMA). The experiment was designed to assess how the inclusion of robotic assistance during situational training exercises (STX) affected the workload perceived by Cadets and to what extent leadership levels influenced this perception.

Across the field training, Cadets did not experience a higher subjective workload when they had Spot within their Platoon. In fact, contrary to Hypothesis 1, Cadets in the Non-Robot Condition had a significantly higher workload than their counterparts. Additionally, there was no significant workload difference between leaders and non-leaders on the STX lanes. Lastly, the complexity of the offensive operation had no impact on the imposed workload either.

All of these results were contrary to the anticipated results based off the literature and the experience of other military experts. This may have been an unintended consequence of two things: the unique environment of Cadets conducting mandatory training, and the Cadets' lack of

experience. As Cadets are concerned about receiving a (graduation required) satisfactory grade on their leadership position more than incorporating a novel robot into the tactical plan. Additionally, as this this is the first time that Cadets are maneuvering a Platoon sized element against a thinking enemy force, most of their attentional resources would likely be focused on this new and complex task. The facts may have impacted their subjective workloads in the STX Lanes.

Application

The Army will continue to balance implementing new technology, incorporating robots to augment capabilities, and the limitations of leaders' capacity on an ambiguous battlefield. As the military modernizes and attempts to keep up with our adversaries, the technology will continue to morph and change; however, leaders will always have to be at the point of friction making the best decisions that they can. With that in mind, incorporating robots into high functioning military teams with their incredible capabilities will continue to be paramount.

Limitations

During the data analysis, the performance dimension of the NASA-TLX was removed from the omnibus workload score to increase the scale's reliability. As seen in Appendix B, Performance is the sole dimension of the NASA-TLX that needs to be reverse coded; that is, a mark on the right side of the scale for Performance indicates better performance whereas a mark on the right side of all the other scales indicates increased difficulty. It seems as though the Cadets did not fully read the scale's instructions, which significantly impacted the initial Cronbach's alpha. Therefore, the researchers chose to analyze the NASA-TLX without the performance dimension.

Another limitation of the research is the lack of an objective performance measure of how the cadets navigated the STX lane. By having an objective performance score, the researchers

could investigate how effectively Spot was used by the Cadets conducting the offensive operations.

Future Research

Future research could explore additional factors influencing the interaction between more experienced leaders and robotic teammates. This may include investigating the learning curve associated with using robotic technologies, assessing the long-term impact on leadership development, and exploring potential adaptations in training protocols based on feedback of more experienced tacticians.

In conclusion, this study contributes valuable insights into the complex dynamics of human-robot interaction within military training contexts. By examining the perceived workload of Cadets and the influence of leadership levels, the research provides a foundation for refining training protocols and advancing the effective integration of robotic teammates into military operations.

References

Braarud, P. Ø. (2021). Investigating the validity of subjective workload rating (NASA TLX) and subjective situation awareness rating (SART) for cognitively complex human–machine work. *International Journal of Industrial Ergonomics*, *86*, 103233.

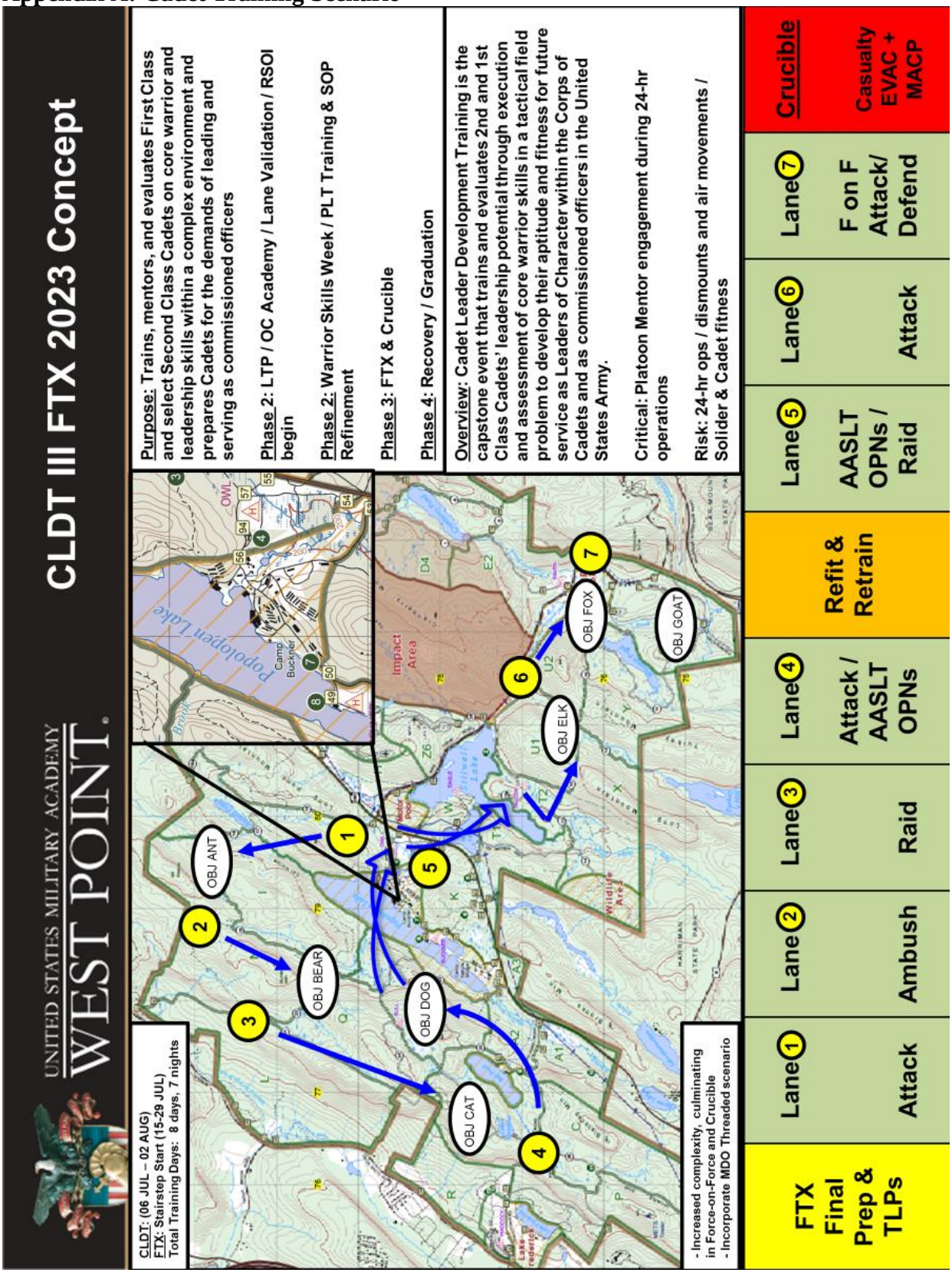
Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904-908. https://doi.org/10.1177/154193120605000909

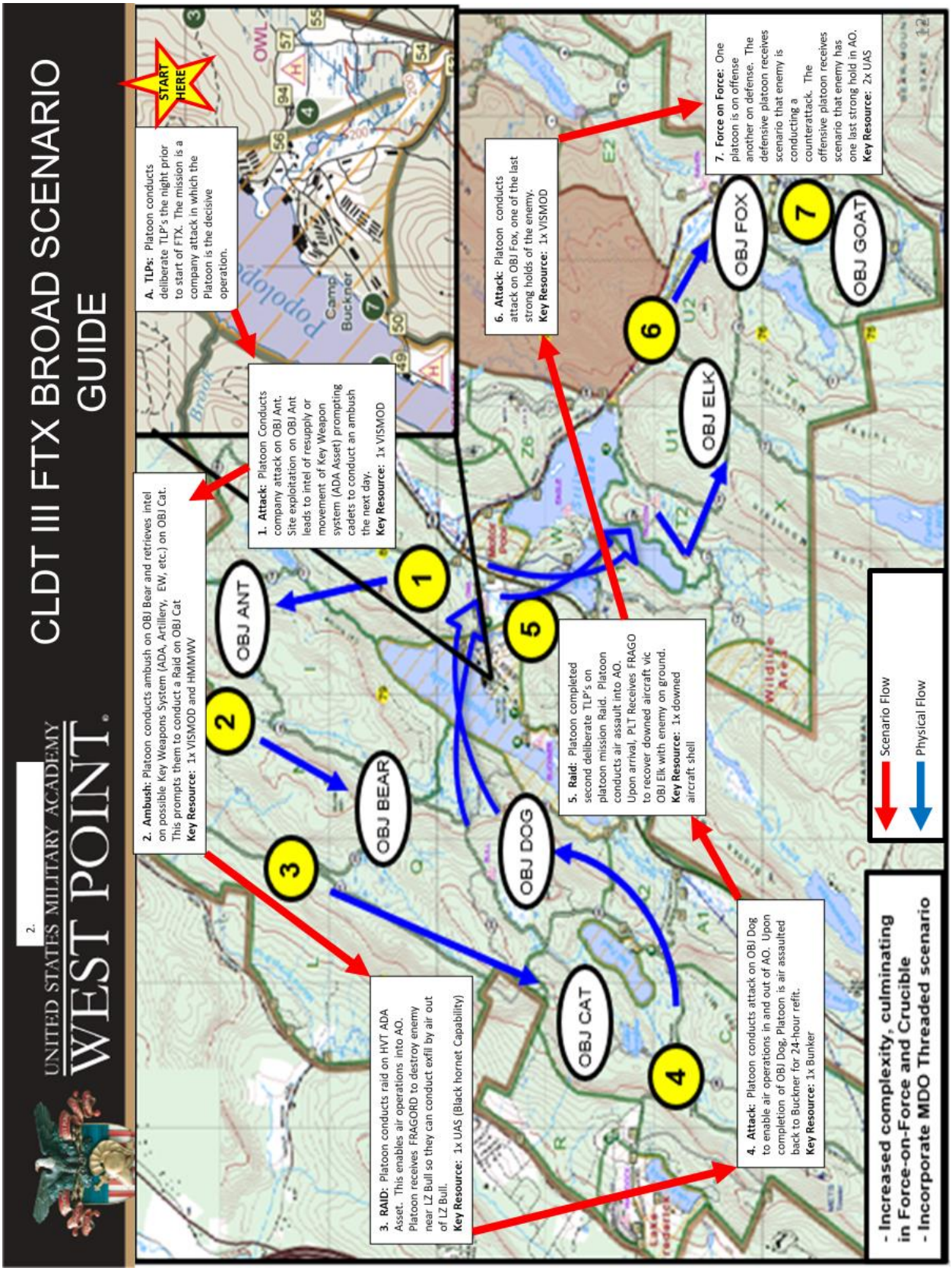
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) *Human Mental Workload*. Amsterdam: North Holland Press.

Hill, S. G., Lavecchia, H. P., Byers, J. C., Bittner, A. C., Zaklad, A. L., & Christ, R. E. (1992). Comparison of 4 Subjective Workload Rating-Scales. *Human Factors, 34*(4), 429-439.

Miller, S. (2001). *Literature Review: Workload Measures* (Document ID: N01-006). Iowa City, Iowa: The University of Iowa.

Appendix A: Cadet Training Scenario




UNITED STATES MILITARY ACADEMY
WEST POINT
CLDT III FTX BROAD SCENARIO GUIDE
START HERE
A. TLPs: Platoon conducts deliberate TLP's the night prior to start of FTX. The mission is a company attack in which the Platoon is the decisive operation.
1. Attack: Platoon Conducts company attack on OBJ Ant. Site exploitation on OBJ Ant leads to intel of resupply or movement of Key Weapon system (ADA Asset) prompting cadets to conduct an ambush the next day. Key Resource: 1x VISMOD
2. Ambush: Platoon conducts ambush on OBJ Bear and retrieves intel on possible Key Weapons System (ADA, Artillery, EW, etc.) on OBJ Cat. This prompts them to conduct a Raid on OBJ Cat Key Resource: 1x VISMOD and HMMWV
3. RAID: Platoon conducts raid on HVT ADA Asset. This enables air operations into AO. Platoon receives FRAGORD to destroy enemy near LZ Bull so they can conduct exfil by air out of LZ Bull. Key Resource: 1x UAS (Black hornet Capability)
4. Attack: Platoon conducts attack on OBJ Dog to enable air operations in and out of AO. Upon completion of OBJ Dog, Platoon is air assaulted back to Buckner for 24-hour refit. Key Resource: 1x Bunker
5. Raid: Platoon completed second deliberate TLP's on platoon mission Raid. Platoon conducts air assault into AO. Upon arrival, PLT Receives FRAGO to recover downed aircraft vic OBJ Elk with enemy on ground. Key Resource: 1x downed aircraft shell
6. Attack: Platoon conducts attack on OBJ Fox, one of the last strong holds of the enemy. Key Resource: 1x VISMOD
7. Force on Force: One platoon is on offense another on defense. The defensive platoon receives scenario that enemy is conducting a counterattack. The offensive platoon receives scenario that enemy has one last strong hold in AO. Key Resource: 2x UAS
OBJ ANT
OBJ BEAR
OBJ CAT
OBJ DOG
OBJ ELK
OBJ FOX
OBJ GOAT
Scenario Flow
Physical Flow
- Increased complexity, culminating in Force-on-Force and Crucible
- Incorporate MDO Threaded scenario

Appendix B: NASA-TLX

The following questions are about your experiences on the mission today. For the following, please circle the value you most agree with where 1 = Very Low and 10 = Very High.

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

Very Low									Very High
1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

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?

Very Low									Very High
1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

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

Very Low									Very High
1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

Performance: How successful do you think you were in accomplishing the goals of the task? How satisfied were you with your performance in accomplishing these goals?

Very Low									Very High
1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

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

Very Low									Very High
1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○