Developing an After-Action Review Tool to Enhance Training through Evidence-based Discussions

Grace Teo, Julian Abich IV
Quantum Improvements Consulting
Orlando, FL
gteo@qunatumimprovements.net,
jabich@qunatumimprovements.net

Gregory Goodwin
Combat Capabilities Development Command
Orlando, FL
gregory.a.goodwin6.civ@mail.mil

Blake Evans
Engineering & Computer Simulations, Inc.
Orlando, FL
bevans@ecsorl.com

ABSTRACT

When squads underperform on training drills, instructors are not always able to diagnose the problems to determine the areas for remediation. This can result in ineffective re-training efforts, wasted resources, and potentially adverse impacts on squad morale. Advanced sensors to capture data in real-time and applications for rapid data processing and visualization can be utilized as a part of the solution. When conducting battle drills, real-time data collected from individual sensors on each Soldier can lead to enhanced training and ultimately benefit squad performance if they are visualized well and made accessible at appropriate times. Data captured on squad movement, communication, and firing during the training drill encourage evidence-based discussions that can help users diagnose issues and identify corrective actions that drive behavioral change. This paper discusses the data visualization concepts for an after-action review (AAR) tool prototype user interface designed for U.S. Army drill instructors, squad leaders and platoon leaders. The solution incorporates features to (i) encourage consideration of training objectives, (ii) support notetaking and grading of the drill, and (iii) visualize key drill outcomes in a way that encourages users to explore relationships and derive actionable insights. Summaries of lessons learned from the prototype design process, and key points from interviews with subject matter experts and peer reviews are discussed. Further, concepts of possible future applications are provided that would benefit drill training, such as rolling up results to inform squad readiness and training needs, as well as concepts for incorporation into a drill data digital ecosystem (e.g., learning management system). Over time, this can yield data which can be utilized to develop higher echelon-level norms and performance standards associated with key performance indicators for different drills. All these serve to maximize training drill effectiveness and increase operational performance of Soldiers on the battlefield.

ABOUT THE AUTHORS

Grace Teo, Ph.D. is a Senior Research Psychologist at Quantum Improvements Consulting. Grace's research involves understanding and improving human performance under various conditions and in different contexts such as working with different technologies, and in teams. Other research interests include decision making processes and measures, vigilance performance, human-robot teaming, automation, and individual differences.

Julian Abich IV, Ph.D. is a Senior Human Factors Engineer at Quantum Improvements Consulting. He has 15 years of experience applying human factors & ergonomics (HF/E) principles, modeling & simulation (M&S) approaches, and instructional design methodology to the assessment, prediction, and improvement of human performance. Continuing to take a user-centered approach, his current work focuses on the application of HF/E and M&S to design and assess the usability and effectiveness of training systems with an emphasis on immersive technologies.

Gregory A. Goodwin, **Ph.D**. is a Senior Research Scientist at the Combat Capabilities Development Command – Soldier Center in Orlando, FL. For the last fifteen years he has worked for the Army researching ways to improve training methods and technologies. He holds a Ph.D. in Psychology from Binghamton University and an M.A. in Psychology from Wake Forest University

Blake Evans is a Product Owner and the PMO Manager at Engineering & Computer Simulations. With over fifteen years of working with or for the Department of Defense, his focus has been on producing exceptional products that meet the needs of the men and women of our armed services. Over the last few years, he has been focused on the implementation of technology to leverage the tracking empirical data in innovative ways to then analyze and revolutionize performance improvement and assess performance of products and processes.

Developing an After-Action Review Tool to Enhance Training through Evidence-based Discussions

Grace Teo, Julian Abich IV
Quantum Improvements Consulting
Orlando, FL
gteo@qunatumimprovements.net,
jabich@qunatumimprovements.net

Gregory Goodwin
Combat Capabilities Development Command
Orlando, FL
gregory.a.goodwin6.civ@mail.mil

Blake Evans
Engineering & Computer Simulations, Inc.
Orlando, FL
bevans@ecsorl.com

BACKGROUND

The after-action review (AAR) is a professional discussion conducted soon after the end of a training exercise or drill to review the training events, highlight well executed items, indicate where errors occurred, and identify areas for improvement (Morrison & Meliza, 1999). This suggests that the quality of the AAR depends on the ability of those in the discussion to understand what happened during the drill and accurately recall key drill incidences when Soldiers performed well and when errors were committed. Currently, AARs are mostly based on what Army instructors and the squad can remember of the drill, as well as any notes taken during execution of live training. These are not always complete or accurate and are egocentric to the individual. They are also constrained by the limits of human perception and information processing by what the instructor and Soldiers see, hear, and retain.

Based on these recollections, notes, and observations of the training drill, instructors make their assessment of the squad's performance. Apart from being based on imperfect information, there is also limited evidence that these largely subjective assessments are standardized across instructors who may be trained differently. The critical issue is that these assessments collectively inform future training activities for the Soldiers and serves as an indication of the squad's readiness (Foo, 2019). If they are not accurate, these assessments can result in ineffective re-training efforts and wasted resources at best, but at worst, could lead to potentially adverse impacts on squad morale or deployment errors that put the Soldier's life and limb at risk.

The recent proliferation in the commercial market of small network-connected sensors such as accelerometers, gyroscopes, GPS trackers, physiology monitors, and portable audio and video recorders has the potential to drastically increase the amount of data available to instructors. Combined with tools like smart targets that register timestamps and location of shots, and systems that can track the trajectory of live or simulated rounds, instructors could be swimming in an ocean of data from even a simple, squad level, battle drill. The Army has begun to conduct research to understand how to take advantage of the data flood that becomes possible from using lightweight sensors. Recent efforts by PEO Soldier at Camp Shelby focused on building a squad performance model using off-the-shelf sensors and researchers at the Soldier center have begun a multi-year program of research called *Measuring and Advancing Soldier Tactical Readiness and Effectiveness* (MASTR-E), which is looking at a wide range of biosamples and sensor-based measures taken from Soldiers conducting live training. The goal of much of the work is to determine whether any of these data sources allow us to monitor, predict, and enhance Soldier lethality with data-driven decision aids and targeted interventions (Benson, 2018).

The above presents numerous research and development opportunities to incorporate data from wearables and sensors into AARs conducted by instructors to train squads. In our interviews with subject matter experts (SMEs) who have extensive experience training squads, we have identified several concepts that, if incorporated into a tool, may support the instructors. In the following section, we present the research context which is training on a common battle drill. We will then discuss these concepts along with notional visualizations illustrating those concepts.

RESEARCH CONTEXT

The work presented in the current paper was part of an Army-sponsored research initiative to examine and understand the relationships among different measures of squad activity and performance on Battle Drill 2A (BD2A) to contribute

to insights for squad training. For the present effort, the team focused on the Support-by-Fire (SBF) Team, which comprised the Team Leader (TL), SAW Gunner (SAW), Grenadier (GRD), and Rifleman (RIF). In BD2A, the SBF Team's role is to provide suppressive fire that forces the enemy to take cover, preventing them from executing their mission while the Assault Team maneuvers to "flank" around SBF Team to get into assault position (see **Error! Reference source not found.**). At timepoint 1 in the figure, the SBF Team directs their fire at the enemy. At timepoint 2, the Assault Team moves into assault position. Right before the assault, the SBF Team would *Shift Fire* away from Assault Team, indicated by the dotted line in the figure, while Assault Team moves in for the assault, at timepoint (3).

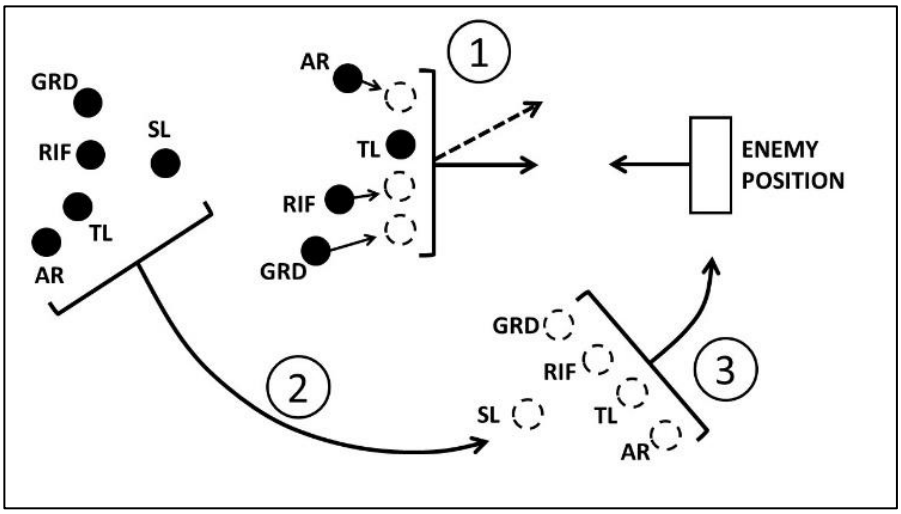


Figure 1: Depiction of BD2A. For both SBF and Assault teams, SL: Squad Leader, TL: Team Leader, AR: SAW gunner, GRD: Grenadier, RIF: Rifleman. (SBF team is directly in front of the enemy position)

CONCEPTS TO BE DEVELOPED IN AN AAR TOOL

1) Augment instructor observations with sensor data

A variety of sensors are available for collecting data on different Soldier behaviors and activity such as:

- ***Verbal communications***: Soldiers' verbal communications are recorded with portable audio recorders and their communications are transcribed in real-time by speech-to-text software, which also analyzes the communication content and identifies critical communications for the drill such as "reload callout", "malfunction callout", "Shift Fire callout", instance where the TL provided explicit directions on rate of fire, and when the TL communicated to the SL (i.e., "reported to higher").

- ***Firing Activity***: Inertia Measurement Units (IMU) fitted on Soldiers' weapons provide data on the timestamp of every shot and the weapon that fired the shot.

- ***Shot Accuracy***: Measures of shot accuracy are available from Location of Misses and Hits (LOMAH) sensors, smart targets such as Marathon targets for live rounds and Multiple Integrated Laser Engagement System (MILES) for simulated rounds.

- ***Soldier's spatial orientation and movement***: Data about Soldiers' helmet orientation/heading within the drill location are provided by 3D motion capture (MOCAP) devices fitted to Soldiers' helmets and attached GPS trackers. Recording devices on overhead drones can capture the bird's eye view of the squad over the different phases of the drill. When processed and analyzed, these can provide information about Soldiers' speed of movement, location and the squad's formation.

2) Develop empirically based performance measures from sensor data

For these data to yield meaningful information about Soldier performance, suitable measures of different aspects of performance must first be developed. For instance, data from IMU and MILES can yield shooting measures such as rates of fire, hit, and miss rates, while data from audio recorders can provide measures of leadership communication. Data from GPS trackers can result in measures such as efficiency of movement. If a tool can be developed that collects these data from these sensors and processes them appropriately, empirical measures of performance over the course of the drill can be obtained from these data. Compared to instructor observations, these measures are more objective and offer a way to empirically assess the squad on previously unobservable constructs such as the Soldiers' coordination of fire. With these objective measures, standards and norms on different aspects of performance can be developed to give empirical definitions to important operational constructs such as squad readiness. Table 1 shows examples of performance measures that may be obtained from various sensor data.

Table 1: Performance measures that can be obtained from sensor data

Sensor data source	Soldier behaviors	Examples of performance measures
1. Audio recorders	Communications	• Leadership communications • Communications on weapon status
2. IMUs on weapons	Firing activity	• Number and timing of shots fired • Firing rate
3. MILES & LOMAH	Shot accuracy	• Hits (vital and non-vital hits) • Misses
4. MOCAP on helmets	Heading and orientation	• Direction of Soldier's attention • Amount of time Soldier was facing in the wrong direction
5. GPS trackers	Movement and location	• Squad formation • Movement efficiency

3) Address training objectives with sensor data

In a report on research-driven recommendations for optimal AARs, it was stated that AARs should be based on a presentation of ground truth from various perspectives and recording of mission action to minimize erroneous understanding of drill events based on false memories (Newlin & Bliss, 2007). This suggests that AARs that involve objective measures derived from sensor data can enhance AARs. In addition, by focusing on specific episodes and incidences in the drill identified with the help of sensor data providing a more accurate portrayal of drill events, AAR discussions can be steered towards particular Soldier actions and the impact of those actions on drill outcomes. This enables actionable goals to be set, which fosters motivation and a sense of direction (Locke & Latham, 2002).

As a training activity, AAR discussions should support training objectives and goals for the drill. This includes training objectives that not only pertain to individuals, but also to the team. For AARs to address the team-based training objectives, AAR discussions should encourage Soldiers to examine the interactions within their team and discover how their actions and behaviors impact each other to produce the performance outcomes of the entire squad. AARs that are not solely focused on individuals' tactical skills or performance, but include discussions on teamwork, are aligned with the concept of integrated AARs (IAARs), which promotes a more team-oriented perspective of training and performance (Townsend et al., 2018).

An example of how a tool can be developed to use sensor-based empirical measures to address a team-based training objective is with training the Support-by-Fire (SBF) team in the squad to coordinate their firing activity. In Battle Drill 2A (BD2A), the volume of suppressive fire produced by Soldiers of the SBF Team is affected when weapons have to be reloaded or when they malfunction. To maintain the volume of suppressive fire, team members must be aware of the firing activity of other weapons so that they can increase their own firing to compensate for any decreases in the overall fire volume. For a particular BD2A training exercise, a training objective for SBF team Soldiers could be to ensure that they inform the team whenever they reload their weapons so that the others can respond appropriately. If the tool can "flag" each time the IMU sensors on a Soldier's weapon indicates that a reload is taking place and

determine if there is an accompanying "Reload" call out from processing the audio recordings of that Soldier's verbal communications, then the instructor can evaluate the extent to which the training objective was achieved by the Soldier. The Soldier will also be able to understand the impact of his verbal communications on the team performance.

4) Design appropriate data visualizations

Including data from the multiple sensors and devices in the AAR discussion can potentially lead to data and information overload since the additional information can easily be overwhelming and jeopardize the instructor's understanding of the drill (Bawden et al., 1999; Jackson & Farzaneh, 2012). Too much data or complex graphs and tables may require too much effort or expertise to interpret. To minimize this, the information should be curated and represented appropriately to reduce cognitive load and facilitate inferential and strategic thinking (Shneiderman, 1996; Tufte, 1990; Ware, 2004). Ideally, data should be presented in a way that readily conveys important information about key aspects of the performance, to help instructors assess the squad and help the squad set realistic, measurable goals that they can achieve through training. Visualizations should be customized to different audiences, which can range from Soldiers, their instructors, and leaders, and appropriate for their intended purpose. For instance, a summary table indicating the squad's performance for the entire drill may be useful to a platoon leader, but a graph of changing firing rates for individual Soldiers will enable an instructor to pinpoint when the Soldier failed to fire appropriately. Finally, it is crucial that visualizations seamlessly link to traditional, doctrinal measures of performance like the Training and Evaluation Outline (T&EO) grade sheet (Foo, 2019). This is so that the various stakeholders can appreciate the relevance of the novel sensor data to existing military performance standards, and that information from those additional data be aligned with T&EO items on which instructors grade squads in various items/actions using the "GO" or "NO GO" format. A possible way to achieve this is to identify the type of behaviors (i.e., firing, moving, or communicating) that is most related to each T&EO item. If more "NO GO" T&EO items were related to firing, then the instructor would review the visualizations pertaining to firing activity, and the same would be done for movement and communication. The following are notional visualizations pertaining to different performance aspects that an instructor may find useful when conducting squad-level AARs.

Firing and Communications

For BD2A, the SBF team's main role is to provide suppressive fire. If there were multiple "NO GO" items in the T&EO grade sheet that pertained to their firing or communications, the instructor may choose to focus on how the squad communicated and fired over the course of the drill. To avoid overwhelming the instructor, an overview of critical drill incidences may first be presented (Figure 2), with further details on that timepoint readily available "on demand" (Figure 3). If the tool were to be developed for the instructor to view these visualizations on a touchscreen tablet for instance, those details can be accessed with a single touch of an icon in the display.

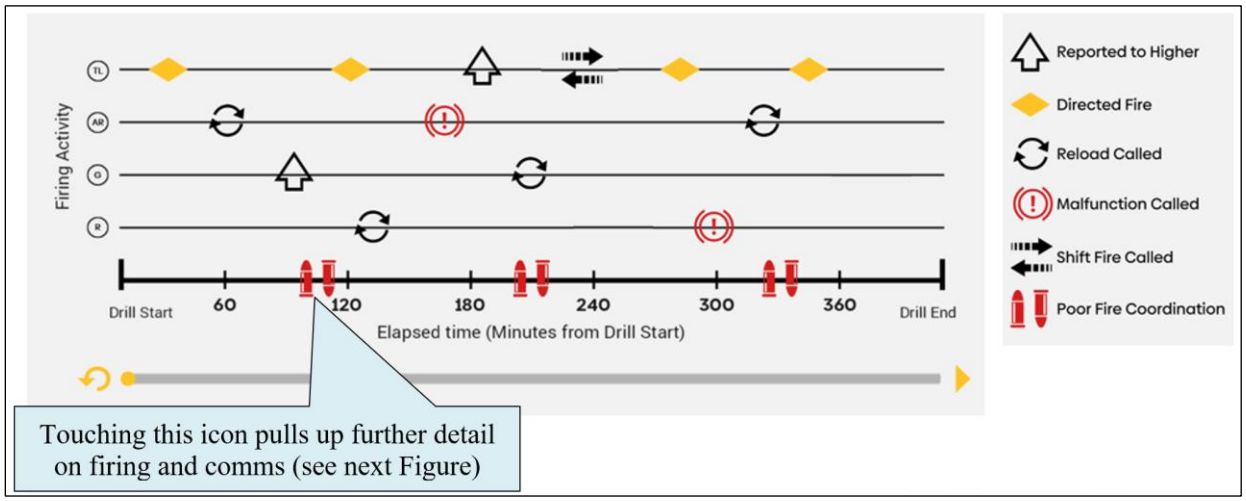


Figure 2: Notional visualization depicting overview of squad's firing and communications

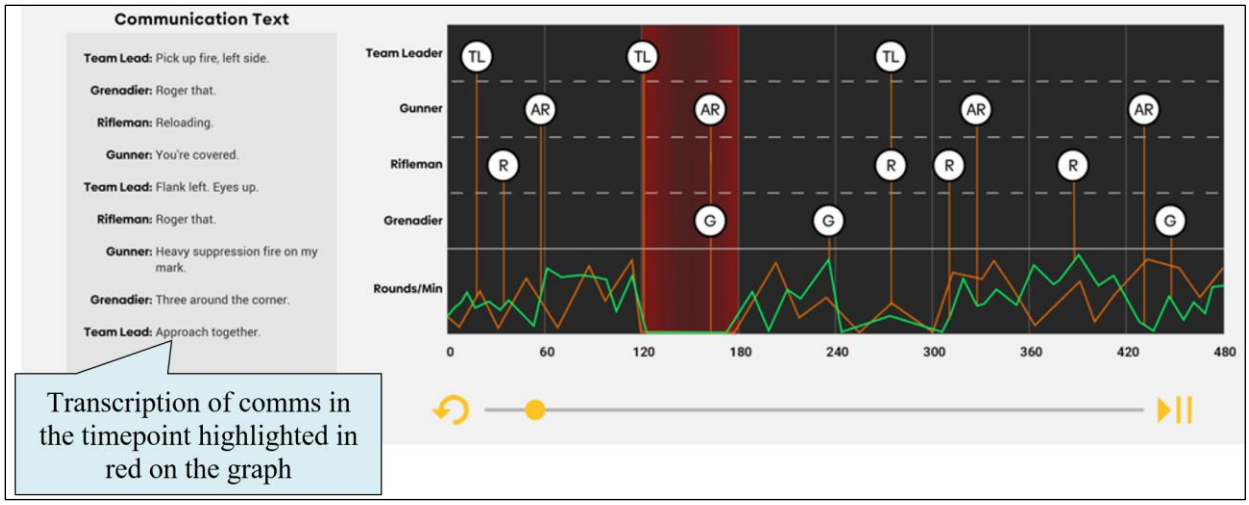


Figure 3: Notional visualization of further details on the firing and communications at a particular timepoint

Movement and Drill Safety

Apart from the suppressive fire provided by the Alpha team at the beginning of the drill, the success of a squad in executing BD2A also lies in the squad's ability to move, maneuver, and keep formation. If the T&EO grade sheet results indicated that many of the squad's "NO-GO" items related to movement (M), then the instructor may wish to review the playback of the Soldiers' movements. The tool can be designed to provide 'playback' of the Soldiers' movement and navigation on a 3D or topographical map (see Figure 4 and Figure 5).

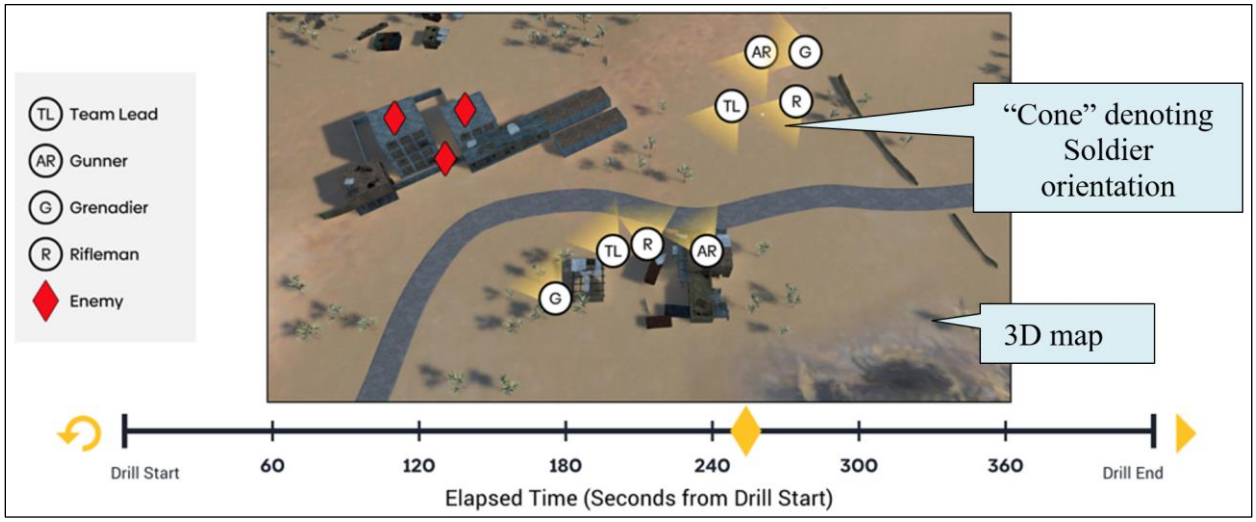


Figure 4: Notional visualization of a 3D map depicting Soldiers' movement and orientation

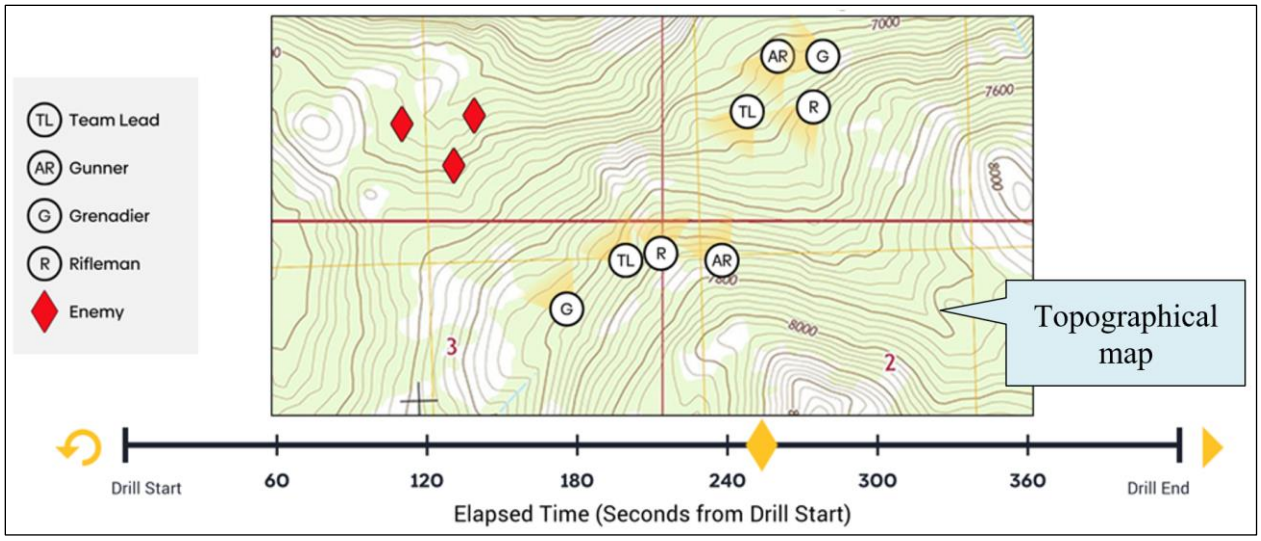


Figure 5: Notional visualization of a 3D map depicting Soldiers' movement and orientation

Another aspect of drill performance is safety. Soldiers are not to engage in any behavior that increases fratricide risk. For instance, they are not to point their weapons at any friendly forces at any time. Sensors on the Soldiers' weapons would allow detection of their orientation. This information can be processed and presented by the tool and discussed during the AAR.

Accuracy

For effective fire suppression, the SBF team's shots must both be well-coordinated and accurate enough to pose a threat to the enemy, requiring them to take cover. The squad will also have a significant advantage if their shots that are well-aimed and took out enemy soldiers. Our SMEs mentioned that information about Soldier's marksmanship in the drill would be valuable in helping them train the squad. They stated that a tool that initially presents an overview of the squad's shot accuracy (Figure 6Figure 5), with details of shot placement for individual Soldiers available on-demand (Figure 7) would be useful.

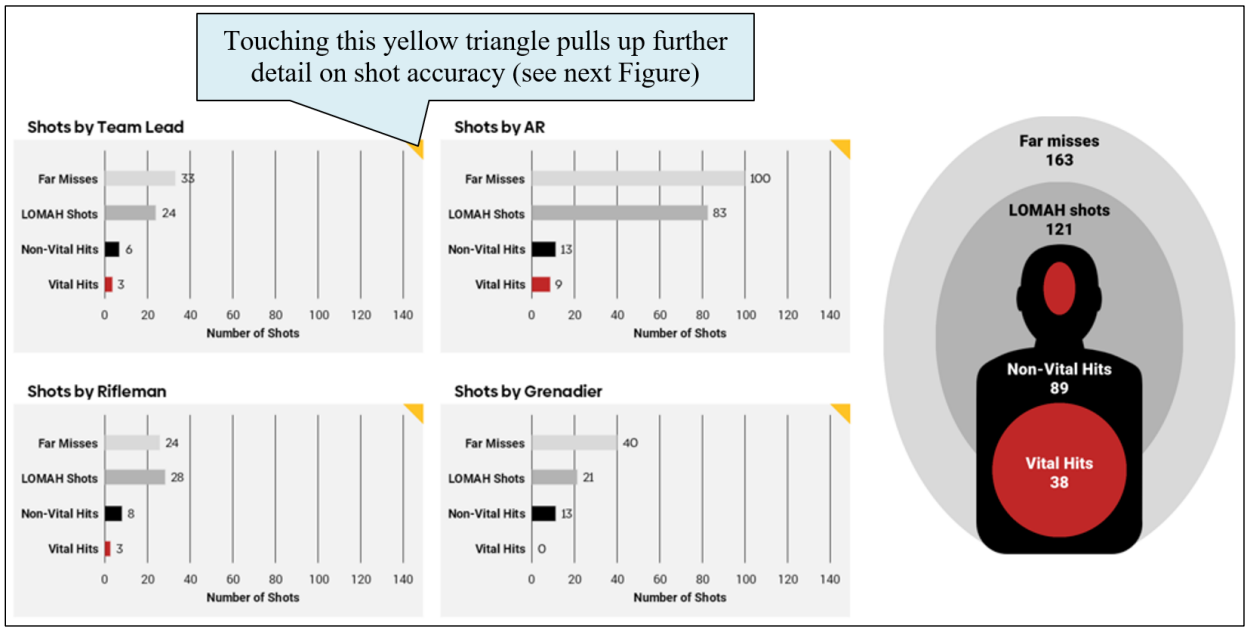


Figure 6: Notional visualization of the overview of the SBF team's shot accuracy

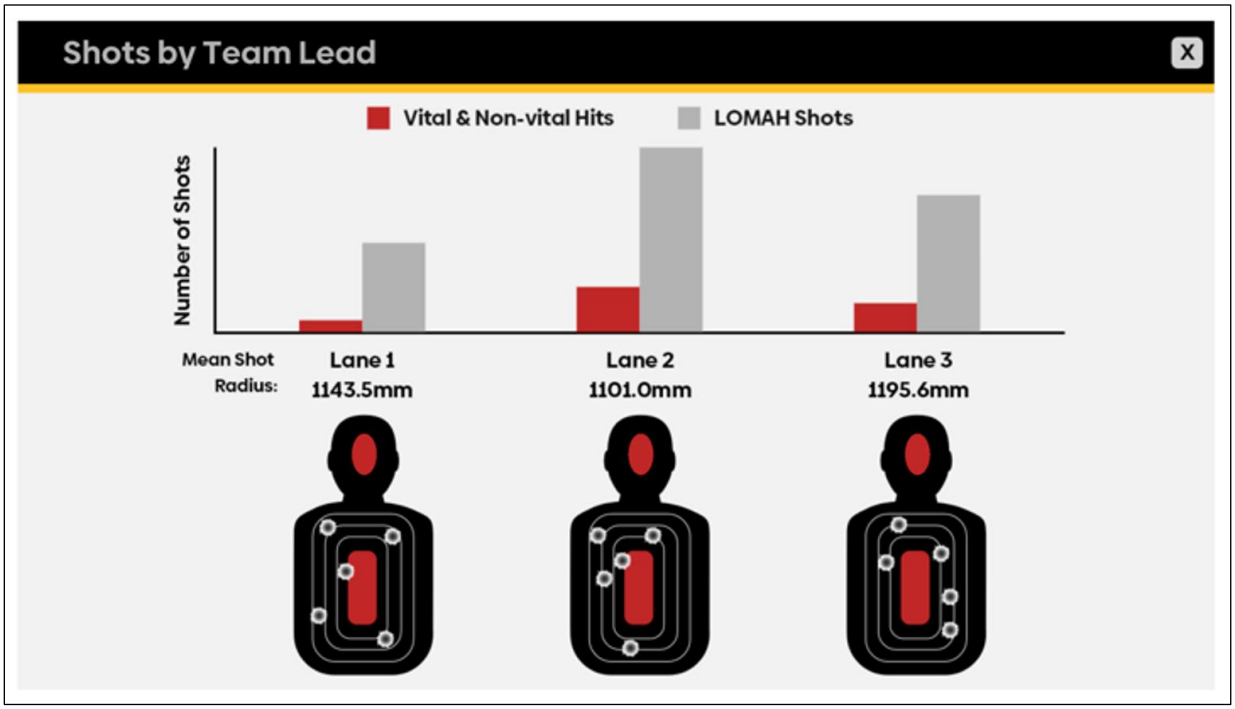


Figure 7: Notional visualization of further detail of an individual Soldier's shot accuracy

5) Support instructor note-taking

Taking notes when a drill is in session can be a challenging task. This is because a squad-level drill typically involves seven to nine Soldiers executing the drill and are typically fast-paced with a level of uncertainty and danger. At each stage of the drill, instructors not only have to attend to how each Soldier performs, they must also observe how they perform as a squad, as well as be vigilant about the Soldiers' safety. SMEs have mentioned that they would use their phones to record some of these notes since writing them down can be inefficient.

The tool can be developed to display data from the sensors in real-time so that instructors can monitor different aspects of the drill that may not be observable, such as firing rates. If the tool were to be on a tablet with a touchscreen, instructors would be able to mark the timeline to note instances of "poor fire coordination", "slow response", or other incidences that affect the squad's performance (see Figure 8). This capability can be made available for the various sensor data such that the instructor can opt to monitor squad firing activity, movement/location, shooting performance or communications in turn when the drill is in progress. In addition, instructors may also take "voice notes", reducing the load on short-term memory, and eliminating the need to look away from the drill to jot down notes.

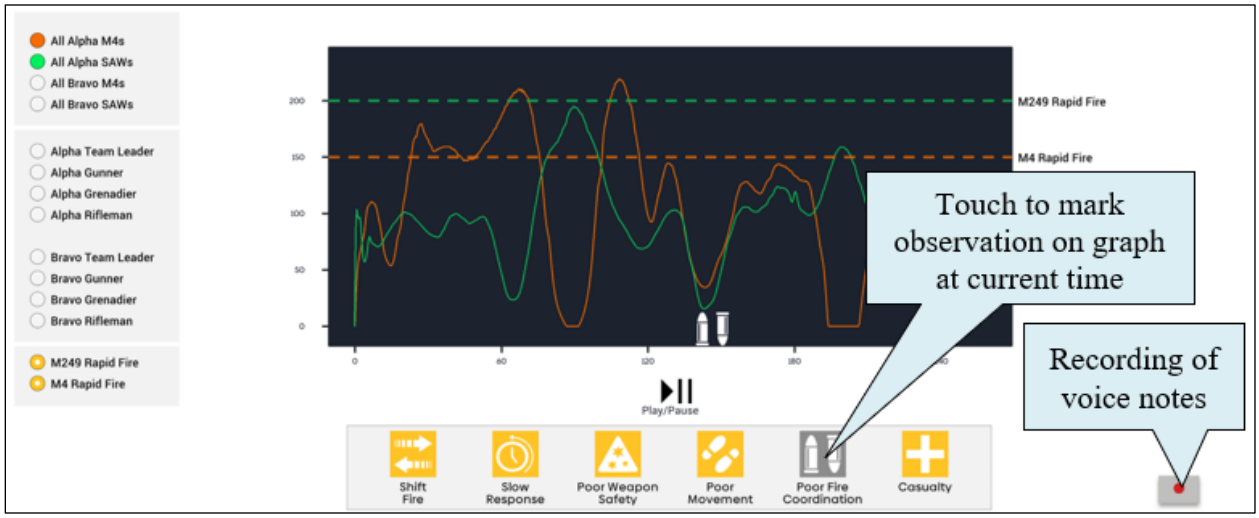


Figure 8: Notional display of real time fire coordination with option to mark the time when notable incidences occurred in the drill. Graph based on actual data

6) Support training-related concepts such as progress tracking, setting performance standards to train to

Since the AAR tool is not an isolated training event, the tool should be developed to enable instructors to retrieve the squad's results from a previous training exercise executing the same drill to show how the squad has improved in performing that drill. By extension, instructors will also be able to compare different squads on the same drill to understand how a particular squad fares relative to others. Comparisons should be matched on factors such as whether drills were conducted in the day or night, and if they were executed *live* or *dry* without ammo (see Figure 9). Over time, with multiple squads' data on the same drill, it would be possible to develop performance standards or expectations of how squad typically progress through the training to reach proficiency. Our SMEs stated that such data can be further rolled up to platoon or company-level reports that indicate Soldier readiness and/or deficiencies in training under specific conditions, such as weather conditions.

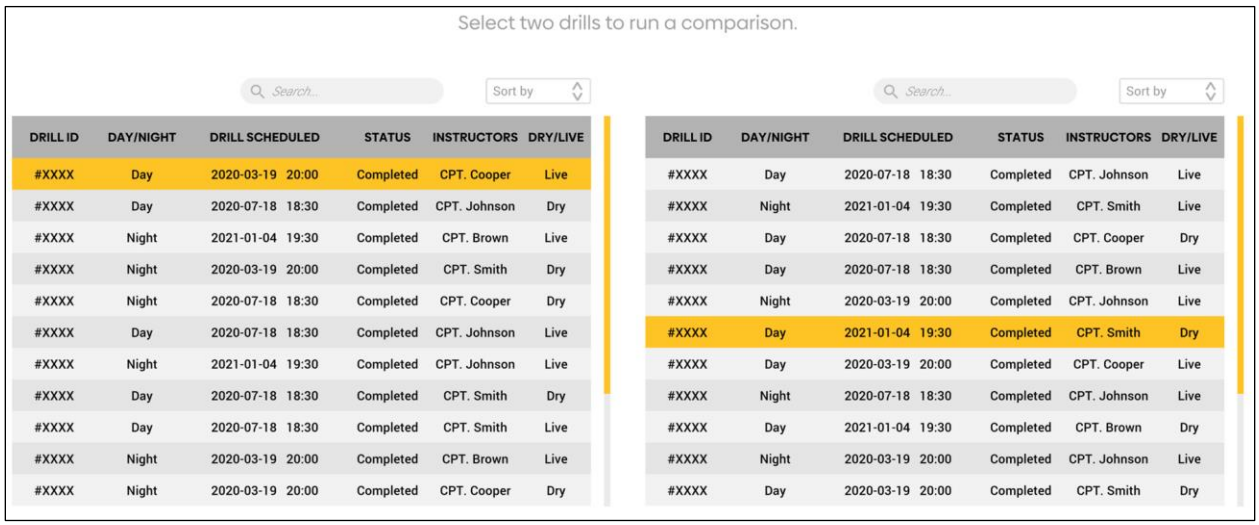


Figure 9: Notional tool interface for comparisons

CONCLUSION

AAR discussions can be enhanced through the inclusion of data that will eventually be available with the use of additional sensors and software that can rapidly analyze and summarize performance data. The goal of this paper was to present a glimpse of what we expect future AAR tools may provide to instructors and Soldiers to help diagnose errors and to enhance their performance with accurate, objective feedback. By envisioning data processing and analytical capabilities that may be available in the near future, several notional visualizations have been drafted in the attempts to get Soldiers and instructors to think about how these data might help to enhance AAR discussions and how best to depict these measures to facilitate training. In addition, it is possible that some of these novel measures may not be particularly useful to instructors or Soldiers and so perhaps would not be needed for an AAR tool. More data is not always better and having Soldiers wear sensors that do not add value to the training event is both a waste of training resources and puts an unnecessary burden on Soldiers when conducting training.

Feedback from the training community about the usefulness of these performance measures for conducting AARs will also help to drive future requirements and prioritize investments in research and development for AAR tool technologies. For example, voice recognition software is constantly evolving to accurately develop transcripts of voice communications by infantry squads during live training events. Technologies for tracking location as well as limb position, weapon orientation, eye tracking, etc. work well in controlled environments but are not as effective in woods with limited network access. Helmets and body armor often interfere with the placement/wearing of these sensors. With the rapid pace of technological advancements, the vision for the sensors and drill data presented herein may become reality sooner rather than later, and more research will be required to determine the value of that investment. The present research is only a start of that research.

ACKNOWLEDGEMENTS

We would like to express our sincere thanks and deep gratitude to Mr. Karl Arbogast and Mr. Fred Dupont, our subject-matter experts (SMEs) who have been generous with their time and help on this effort. Their valuable inputs and feedback have helped shaped key aspects of the AAR tool prototype. This research was sponsored by the U.S. Army Combat Capabilities Development Command - Soldier Center (CCDC SC). 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 CCDC-SC or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

Bawden, D., Holtham, C., & Courtney, N. (1999). Perspectives on information overload. In *Aslib Proceedings* (Vol. 51, Issue 8, pp. 249–255). Association for Information Management (ASLIB). https://doi.org/10.1108/EUM0000000006984

Benson, J. (2018). *Natick leads first-of-its-kind, Soldier-readiness study*. Www.Army.Mil. https://www.army.mil/article/205789/natick_leads_first_of_its_kind_soldier_readiness_study

Foo, H. S. (2019). *Training and Evaluation Outlines (T&EO): Usage and Scoring Method Preference for Task Steps and Sub- steps* (ARI Research Note 2019-01; p. 16). U.S. Army Research Institute.

Jackson, T. W., & Farzaneh, P. (2012). Theory-based model of factors affecting information overload. *International Journal of Information Management*, *32*(6), 523–532. https://doi.org/10.1016/j.ijinfomgt.2012.04.006

Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation A 35-Year Odyssey. *Psycnet.Apa.Org*, *57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Morrison, J. E., & Meliza, L. L. (1999). *Foundations of the After Action Review Process*.

Newlin, E. A., & Bliss, J. P. (2007). *Integrated theory of after action review foundations*. Final Report, Project CR03). Portsmouth, VA: MYMIC, LLC.

Shneiderman, B. (1996). Eyes have it: A task by data type taxonomy for information visualizations. *IEEE Symposium on Visual Languages, Proceedings*, 336–343. https://doi.org/10.1109/vl.1996.545307

Townsend, L., Johnston, J., Ross, W. A., Milham, L., Riddle, D., & Phillips, H. (2018). An Integrated After Action Review (IAAR) Approach: Conducting AARs for Scenario-Based Training Across Multiple and Distinct Skill Areas. *International Conference on Engineering Psychology and Cognitive Ergonomics*, 230–240.

Tufte, E. R. (1990). *Envisioning information*. https://doi.org/10.2307/3325378
Ware, C. (2004). Perception for Design. In *Information Visualization* (Vol. 22).