Taking a Data-Informed Approach to Squad Training Evaluations

Michael King, Julian Abich IV
Quantum Improvements Consulting
Orlando, FL
mking@quantumimprovements.net
jabich@quantumimprovements.net

Clifford Hancock, Meghan O'Donovan
U.S. Army Combat Capabilities Development Command – Soldier Center
Natick, MA
clifford.l.hancock4.civ@army.mil
meghan.p.odonovan.civ@army.mil

Gregory Goodwin
U.S. Army Combat Capabilities Development Command – Simulation and Training Technology Center
Orlando, FL
gregory.a.goodwin6.civ@army.mi

ABSTRACT

The U.S. Army S&T is interested in using small, wearable sensors to automatically collect behavioral and physiological data from Soldiers in training to create a digital record of their performance and to augment trainer evaluation practices as they provide assessments and feedback. Collecting large amounts of data from wearable sensors is increasingly practical. Still, guidance is needed on how to visualize these data to support informed evaluations by expert observer-controller/trainers (OC/Ts). This paper describes our process for developing an example of such visualizations using an Army use case. Our team collected physiological and behavioral data to evaluate Army squad performance in a common infantry battle drill. In addition, we interviewed OC/Ts to determine what information they typically use to assess performance. Using their identified data needs to guide us, we developed a measurement framework that augments their evaluation process. Results from this study are presented and discussed in terms of how they promote a well-balanced, data-informed approach toward squad training and evaluation.

ABOUT THE AUTHORS

Michael King, Ph.D., is a Research Psychologist II at Quantum Improvements Consulting (QIC). At QIC, Michael has led various training and human performance projects. Michael's research interests at QIC center on human performance, training, and the predictors of success in team-based performance. Michael holds a Ph.D. in Experimental Psychology from Case Western Reserve University.

Julian Abich IV, Ph.D., is a Senior Human Factors Engineer at Quantum Improvements Consulting. He holds a Ph.D. in Modeling and Simulation. He has expertise in applying human factors and ergonomics (HF/E) principles, modeling and simulation (M&S) approaches, and instructional design strategies. He has extensive experience conducting user experience evaluations across various human-computer interaction and training applications.

Clifford Hancock, M.S., is a biomechanical engineering researcher with the U.S. Army Combat Capabilities Development Command – Soldier Center (Natick, MA, USA). He received his M.S. in Biomedical Engineering from the University of Texas at Arlington in 2013 and his B.S. in Engineering Science and Mechanics from the Virginia Polytechnic Institute and State University (Virginia Tech) in 2011.

Meghan O'Donovan, M.S., is a principal investigator and project officer with the U.S. Army Combat Capabilities Development Command – Soldier Center (Natick, MA, USA) who has led several efforts in the areas of human performance, military equipment evaluation, and performance augmentation. She currently leads the U.S. Army Small Unit Performance Analytics program (SUPRA). She holds an M.S. in Biomedical Engineering with a concentration in Biomechanics from the University of Rochester.

Gregory Goodwin, Ph.D., is a Branch Chief 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.

Taking a Data-Informed Approach to Squad Training Evaluations

Michael King, Julian Abich IV
Quantum Improvements Consulting
Orlando, FL
mking@quantumimprovements.net
jabich@quantumimprovements.net

Clifford Hancock, Meghan O'Donovan
U.S. Army Combat Capabilities Development Command – Soldier Center
Natick, MA
clifford.l.hancock4.civ@army.mil
meghan.p.odonovan.civ@army.mil

Gregory Goodwin
U.S. Army Combat Capabilities Development Command – Soldier Center STTC
Orlando, FL
gregory.a.goodwin6.civ@army.mi

INTRODUCTION

Over the past decade, data-driven methodologies have become essential across various industries, where advanced analytics and performance metrics provide a competitive edge. The U.S. Army has been researching these capabilities, particularly in the training and education domain, by exploring sensor-based data collection to understand individual and team-level performance better. However, it is unclear what the best uses of these technologies are. The Small Unit Performance Analytics (SUPRA) initiative attempted to answer these questions by collecting and analyzing squad sensor data and executing multiple Battle Drill 2A: React to Contact (BD2A) iterations. One of the objectives of SUPRA was to determine which performance measures would help differentiate the levels of performance across squads. An additional question was whether those sensor-based measures would be useful in training and operational contexts.

During the SUPRA study, researchers learned that there is a need to provide sensor-based data and visualizations to aid observer-controller/trainers (OC/Ts) in squad assessments. This need became especially clear after a simulated fratricide incident in one of the BD2A trials. This event was not observed by OC/Ts present and thus did not impact that squad's performance evaluation (King et al., 2023). Access to specific data produced by wearable sensors would have benefitted the OC/Ts by providing an array of objective performance data. However, it is not as simple as just giving OC/Ts data. Interpreting raw data is challenging, arduous, and time-consuming. Therefore, the data must be presented in a way that facilitates efficient assessment and feedback. Assessment approaches should integrate objective data with expert judgment to overcome the limitations of either approach alone and provide a comprehensive understanding of Soldier performance. This paper describes our process towards designing data visualizations that can augment OC/Ts in their squad assessments in a data-informed rather than data-driven approach. This approach highlights the importance of prioritizing the end user's needs, preferences, and goals, in this case, the OC/T, throughout the visualization design process.

Background

As part of the SUPRA initiative led by the Combat Capabilities Development Command (DEVCOM) under the Measuring and Advancing Soldier Tactical Readiness and Effectiveness (MASTR-E) program, researchers collected and analyzed squad sensor-based data from 18 dismounted infantry squads executing BD2A following a Movement to Contact operation. This exercise involves the squad locating and suppressing the enemy, establishing supporting fire, and assaulting the enemy position using fire and maneuver. The drill commences when the Alpha team, also known as the Support-by-Fire team, initiates fire or when the squad is fired upon by the Opposing Force (OPFOR). As soon as this occurs, the Alpha team produces suppressive fire that forces the enemy to keep their heads down, preventing that enemy from executing their mission. Suppressive fire presupposes a specific volume and accuracy of coordinated firing from the Alpha team's M4 rifles and Squad Automatic Weapons (SAW). There should be a continuous volley of shots fired toward the enemy's general direction at all times, forcing them to take cover and fixing them in place (see point 1 in Figure 1). While the enemy is suppressed, the Bravo (i.e., Assault team) moves in (see point 2 in Figure 1) to flank either side of the enemy position (see point 3 in Figure 1). Before the Bravo team's

assault, the Alpha team must shift their fire away from the direction of the intended assault and lift fire once the Bravo team moves in for the assault (see dotted line in point 1 in Figure 1). After the Bravo team has swept through the enemy position, the Alpha team follows to ensure the enemy has been eliminated.

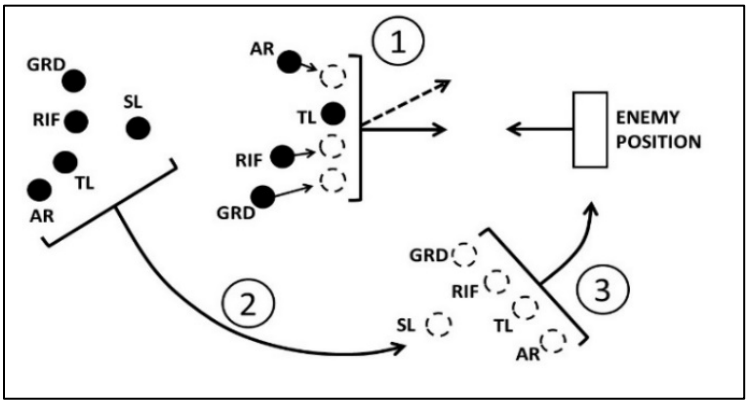

Figure 1. Depiction of BD2A with duty position labels: squad leader (SL), team leader (TL), automatic rifleman (AR), grenadier (GRD), and rifleman (RIF).

The SUPRA program leveraged a range of sensor-based measurement techniques (for a detailed description of the SUPRA methodology, see O'Donovan et al., 2023), including (a) audio recordings to capture squad communications, (b) inertial measurement unit (IMU) data to record weapon activity, weapon orientation, and helmet orientation, (c) camera-based motion tracking, (d) GPS tracking to examine squad movement and formations, and (e) several physiological measures including cardiovascular and blood chemistry. These measurements were able to provide evidence for most aspects of squad BD2A performance, such as whether leaders gave key communications, the consistency of squad formation spacing during flanking movements, the extent to which a squad has 360-degree security at a given time, and how well a squad timed their firing to ensure suppression of the enemy was achieved. While wearable sensors provided detailed data points on Soldiers' behaviors, OC/Ts observed the drills and assessed how well each squad performed in the drill.

OC/Ts are critical in assessing small team performance in the Army. During training exercises, OC/Ts closely monitor and evaluate squad performance. They use standardized evaluation criteria organized in training and evaluation outlines (T&EO) issued by the Army. One of the essential roles of OC/Ts is providing constructive feedback to the squads they observe during AARs. Assessing squad performance relies heavily on what the OC/Ts can observe and remember throughout the drill. In SUPRA, there were limited opportunities for OC/Ts to incorporate sensor data into their evaluations. Thus, these ratings were subjected to bias and error. These biases are not unique to OC/Ts conducting BD2A squad evaluations and are nearly impossible to overcome for any observer evaluation. To fully understand squad performance, assessments should be complemented by sensor-based data sources. Relying solely on observer ratings may limit the richness and depth of the performance assessment.

The SUPRA effort successfully demonstrated that various data streams can be collected during live squad-level force-on-force battle drills and that these data may fill the gaps in OC/Ts' ability to assess squads accurately. However, how exactly OC/Ts can incorporate these data into their evaluation process is unclear. Therefore, the goal of the present study was to inform the design of data visualizations that would assist OC/Ts in evaluating live squad-level force-on-force exercises, contributing to a data-informed squad evaluation approach.

APPROACH

To develop a data-informed assessment approach for Army OC/Ts, we employed a user-centered design process that involved two core phases of interviews and feedback sessions with active duty OC/Ts. First, we conducted semi-structured interviews to identify gaps in OC/Ts' current assessment workflow for BD2A to understand better where data could enhance their assessment capabilities from their perspective. We created concept data visualizations based on the feedback to illustrate how objective data could inform these identified gaps. Subsequently, we conducted a second round of semi-structured interviews with different OC/Ts to gather input on these concepts. These discussions

aimed to evaluate the visualizations' effectiveness and solicit suggestions for improvement. By iterating on the feedback from both sets of interviews, we refined our approach to ensure that final OC/T recommendations are closely aligned with their needs and that data is informing their evaluations where they feel it can best be utilized. This process serves as a framework for developing a user-informed approach toward small team assessment. Further, this process highlights a replicable method, agnostic of performance space, for identifying user requirements and developing data visualizations that effectively support their goals.

PHASE 1: OC/T DATA NEEDS INVESTIGATION

During Phase 1, researchers conducted semi-structured interviews with OC/Ts to establish what information they use to evaluate performance during BD2A and solicit ideas for how additional data may help. Researchers asked OC/Ts during these discussions to reveal their processes for assessing squad performance during BD2A. In addition, they were asked targeted questions regarding where they believe they have gaps in their ability to evaluate squads accurately and what type of tools, information, or data they believe would help augment their assessments. Researchers then worked to develop data visualization concepts of BD2A, utilizing data from the SUPRA project to fill in the gaps of OC/Ts' assessments where data visualizations are needed the most.

Methods

OC/Ts were interviewed individually for 60-90 minutes. The interviews were conducted remotely by two researchers. The interview began with an explanation of its goals and format, followed by a step-by-step discussion of T&EO checklists for BD2A. Questions focused on current evaluation methods, gaps in information, and potential data sources for improvement for each T&EO assessment. The interview concluded with demographic questions, and OC/Ts were thanked and dismissed.

Participants

We interviewed 10 participants with prior experience assessing squads completing BD2A as an OC/T. The participants were current and former OC/Ts stationed at the Joint Readiness Training Center (JRTC) at Fort Johnson, LA, at the Sergeant First Class (E-7, n = 8) and Staff Sergeant (E-6, n = 2) ranks. The total estimated number of battle drills evaluated by each OC/T ranged from 6 to 500, with Staff Sergeants on average having evaluated 25 battle drills in their career (M = 25, SD = 21.2) and Sergeants First Class on average having evaluated 135.5 battle drills in their career (M = 135.5, SD = 166.4).

Results

We conducted a structured thematic analysis to analyze the data from the Phase 1 interviews. First, researchers took detailed notes during interviews to capture essential content and answers to specific questions asked by interviewers. Researchers then reviewed the recorded conversations to identify initial qualitative data from OC/T responses to questions, representing recurring concepts and unique insights related to evaluation challenges, criteria, and data needs for effective BD2A assessments. This data was then organized into a dataset and discussed and refined through iterative consensus meetings to develop a coding framework. Using this framework, the researchers systematically coded the entire dataset, ensuring consistency and coverage of themes. Once coding was complete, we grouped related codes into overarching themes that captured the essence of the OC/Ts' assessment processes and perspectives on data that may enhance these processes. Themes were refined and validated against the data to ensure they accurately represented the interview content. This process resulted in well-defined themes that provided a detailed understanding of the current evaluation practices, challenges, and potential areas where gaps exist in squad performance assessments. These analyses identified several areas where data visualizations would help their evaluations. See Table 1 for a summary of the thematic analysis. The most prominent areas were GPS positioning data, communication, firing rate, and shot accuracy.

Table 1. OC/T Data Needs Thematic Analysis Results.

Theme	Details
GPS Locations and Movement (n = 10)	• Seeing GPS locations and movements of Soldiers from MILES (i.e., dots moving on a map) • Timestamp trackers for GPS movement of individual Soldiers

	• Visualizations of their movement throughout the drill • Aerial shots would be helpful. Soldiers often don't get to see the "big picture" and where the rest of the platoon and company is during a drill
Communications (n = 7)	• Communication transcriptions of the entire squad • Squad Leader communications would be especially helpful • Timestamps for communication logs
Firing Rate (n = 9)	• Visualizations of firing rates • Shot counts • Ammo depletion counter • Rate of fire and ammo expenditure can help to evaluate the extent to which a squad was a "well-tuned orchestra"
Shot Accuracy (n = 4)	• Target hit-and-miss data • Visualizations showing details of kills and deaths (who killed whom and when) • Sensors on the body to indicate where hits were taking place. It helps to know what portions were exposed and what was covered or concealed
Visualization Overlays (n = 7)	• Communications overlayed with rates of fire • Communications overlayed with GPS position and movement • Video of movements overlayed with firing rate

Note. The Theme column shows the number of OC/Ts, out of 10, who made comments reflecting an area where data visualizations would help their evaluations.

CONCEPT VISUALIZATIONS

The thematic analysis results guided the development of concept data visualizations and the overall data-informed evaluation framework for OC/T squad-level assessment. Common design principles and patterns familiar to Soldiers, drawing from industry standards that can be seen in existing assessment tools (Stone, 2017; Sinatra et al., 2020) and first-person shooter video games (Zagata, & Medyńska-Gulij, 2023) were incorporated. Visualizations were created using Microsoft PowerPoint. These concept visualizations were created solely for research purposes to drive discussion and feedback with OC/Ts in Phase 2 of this study. See below for examples of these data visualizations in each key area in Table 1.

GPS Positioning

For GPS positioning data, OC/Ts expressed a desire to see Soldiers' locations, movements, and aerial shots of movement throughout the drill. Researchers included visualizations with head orientation cones to aid in assessing squad security. Concept visualizations were created to showcase these features using fabricated Soldier GPS data and Google Maps images of Fort Johnson. Figure 2 provides an example of such a GPS positioning visualization. In this playback visualization, the squad's movement is displayed, with locations of friendly (1) and enemy (2) units updated to show movement. Head orientation (3) updates to provide an approximation of gaze directions, and a "snail trail" (4) shows the history of squad members' movements. A compass (5) with combined head orientation cones displays overall squad security levels. A menu (6) also allows users to toggle features on and off.

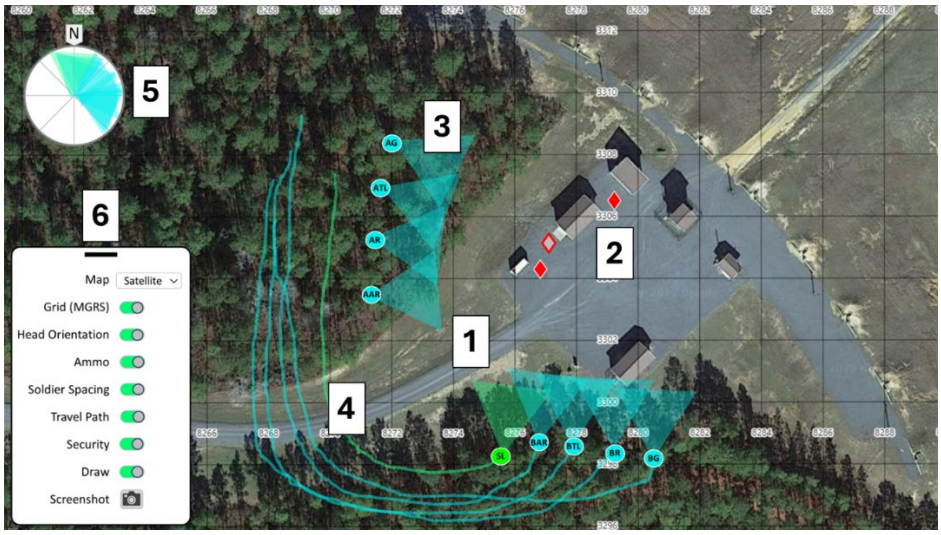

Figure 2. GPS Positioning Visualization Concept provided to OC/Ts.

Communication

OC/Ts unanimously (n = 10) expressed that communication transcripts of the entire drill with timestamps would be helpful for communication data. We decided to display transcripts on a GPS map showing Soldier movements so that OC/Ts have the appropriate context for communications at any given time. Visualizations showcasing these features were created using Soldier communications data during an iteration of BD2A in the SUPRA study and Google Maps images of Fort Johnson. See Figure 3 for an example of a visualization showcasing communication data used in this study. Figure 3 shows a scrolling transcript on the left (1) while the squad movement animation plays in the panel to the right, just as in Figure 2. Audio recordings of squad members can be heard (2). At the same time, the animation plays, and the transcribed communication is highlighted on the left panel and appears briefly as a text bubble adjacent to the symbol of the person making the statement on the right (3).

Figure 3. Communication Visualization Concept provided to OC/Ts.

Firing Rate

OC/Ts expressed that they would like visualizations of firing rates over time and individual Soldier shot counts for firing rate data. They stated the rate of fire and ammo expenditure data could help to evaluate the extent to which a squad was a "well-tuned orchestra" during support by fire activities in BD2A. In addition, OC/Ts stated that a key part of their assessments of squad suppressive fire is how well they maintain proper firing rates at different times in

the drill. See below for an example visualization showcasing these features. This visualization depicts firing rate graphs showing squad fire over time (1), with menu options to toggle whose data is displayed (2). In addition, the graphs overlay thresholds that squad firing rates need to reach to maintain sustained and rapid fire (3). On the right side of the visualization, you can find options to toggle the appearance of firing rate levels or phases of the drill (4). These visualizations were created using Soldier firing rate data during an iteration of BD2A in the SUPRA study. Firing rate data was collected using IMU sensors mounted on each Soldier's rifle (see O'Donovan et al., 2023 for detailed data collection techniques). See Figure 4 for an example of a visualization showcasing firing rate data used in this study.

Figure 4. Firing Rate Visualization Concept provided to OC/Ts.

Shot Accuracy

For shot accuracy data, OC/Ts expressed that they would like to be able to see visualizations of enemy target hits and missed, fire accuracy data, visualizations showing details of kills and deaths (who killed who and when), and visualizations indicating where hits were taking place on Soldiers. Visualizations showcasing these features were created using fabricated Soldier shot accuracy data. See Figure 5 for an example of a shot accuracy visualization used in this study. This example showcases various metrics of shot accuracy (1), including vital and non-vital hits, near and far misses, and hits and kills probability metrics (2). OC/Ts mentioned that shot accuracy data would be more relevant for specific team members based on their roles and weapon systems. For example, a SAW (Squad Automatic Weapon) is intended for suppressive fire and may not result in as many kills or hits, but this does not mean the Soldier was not performing their role effectively. Because of this, OC/Ts need the ability to toggle whose metrics are displayed to increase the data's utility (3).

Figure 5. Shot Accuracy Visualization Concept provided to OC/Ts.

PHASE 2: OC/T DATA VISUALIZATION FEEDBACK

During Phase 2 of the study, researchers conducted semi-structured interviews with OC/Ts to solicit feedback on several data visualization concepts with varying amounts of data and different data sources to determine which concepts would be most useful to them and how to improve them for future use. Here, OC/Ts were asked questions regarding the likelihood that they would utilize the visualizations to augment their squad assessments or during an AAR session with their squads. OC/Ts additionally gave feedback regarding what they liked about each visualization and what they would like to see changed or introduced to make them more effective.

Methods

Interviews were conducted in person at JRTC and individually with OC/Ts. The interview started with an explanation of its goals and format, followed by a presentation of data visualization concepts. OC/Ts were asked to evaluate the concepts' usefulness for squad performance evaluation and AAR feedback with squads and provide feedback on improvements. OC/Ts were asked not to consider the feasibility of implementing and fielding these capabilities. Utility ratings were given on a four-point scale from Very Useful to Not Useful. The interview concluded with demographic questions, and OC/Ts were thanked and dismissed.

Participants

We interviewed 10 participants with prior experience assessing squads completing BD2A as an OC/T. Like Phase 1 interviews, these participants were current and former OC/Ts stationed at JRTC at Fort Johnson, LA. The sample consisted equally of individuals at the Sergeant First Class rank (E-7, n = 5) and Staff Sergeants (E-6, n = 5). The total estimated number of battle drills evaluated for each OC/T ranged from 2 to 500, with Staff Sergeants evaluating 28 battle drills on average in their career (M = 28, SD = 24) and Sergeants First Class evaluating 340 battle drills on average in their career (M = 340, SD = 156).

Results

Conclusions were derived from the OC/Ts' feedback on the concepts' utility for two functions: squad performance evaluation and AAR feedback. Performance evaluation relates to the Training and Evaluation Outline (T&EO) checklist items, and AAR feedback involves providing coaching on improving performance. While the same information can sometimes be used for both purposes, that is not always the case. Participants rated the utility of these concepts on a four-point scale, ranging from Very Useful to Not Useful. The quantitative data from these utility ratings were analyzed using descriptive statistics to determine the overall perceived usefulness of each concept for informing OC/T squad evaluations and augmenting AAR feedback separately. These ratings provided a straightforward numerical summary of the concepts' effectiveness as perceived by the OC/Ts. See Table 2 for an overview of the utility ratings.

Table 2. Percentage of Sample Rating the Utility of Visualizations for Evaluation and AAR Uses.

Concept	**Utility Ratings**			
	Very Useful	Moderately Useful	Somewhat Useful	Not Useful
GPS Positioning				
Evaluation	70%	30%	0%	0%
AAR	90%	10%	0%	0%
Communication				
Evaluation	80%	10%	10%	0%
AAR	90%	10%	0%	0%
Firing Rate				
Evaluation	70%	20%	0%	10%
AAR	70%	10%	20%	0%
Shot Accuracy				
Evaluation	80%	10%	10%	0%

AAR	80%	10%	10%	0%

In addition to giving utility ratings, researchers sought feedback and recommendations from OC/Ts on improving the presented concept visualizations to inform evaluations and provide better feedback. Table 3 summarizes each feature categorized by data type, the justification for its inclusion in concept visualizations, and additional OC/T recommendations for improving future visualizations.

Feature	Justification	OC/T Recommendation
GPS Positioning		
Satellite Map View	Satellite images clearly show roads, structures, tree lines, fire breaks, and other linear danger areas (LDAs) that can assist OC/T judgments of Soldier movement.	• Add an arrow pointing north to provide a cardinal reference point. • 3D maps would be useful for urban areas. • Need to ensure that maps are up to date. • Include topographic map options with contour lines to provide terrain elevation information and relief features.
Soldier Icons	Use standard Army icons for friendly and enemy forces.	• Show the Soldier position (e.g., prone, kneeling, standing). • Identify team leaders with different colors. • Indicate Soldiers' weapon system. • The ability to toggle enemy icons on/off to better train units without giving away enemy location.
Head Orientation	Helps OC/Ts evaluate squad security coverage.	• Ability to toggle features on and off for individual squad members. • Show the Soldier's accurate viewing distances with elevation considered. • Include the option to view muzzle orientation rather than head orientation.
Travel Path Visuals	Help visualize the squad's travel path and their spacing during movement.	• Have the ability only to see movement when Soldiers are engaged to reduce clutter. • Include the enemy travel path as an optional overlay to show where they were in relation to the squad. • Show elevation if they are in a building. Include a timer to show how long they are up and kneeling or in prone when engaged or bounding.
Toolbar	Include a toolbar that allows OC/Ts to toggle on and off different features, draw on still images, and take screenshots of an exercise.	• Ability to zoom in and out of the maps. • Ability to add notes and reference the T&EO checklist to align feedback to a squad's needs. • Include a protractor tool where angles of fire and movement can be assessed. • Include an overlay visualization to show weapon ranges and fields of fire.
Communication		
Playback	It is useful to see who is speaking and how the squad communicates quickly. This could help OC/Ts give feedback on communication discipline, especially for LACE (liquid, ammunition, casualties, equipment) reports.	• Include the ability to track non-verbal communication. • Include the ability to turn on and off audio of communications. • Highlight who is speaking on the GPS Solider icons.

Text Bubbles	It helps to show firing rate changes during key drill phases.	• It is helpful to isolate the "mad minute" and shift fire phases of BD2A, but otherwise, this isn't an essential visualization feature.
Speaker Filter	It helps to visualize how Soldiers follow the rate of fire they are ordered to by the team leader and reduces the mental math needed from OC/Ts to calculate firing rates in real-time.	• No additional recommendations.
Firing Rate		
Firing Rate Graphs	It helps to show OC/Ts a visual of squad fire control during an engagement.	• The firing rate can be used as an overlay window with the communications and map visuals to see it all together. • Add time scale marks for minutes, even if seconds are chosen as the scale. • Firing rate graph lines should look like heartbeats on an electrocardiogram (ECG), not smoothed out.
Drill Phase Filter	It helps to show firing rate changes during key drill phases.	• It is helpful to isolate the "mad minute" and shift fire phases of BD2A, but otherwise, this isn't an essential visualization feature.
Firing Rate Levels	It helps to visualize how Soldiers follow the rate of fire they are ordered to by the team leader and reduces the mental math needed from OC/Ts to calculate firing rates in real-time.	• No additional recommendations.
Shot Accuracy		
Shot Accuracy Data	In AARs, showing Soldiers their data will help them better understand areas of improvement. This data could breed healthy competition among squads as well.	• Ability to show how Soldiers performance changes over time to indicate training effectiveness. • It is useful to filter the type of shooting data collected, whether from MILES with blanks, live fire, or live rounds impacting paper, metal, or electronic targets. • Hits and misses may be sufficient for AAR feedback.

DISCUSSION

Our results reflected strong support by the OC/Ts we interviewed for digital representations of squad performance. In the interviews, their support seemed to stem from two major capability gaps they experienced. First, the scope of training they typically observe makes it challenging to observe all squad activities and behaviors directly. Even at the squad level, trainees are distributed over a large enough area that an OC/T may, at best, only be able to observe a single fire team at a time. At other times, they cannot observe training due to physical obstructions like walls or vegetation, even if they are near the trainees. The second reason they supported these representations is that it frees them from relying on their (and the trainees') memories of events. A common sentiment was that they liked providing objective data that would verify what occurred instead of relying on (sometimes conflicting) subjective opinions of what transpired. They felt this would set the stage for more constructive discussions of unit performance.

Several issues arose regarding the data visualizations we provided that will likely need additional research and evaluation to be resolved. For example, OC/Ts had mixed opinions on whether head or weapon orientation was more useful. Some felt that head orientation was more important because it showed where they were looking, while others felt that weapon orientation was more important because it showed where they were truly providing security. Of

course, one option is to try to provide both data points; this will require additional sensors and data processing, which have inherent costs that must be considered. Similarly, there was disagreement on whether audio recordings are needed for every trainee or only for leaders. Again, it is possible to record from everyone (more recordings and more data) but allows leaders to filter just leaders, but this also imposes additional costs for sensors and increases the data processing load that needs to be weighed.

CONCLUSION

In the context of Army training and evaluation, the discussion surrounding the implementation of a data-informed approach for squad assessment holds significant implications for enhancing training effectiveness and overall mission readiness. By embracing this approach, the Army stands to benefit from a more comprehensive and nuanced understanding of squad performance, informed by a synthesis of objective data and expert judgment. The integration of data analytics techniques offers the promise of unlocking insights that may not be readily apparent through traditional assessment methods, providing OC/Ts with a more granular understanding of individual and squad performance. By leveraging the expertise of OC/Ts in combination with this objective data, squad assessments can be augmented with contextual insights and perspectives from years of experience, ensuring that the assessment process remains accurate and balanced. However, the process by which this data-informed approach is developed must involve iterative discussions with experienced OC/Ts, as shown in the process described in the present study.

The present study offers a case study demonstrating a process researchers can use to develop a data-informed approach to enhance OC/T assessments. Researchers developed a holistic understanding of squad performance through SUPRA sensor-based measures and in-depth interviews with OC/Ts. Understanding the end-user enabled us to conceptualize data visualizations to enhance squad assessments by addressing OC/Ts' current limitations from their perspective. We further interviewed OC/Ts for their feedback regarding the concepts, taking a user-centered design approach towards data visualization development.

The synthesis of objective data and subjective OC/T knowledge enhances the accuracy of performance assessments and promotes a culture of continuous improvement and adaptability within Army training and education. However, challenges remain in implementing this type of approach. In developing technologies and tools for assessment, collaboration between researchers and OC/Ts is crucial to ensure that end-user perspectives are integrated. By incorporating OC/Ts' valuable insights alongside data-driven methodologies, a more comprehensive and effective data-informed approach to assessment can be achieved. This collaborative process enhances the relevance and usability of assessment tools and ensures they align closely with the practical needs and challenges faced by trainers in the field. Following the method and approach outlined in this paper can serve as a blueprint for researchers aiming to integrate OC/T perspectives into developing assessment aids. By systematically engaging with OC/Ts throughout the research and development phases, researchers can ensure that their solutions are not only data-driven but also informed by the practical insights and needs of those who will use them in the real world. This collaborative approach enhances the applicability and effectiveness of assessment methodologies, fostering a more robust and insightful framework for evaluating squad performance in military contexts.

ACKNOWLEDGMENTS

This research was sponsored by the U.S. Army Combat Capabilities Development Command - Soldier Center (DEVCOM 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 DEVCOM SC or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes, notwithstanding any copyright notation.

AUTHORSHIP STATEMENT

All authors were involved in research planning, data collection, data analysis, report writing, and editing.

REFERENCES

King, M., Sikorski, E., & Goodwin, G. (2023). Diagnosing critical incidents during army battle drills. *Proceedings of the 14th MODSIM World*, Norfolk, VA.

Morgulev, E., Azar, O. H., & Lidor, R. (2018). Sports analytics and the big-data era. *International Journal of Data Science and Analytics*, 5, 213-222.

O'Donovan, M. P., Hancock, C. L., Coyne, M. E., Racicot, K., & Goodwin, G. A. (2023). Assessing the impact of dismounted infantry small unit proficiency on quantitative measures of collective military performance part 1: Recommended test methodologies. *Defense Technical Information Center (DTIC)*, TR23-013.

Sinatra, A., Graesser, A. C., Hu, X., Goldberg, B., & Hampton, A. J. (Eds.). (2020). Design recommendations for intelligent tutoring systems: Volume 8- data visualization. U.S. Army Combat Capabilities Development Command–Soldier Center.

Stone, A. (2017, December 14). Virtual battle spaces can help rapid real-world tactical decisions. C4ISRNet. https://www.c4isrnet.com/it-networks/2017/12/14/virtual-battle-spaces-can-help-rapid-real-world-tactical-decisions/

Zagata, K., & Medyńska-Gulij, B. (2023). Mini-map design features as a navigation aid in the virtual geographical space based on video games. *ISPRS International Journal of Geo-Information*, 12(*2*), 58.