Research Product 2020-02

Soldier Performance and Talent Assessment: Mobile Application Development

Victor J. Ingurgio
U.S. Army Research Institute

David R. James
Gustavo Silva
Brandon Mitchell
Anthony Scoppa
Northrop Grumman Corporation

January 2020

United States Army Research Institute for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved:

MICHELLE ZBYLUT, Ph.D.
Director

Research accomplished under contract
for the Department of the Army by

Northrop Grumman Corporation

Technical review by

Dr. Thomas Rhett Graves, U.S. Army Research Institute

Dr. Melissa Stuhlman, U.S. Army Research Institute

NOTICES

DISTRIBUTION: This Research Product has been submitted to the Defense Information Technical Center (DTIC). Address correspondence to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-ZXM, 6000 6th Street (Building 1464 / Mail Stop: 5610), Fort Belvoir, VA 22060-5610.

FINAL DISPOSITION: Destroy this Research Report when it is no longer needed. Do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Research Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
January 2020	Final	December 2017 – July 2019

4. TITLE AND SUBTITLE

Soldier Performance and Talent Assessment: Mobile Application Development for

5a. CONTRACT NUMBER
W911NF-18-F-0005

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
622785

6. AUTHOR(S)
Victor J. Ingurgio;
David R. James, Gustavo Silva, Brandon Mitchell, Anthony Scoppa

5d. PROJECT NUMBER
A790

5e. TASK NUMBER
433

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral & Social Sciences
6000 6th Street (Building 1464/Mail Stop 5610), Fort Belvoir, VA 22060-5610; and
Northrop Grumman Systems Corp.
3565 Macon Rd., Columbus, GA 31907

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral & Social Sciences
6000 6th Street, Building 1464
Fort Belvoir, VA 22060

10. SPONSOR/MONITOR'S ACRONYM(S)
ARI

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Research Product 2020-02

12. DISTRIBUTION/AVAILABILITY STATEMENT:
Distribution Statement A: Approved for public release; distribution is unlimited. Final Disposition: Destroy this Research Report when it is no longer needed. Do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

13. SUPPLEMENTARY NOTES
ARI Research POC: Dr. Victor J. Ingurgio, Fort Benning Research Unit

14. ABSTRACT

The purpose of this research was to support Army leaders with tools that provide them with a better understanding of Soldier performance during critical assessment events. This research developed a mobile assessment tool that is flexible, scalable, and editable and that can support differing Army events, competitions, and assessments of individual and collective task training and testing. The model for the solution was the Expert Infantryman Badge (EIB). The mobile assessment tool allows cadre to rate over 1,000 EIB Candidates, using tablets to access digital rubrics to enter Candidate scores. These scores are digitally transferred in (near) real-time to a tracking application and displayed on a data analytics dashboard in the tactical operations center, providing leaders with an instantaneous overview of EIB candidates' performance. This system was tested and the user interface was validated during the train up week of multiple EIBs. The results from the data analytics can be used by Army leaders to focus their training and ultimately increase the proficiency of Soldiers and future Army leaders.

15. SUBJECT TERMS
mobile technology, mobile applications, digital assessment, performance data analytics, real-time performance data capture, talent management and assessment, Expert Infantryman Badge, Soldier tasks, training and evaluation outline (T&EO)

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Dr. Jennifer S. Tucker
a. REPORT	b. ABSTRACT	c. THIS PAGE	Unlimited	73	19b. TELEPHONE NUMBER
Unclassified	Unclassified	Unclassified	Unclassified		706-545-2490

ARI Research Product 2020-02

Soldier Performance and Talent Assessment: Mobile Application Development

Victor J. Ingurgio
U.S. Army Research Institute

David R. James
Gustavo Silva
Brandon Mitchell
Anthony Scoppa
Northrop Grumman Corporation

Fort Benning Research Unit
Jennifer S. Tucker, Chief

January 2020

ACKNOWLEDGEMENTS

The authors are grateful for the collaboration and support of the U.S. Army Infantry School, Fort Benning, GA, and the EIB Test Managers, for their guidance, feedback, and support that made this product possible. The authors would also like to acknowledge the Soldiers and leaders at Forts Benning, Hood, Polk, and Joint Base Lewis-McChord for their professionalism and support during our data collection efforts.

MOBILE APPLICATION DEVELOPMENT FOR SOLDIER PERFORMANCE AND TALENT ASSESSMENT

CONTENTS

CONTENTS (cont.)

Page

TABLES

LIST OF FIGURES

CONTENTS (cont.)

Page

Mobile Application Development for Soldier Performance and Talent Assessment

Background

Individual and unit mission readiness is a top priority in the U.S. Army according to General Mark Milley, the Army Chief of Staff, in the *Army Readiness Guidance, Calendar Year 2016-17* (Milley, 2016). His handwritten note at the bottom of this document left no doubt, "Readiness is # 1…and there is no other # 1." This sentiment has been echoed down the chain-of-command throughout the Army. However, readiness is not a permanent condition. To be assessed as ready, individuals and units must demonstrate performance that meets a standard and these results must be available for reporting and monitoring and include actions to sustain or inform remediation. The *Talent Management Concept of Operations for Force 2025 and Beyond* emphasizes that talent management is a required capability that impacts readiness (U.S. Army Combined Arms Center [CAC], 2015, p. IV). This concept also implies that Human Capital-Big Data, people analytics, and profiling high performers are important aspects to consider for talent management systems. However, no talent management system currently exists that facilitates the linkage between training performance and talent management.

Almost every activity in the Army relates to a task or subtask that can be trained, practiced, tested or evaluated, and measured against some standard. Regardless of the training audience or venue, each of these assessments share common features. First, there is a standard of performance. A training and evaluation outline (T&EO) establishes that standard. Second, there is an evaluator or team of evaluators. The evaluators may be unit leaders, a designated evaluator, or an Observer/Controller Trainer at a major training center. Lastly, there is an evaluation schema and method of capturing the assessment.

In many situations, these assessments are paper-based and generate mountains of files and forms, which in turn create data storage and data entry issues. Some assessments create a massive administrative burden and backlog. During special occurrence competitions, commanders, evaluators, event administrators, and competitors, all eager for timely and accurate information, must wait for the backlog to be reduced and the results to be posted. The Expert Infantryman Badge (EIB) testing is one such example of an Army assessment that occurs repeatedly throughout the year and impacts numerous infantry organizations and Soldiers and effects overall unit readiness. The assessment is a test of Soldier's individual task proficiency using a T&EO that clearly lays out the task, conditions, and standards and is assessed using performance measures, steps, and sub-steps. Both the previous version of the EIB manual (circa 2016) and the current version (2018/2019) generate massive amounts of data as each Soldier's performance is documented on a paper T&EO. Further, the analysis of performance is conducted post-hoc and limited to the task level.

The Expert Infantryman Badge

The EIB, established by Army Chief of Staff George C. Marshall in 1944 (*History of the Expert Infantryman Badge*, Dec 3, 2018), is awarded to recognize Infantry Soldiers who have demonstrated a mastery of critical tasks (Department of the Army [DA], 2019, p.7). The EIB has

changed over the decades as new equipment was fielded and critical tasks were updated. The current (2019) EIB is comprised of four events, the EIB Physical Fitness Assessment (EPFA), a Land Navigation Course (Day/Night), Individual Task testing stations, a 12-mile Foot March and Final Event, which encompass a total of 46 individual tasks. Soldiers who are eligible must complete all events to standard to be awarded the EIB. Attrition is high. Since the inception of the new EIB standard in 2018, the success (award) rate as of July 2019 is 16% (1,345/8,597) (*Expert Infantry Community*, July 15, 2019).

According to the U.S. Army Maneuver Center of Excellence, the EIB is a talent discriminator when Infantry Soldiers are considered for selection for promotion to the next higher rank. For example, from 2012 to 2018 Sergeant First Class (SFC) centralized promotion board analysis identifies that 78% (4,549/5,864) of Staff Sergeants (SSG) in the Career Management Field 11 (11B Infantrymen and 11C Mortar men) selected for promotion to SFC had been awarded their EIB (*Board Analysis*, July 9, 2019). The 2018 board analysis further defines the "best qualified SSG" as:

> An exceptional SSG that is determined to be best qualified for promotion will have at least 24 months' rated time in an authorized leadership position; **will have earned the EIB**; will have scored at least 270 on the Army Physical Fitness Test (APFT); will have completed some college classes; will have graduated from at least five MOS-enhancing courses; will have graduated from either Bradley Master Gunner Course, Battle Staff Non-Commissioned Officer (NCO) Course, or the Ranger Course; and will have served in both priority Operational Force and priority Generating Force assignments. (p.1)

The EIB proponent is the United States Army Infantry School (USAIS) at Fort Benning, GA. The USAIS maintains, modifies, and publishes the doctrinal reference for the EIB. USAIS Pamphlet 350-6 *Expert Infantryman Badge* establishes:

> ...policies, procedures, and standards for the Expert Infantryman Badge (EIB). The EIB test measures a Soldier's physical fitness and ability to perform to standards of excellence in a broad spectrum of critical Infantry skills. Detailed instructions in this pamphlet ensure Army-wide uniformity. EIB training and testing is intended to be rigorous, mission-focused, and conducted under realistic conditions. (DA, 2019, p.6)

Each Infantry Unit's EIB is validated by the USAIS EIB Test Manager but executed by the testing unit. According to the U.S. Army Maneuver Center of Excellence (*Board Analysis*, July 9, 2019), between September 2018 and Mar 2019, 11 Infantry Brigades or Separate Battalions executed an EIB with a total of 6,448 Soldiers eligible for testing. As presented above, of these eligible Soldiers, 16% (1,019) were awarded their EIB. From a talent assessment perspective, these awardees are the first to be considered when identifying Soldiers for positions of increased responsibility. From an analysis perspective, the question should concern why 84% of the Soldiers failed, and where should Infantry leaders focus their training to increase proficiency and identify future Infantry leaders.

Current analog processes enable the EIB Test Manager to drill-down to the individual task title level. For example, during one Unit's EIB, the first time "GO" rate for the task *Adjust*

Indirect Fire was 68%, however, the ability to identify specifically at what sub-task/step within the task the 32% of failures occurred is non-existent. Moreover, this macro level of analysis is collated after the fact, at the completion of testing, and only provides leaders with the knowledge that additional practice is needed on the subject task.

Problem Definition

As with the EIB, it is often the case that Army testing events occur in a field setting where important evaluations are taken that reflect the key elements of an individual or team skill. Evaluators have had limited ways to capture such important moments, mostly via pen and paper grading sheets; however, with the development of mobile technology, the options have expanded considerably. This effort was undertaken to design and develop a technological solution to meet a broad spectrum of challenges that include connectivity, flexibility, security, and user-friendly interfaces, displays, and analytics.

Research Objective

The objective of this research was to develop a single technology solution that was flexible, scalable, and editable and could support differing Army events, competitions, and assessments of individual and collective task training and testing. The model for the solution was the EIB. To facilitate the research objective, our researchers were comprised of an Army Research Institute for the Behavioral and Social Sciences (ARI) Senior Research Psychologist, a retired Infantry Sergeant Major with over 20 years of EIB experience, a senior software architect, a digital transformation and product design lead, and a software engineer.

Research Solution

The research described in the following sections enabled the researchers to develop a single solution for multiple Army individual task events/competitions. The Soldier Performance Application for Readiness and Talent Assessment (SPARTA) is an enterprise platform for hosting mobile assessment applications. The SPARTA platform (Figure 1) is a flexible, scalable, and secure solution.

Figure 1. SPARTA splash screen. SPARTA was designed as an iOS discoverable and user friendly application. Design principles enabled EIB evaluators to easily navigate between tasks and evaluation criteria.

Structure of the Report

The results of this research are presented in three sections. The first section discusses gathering requirements for the EIB exemplar test, the second section presents the development process for and details of SPARTA, and the third section describes the piloting and subsequent refinement of SPARTA.

Requirements Gathering

Our researchers executed a four-phase approach to this task with the overarching goal of identifying the digital requirements to support the creation, assessment, and analysis of an EIB. We initially gathered requirements by reviewing USAIS 350-6 to determine the breadth and width of each EIB event, along with established business rules. We then met with the USAIS EIB Test Manager to determine roles within support and execution of an EIB and to clarify questions from the doctrinal review. Next, we conducted multiple on-site observations of Infantry units conducting EIBs to confirm assessment workflow and, lastly, discussed with U.S. Army Infantry School (USAIS) leaders, the EIB Proponent, the level of analysis required.

USAIS Pamphlet 350-6 Expert Infantryman Badge

The USAIS, as the proponent for the EIB, publishes the EIB Manual (USAIS 350-6) in order to standardize EIB testing across Infantry Units. In addition to the publication, the EIB Test Manager created an EIB website[1] as a one-stop location for all information pertaining to the EIB test. We used both resources during our initial review to identify design and content requirements.

Our initial review was of the August 2016 version of USAIS Pamphlet 350-6; however, three months into the research effort, the USAIS conducted a wide-ranging revision of the EIB test and published the revised version in 2018. While the events and method of evaluation (T&EO) stayed the same, the tasks and sub-tasks within the events changed, with the majority of the changes applied to the Individual Testing Stations. Additionally, and more importantly, the rules to eliminate a Soldier from testing changed. Table 1 compares the 2016 events to the 2018 events with the changes in bold text.

[1] The EIB website can be found at https://www.benning.army.mil/Infantry/EIB/. The site provides updates to the doctrinal publication, steps for Infantry Units to request approval to test, and the appropriate forms and Microsoft™ Excel spreadsheets to use to collate results.

Table 1

Comparison of 2016 EIB and 2018 EIB

	2016	2018
Army Physical Fitness Test **renamed as the EIB Physical Fitness Assessment**	Push-ups Sit ups 2-mile run Age and gender specific scoring	Push ups Sit ups **4-mile run** **Age and gender neutral**
Land Navigation Day/Night	Self-correcting course	**Non self-correcting course**
Individual Task Testing	3 Lanes – Medical (M), Patrol (P), and Weapons (W) 10 Stations per lane 1 Task per station	3 Lanes – Medical (M), Patrol (P), and Weapons (W) 10 Stations per lane **2 Tasks per station for W1-W4** **3 Tasks for station W5** 1 Task each for the remaining 25 stations
Unit Options	Units must select the 30 tasks from 45 total tasks	**No tasks are optional** **Units can choose the weapon system within certain tasks based on assigned/available equipment**
Rules	Cumulative strikes across individual testing lanes – third NO GO and you're out	Cumulative strikes **within** individual testing lanes – **second** NO GO from one lane and you're out **Strikes start over when changing lanes**
12-Mile Foot March	12-Miles with 35 pound pack	12-Miles with 35 pound pack
Final Event	Objective Bull medical tasks	**Assemble and Disassemble an M4/M16 Rifle**

Note: W1 through W5 are the first five stations in the Weapons lane.

The changes to the EIB, while seemingly not significant, did impact the design of SPARTA. Initially, we began designing a relational database model based on the 2016 review and method of evaluation, the T&EO and recognized that we had to modify the model to accommodate these changes. The most significant changes we identified were that (a) SPARTA would have to be flexible in its ability to handle not just this change, but any future changes, (b) evaluators for W1-W5 would now evaluate two or more tasks sequentially with a separate workflow based on Soldiers failing the second task within the sequence, (c) groups of tasks, e.g., W1-W5, would have cascading effects to our database model, the mobile model, and application

user interface (UI), and (d) the differing rules for eliminating a Soldier would result in the development of a separate rules table or engine. The review of the 2018 EIB Manual draft document required a meeting with the EIB Test Managers to confirm the changes and clarify our questions.

USAIS EIB Test Managers

The USAIS has appointed two senior NCOs as the EIB Test Managers. The Test Managers are responsible for maintaining the EIB Manual, recommending and making any changes to it, processing a unit's request to conduct an EIB, validating all EIB test sites, and issuing official orders and certificates to Soldiers who are awarded the EIB. We held a meeting with the EIB Test Managers to discuss theirs and the unit's roles in executing an EIB and to clarify the administrative changes to the 2018 EIB Manual.

EIB roles. Through discussions with the EIB Test Managers, we identified four distinct roles that would have to be considered for inclusion into SPARTA. Those roles were the test creator, the test user, the test evaluator, and the analyses users. Each role would have specific administrative permissions and would interact with SPARTA in unique ways.

Test creator. We identified that there should be one test creator to create the "gold copy" of a test/competition/event. In this case, it would be the EIB Test Manager. They would create the EIB in total, e.g., naming the events, creating the business rules, establishing standards, providing options, creating the T&EOs for the tasks. The EIB Test Manager would create the test based on the source document (USAIS PAM 350-6) and they would be the single point for all changes.

Test user. We identified that there could be many test users, all able to create copies or "instances"[2] of the test for their own use, each instance a unique entity. In this case, it would be the EIB unit administrator from the requesting unit. The EIB unit administrator would have access to the test to select from any of the options created by the EIB Test Manager. For example, one EIB test user would create an instance of the EIB test and select the M9 Pistol option from the two options available for task W2, either the M9 or M17 Pistol; whereas, another EIB test user could select the M17 pistol option for their instance. The selection of these options is based on the requesting unit's authorized equipment and options would currently apply to multiple tasks (W1, W2, W10, and P9). There are no options available for the remaining 26 tasks.

Test evaluator. We identified that there would be many test evaluators per instance. For example, at a minimum, there are two test evaluators for each task assessed during the Individual Task testing totaling 60 test evaluators per instance—i.e., 30 tasks (3 lanes at each of 10 Stations) multiplied by 2 test evaluators. Further, there would be 120 test evaluators during two simultaneous EIB test instances.

[2] We elected to use the word 'instance' as each copy of the test is a single use, of which there could be multiple instances occurring at the same time (i.e., one EIB CONUS and one EIB OCONUS concurrently testing on the same dates).

***Analyses user*.** We identified that there could be many analyses users. For example, during an EIB event, many Army leaders are interested in their Soldiers' individual task proficiency and are frequently on-site, usually at the Tactical Operations Center (TOC), looking for status information. Additionally, each unit that conducts an EIB is required to provide performance results to the EIB Test Manager for collation, analyses, and presentation within State of the Infantry briefings.

EIB 2018 rules clarification. The main points of discussion and clarification with the EIB Test Managers related to the business rules regarding candidate elimination. Within the EIB, Soldiers who fail/NO-GO an event or task are subject to elimination from the EIB test. We have identified multiple rules for events and tasks and clarification is warranted.

The current USAIS EIB reference states that Candidates [read Soldiers] are eliminated from EIB under the following criteria. First, they can fail to pass any big event (EPFA, Day/Night Land Navigation, 12-Mile Foot March, or Final Event). Second, they can receive more than one NO-GO within one lane over the course of the 30 Individual Testing Stations (Soldiers are allowed three total NO-GOs/retests; one per lane, per day). Third, they can fail to return for a retest within one hour. Finally, they can break a rule of the competition, i.e., any unsafe act or integrity violation that is clearly defined and briefed by the EIB Board prior to training (DA, 2019, p. 11). Additionally, the reference also states that "If a task has multiple parts (W1-W5), the Soldier does not have to retest the portion(s) they have already passed but will restart at the beginning of the part they failed" (DA, 2019, p.16).

In clarifying each of these rules, the EIB Test Managers reiterated that each big event (e.g., EPFA) or each task (e.g., W2 – Pistol and Shotgun) must have a method of eliminating the Soldier from the test. However, they stated that there are different rules for what the EIB manual calls the "big events" than for the Individual Task testing. The big events are a one-and-done elimination with no retest opportunity. For example, the EPFA comprises three tasks, the push-ups, sit-ups, and 4-mile run, which are completed in this order. If a Soldier fails to meet the standard for the push-ups (49 push-ups in 2-minutes), then the Soldier is eliminated from the EIB event with no retest. This elimination criterion applies to each of the tasks within the big events. In contrast, during Individual Task testing, the Soldier gets to retest (within one hour) on a task they have failed/received a NO-GO, but only for two tasks in each lane (typically, one lane per day).

We initially found these business rules regarding elimination confusing, but after talking with the EIB Test Managers and gaining a thorough understanding, we determined that we would have to add a test rule table into the SPARTA platform to accommodate these rules, plus additional rules for any other similar tests. Discussing the EIB with the EIB Test Managers provided clarity; however, seeing the EIB being executed would provide a contextual understanding and could highlight requirements yet to be discovered. With this in mind, we made a request of the EIB Test Managers to observe an EIB during both testing and train-up days.

EIB Observations

Our researchers coordinated with the EIB Test Manager to schedule EIB observations. We were allowed to observe two EIBs. These observations provided context as to how an EIB is conducted. The primary goal of our initial observation was to observe and document the workflow related to the assessment of a single EIB task. We wanted to follow the data. The goal of the observations at the second EIB was to confirm the documented workflow using a prototype application. We focused our observations on just the Individual Task test lanes.

Initial EIB observations. Our researchers observed an Armored Brigade Combat Team conducting EIB training[3]. The unit conducting the EIB was using the 2016 EIB standard as the new standard had yet to be released. By happenstance, the EIB training that our researchers were able to observe was conducted using the cradle-to-grave concept outlined in USAIS 350-6. This concept is employed when the testing unit cannot provide the 100+ verified EIB badge-holders required to execute a traditional EIB (DA, 2019, p.9) and could offer a different workflow. While the set-up of the training site and the conduct of training differed, i.e., only one lane and its 10 stations were set-up each day, the assessment process was the same as the traditional method of conducting an EIB. The traditional EIB training method includes all three lanes (30 stations) set-up every day.

We identified that the assessment process for a single EIB task starts with instructions to the Soldier and follows one of four paths: GO, NO-GO, Protest, or Safety Violation. As the Soldier performs the task, the evaluator follows one of these four paths, concluding with documentation of the results.

A single EIB task (station) requires the Soldier to perform a hands-on procedural task that follows a prescribed set of steps. The T&EO for each task will list these steps. The workflow begins when the Soldier arrives at the testing station. The evaluator documents the Soldiers' rank and name, reads the standardized "Instructions to the Soldiers" (task, conditions, standards, and any other administrative information) to each Soldier, and then observes the Soldier performing the steps. The EIB badge-hold evaluator has previously been validated as a SME for the task they are evaluating.

If the Soldier successfully performs all steps correctly, and for most tasks, in the prescribed order in the prescribed amount of time, the evaluator will give the Soldier a GO. If the Soldier does not perform the steps correctly, performs the task steps out of order or exceeds the time allowed, the evaluator will give the Soldier a NO-GO. The Soldier would then return to the holding/practice area for additional time to practice. These are the two most common paths. The Protest path is an off-shoot of a NO-GO where the Soldier protests a NO-GO result and an adjudication process begins that ultimately ends with either a NO-GO (and a task re-try within 60 minutes) or a GO (retest with no strikes). The Safety (or Integrity) Violation path is related to the task whereby the evaluator gives the Soldier a NO-GO for performing an unsafe act (e.g., removing the buffer mechanism of a 240B machine gun with the bolt and operating rod charged to the rear), and the Soldier is immediately eliminated from testing.

[3] EIB training is conducted one or two weeks prior to testing where Soldiers practice all tasks under test conditions.

At the completion of the task, the evaluator records the result and repeats the process with the next Soldier. If the Soldier passed the task, he or she will move on to the next station; if the Soldier fails the task, he or she is provided with an opportunity to return to the holding/practice area and must re-test within 60 minutes. If the same Soldier fails the task the second time, or fails to report within the 60-minute window, he or she is eliminated from the EIB and escorted off the testing site.

We observed multiple tasks being conducted and assessed to determine if each evaluator used the same workflow. We determined that with minor changes to the instructions to the Soldier, each evaluator did follow the same workflow. This workflow became the basis of the user interface (UI) design for the prototype application.

However, with the change in EIB standards from 2016 to 2018, as depicted in Table 1, we had to re-think the workflow for a station that was now assessing two or more sequential tasks (e.g., W2 - Pistol and Shotgun). We revisited with the EIB Task Managers to discuss how the workflow would be affected. They described a different workflow than would have been previously applied. For example, if the Soldier passes the first task and fails the second or subsequent task, then when re-testing, the Soldier would now restart at the beginning of the task he or she failed, without having to re-do the one passed.

We incorporated the workflow – select a Soldier, read instructions, evaluate, record results – into a prototype EIB evaluator mobile application in preparation for our next opportunity to observe an EIB in action.

Prototype EIB evaluator Application. At this point in our requirements gathering process, we determined that we had enough information to develop a prototype EIB evaluator application (Figure 2). We intended to place the prototype in evaluator's hands as a means of gathering user experience (UX) feedback and to confirm the assessment workflow.

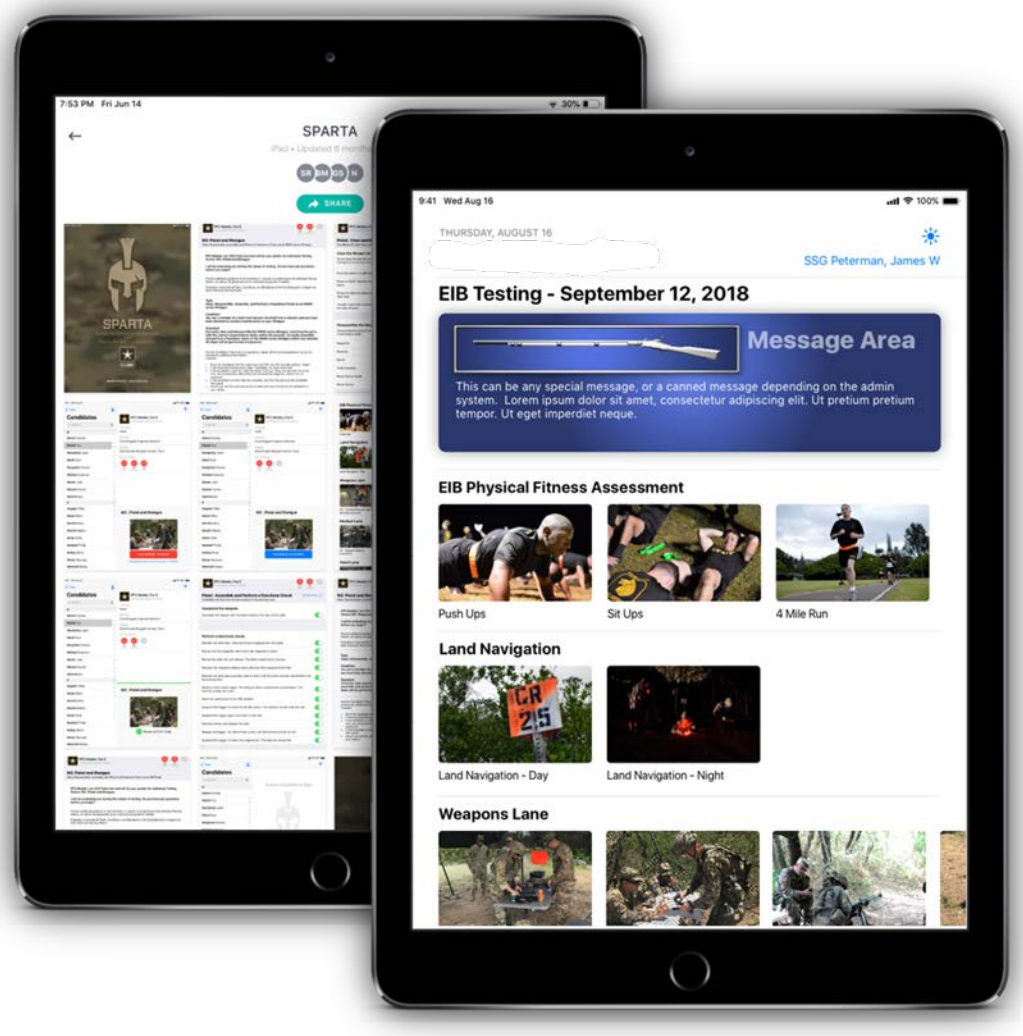


Figure 2. Prototype application. The prototype was developed using collaborative software to gain UX feedback from EIB evaluators.

We developed the prototype using InvisionApp™[4] product design, workflow, and collaborative software that enabled our designer to create the look and feel of an iOS application that could be displayed on a standard 9.7-inch iPad without coding the application. The software touch points enabled users to navigate through the application workflow and either provide verbal or text-based feedback specific to each screen. The value of using this software was the ability to make faster UI changes based on feedback without the added expense of re-coding the application.

The station we replicated within the prototype was Station W2 – Pistol and Shotgun that required evaluators to assess two sequential tasks. We selected this station as it was one of the significant changes between EIB 2016 and 2018 that had yet to be implemented, and we needed to confirm the assessment workflow. The prototype included the workflow for evaluators to login to the application, select the task group to be assessed (their Lane/Station), select the

[4] For more information visit InvisionApp™ at https://www.invisionapp.com/

Soldier to evaluate, evaluate the first task, evaluate the second task, and send the results to the server in real-time.

We demonstrated the prototype first to the EIB Test Managers to gain their feedback and made minor modifications with a quick turnaround. For example, the Test Managers requested a change to the task timer (change from hh:mm:ss to just mm:ss), the addition of the Protest button to more clearly indicate the process, and wording changes throughout the UI. Our designer quickly made the changes, and we re-presented the prototype to gain approval. Once we had gained overall approval of the prototype, we next planned for the opportunity to put the prototype in the hands of EIB evaluators at the second EIB observation to gain end-user feedback and also to confirm the designed workflow.

Second EIB observations. The purpose of these observations was two-fold. First, we needed to observe the piloting of the new EIB standards; second, we needed to confirm the new task evaluation workflow.

The USAIS had previously coordinated with an Infantry Brigade to be the first unit to execute an EIB with the new standards. Our researchers were invited to observe the piloting of the new standards as any modifications would affect the design and implementation of the SPARTA mobile application. One area that was ripe for modification was the time standards associated with the new tasks. Initially, the time standards had been determined by an EIB working group guided by the USAIS. The USAIS and EIB Test Managers expressed interest in collecting actual Soldier performance time for each task in order to validate and adjust the time standards as necessary. Our researchers offered to assist in this endeavor while at the same time observing the new standards being implemented.

Our researchers conducted the training observations while collecting the time validation data. We were afforded the opportunity to observe the workflow for three out of the five events during the unit's train-up. We observed the EPFA, the day portion of the Land Navigation, and the Individual Testing Stations. The remaining two events, the 12-Mile Foot March and the Final Event, were not conducted during the training period.

We identified that while the EPFA and the Land Navigation did not use a T&EO to evaluate the Soldiers, the two events followed the task-condition-standard format, had specific scoring criteria, and Soldiers were assessed as having either passed (GO) or failed (NO-GO). We also confirmed that the business rule for elimination for these two events differed from the Individual Testing Stations in that Soldiers were not afforded the opportunity to re-test and a single NO-GO would eliminate them from testing. In conversation with an EIB evaluator, this was also confirmed as the rule that applied to the 12-Mile Foot March and Final Event.

We next observed the Individual Testing Stations. One researcher was assigned to each lane (Medical, Patrol, and Weapons) to observe the assessment workflow and collect task performance time. The researcher with the most EIB experience (20+ years) observed the Weapons lane as this was where the new task workflow (multiple tasks) would be implemented.

Researchers utilized a Time Data Collection Sheet (Appendix C) to document the performance time for at least 10 Soldiers per station for each of the 30 stations. Time results were provided to the EIB Test Manager with a plan to collect more data from additional EIBs. Our primary reason was to observe the new EIB standards being implemented and to confirm the new multi-task workflow. We identified no change to the assessment workflow on the 20 stations being conducted on the Medical and Patrol lanes. The prototype application depicting the multi-task workflow was reviewed by the Station Non-Commissioned Officer in Charge (NCOIC) and the two evaluators assigned to Station W2 - Pistol and Shotgun. All three confirmed that the workflow matched what they were implementing during the EIB, with one exception. The workflow design included the timer function designed for the task as we had originally been provided. However, on the day of execution, the Brigade chain of command had made an EIB Test Manager approved change to the time standard from 45 to 60 seconds. This change was duly noted.

Analysis Requirements

In our discussions with the EIB Test Managers and with Army leaders during data collection efforts, we identified their desire to analyze the data at various levels. At the test site, we identified two analysis consumers, the TOC NCOIC and the Army leaders whose Soldiers were being tested. Both described interest in a "big picture" level of analysis that is provided currently by the EIB Unit Tracker Microsoft™ Excel spreadsheet. Further, the EIB Test Managers were interested in the more detailed causes of failure (i.e., time, safety, integrity, and performance steps or sub-steps).

Based on the recommendations and information provided by all stakeholders, and with access to the EIB Unit Tracker Microsoft™ Excel spreadsheet, we felt that we had gathered the necessary requirements to begin the development of the SPARTA mobile application and platform. The following section describes our development process.

SPARTA Development

Field Manual (FM) 7-0 *Train to win in a Complex World* is the U.S. Army's source document for training units. It introduces training concepts and processes, and describes how to plan, prepare, execute, and assess each training event (DA, 2016, p. vii). The SPARTA platform was designed to facilitate the assessment of individual training. The T&EO used to measure observed task proficiency, was the model for the SPARTA platform. The T&EO is the Army's source for individual and collective task training standards. The T&EO consists of the major procedures (steps or actions) a unit or individual must accomplish to perform a task to standard (DA, 2016, p. B-1).

Training and Evaluation Outlines are available for download from four approved Army sources, the Digital Training Management System (DTMS), the Army Training Network (ATN), the Combined Arms Training Strategy (CATS), and the Central Army Registry (CAR). The T&EOs are available in either an Adobe™ portable document format (pdf) or a Microsoft™ Word document. In either format, the Army leader or task evaluator will complete the evaluation, typically using paper and pencil, and are required to collate the results for reporting.

Completed T&EOs form the backbone of bottom-up feedback that company commanders and First Sergeants review at the weekly training meeting. The T&EOs provide the commander the necessary objective evaluations to assess unit training proficiency and ultimately to assess training readiness. (DA, 2016, p. B-1)

Overview

The SPARTA platform (Figure 3 depicts the high-level SPARTA architecture) was designed to provide leaders and evaluators with a digital T&EO that would facilitate and streamline evaluations and collation of individual task proficiency with real-time results for reporting. Further, it was designed as a data-driven enterprise mobile platform to meet the following eight criteria:

1. Designed for the user – role based
2. Be discoverable, simple navigation and recognized control buttons
3. Require minimal user training
4. Be flexible and configurable to accommodate a variety of events
5. Be scalable, on demand, as multiple units are conducting events simultaneously
6. Have a secure environment
7. Provide real time analytics and metrics, and
8. Provide exportable reports

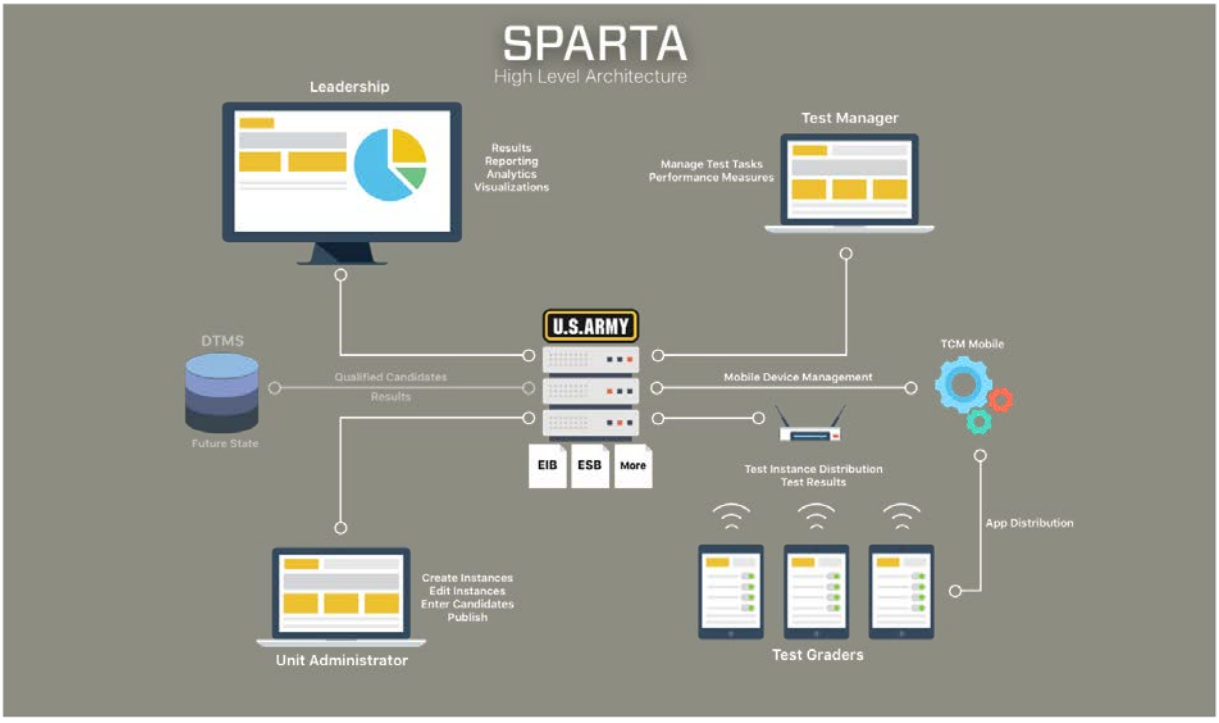


Figure 3. SPARTA high-level architecture. SPARTA was conceived as an enterprise platform hosting both mobile and web application that communicate with a data repository.

The SPARTA platform, as depicted in Figure 3, is comprised of a physical and mobile data model, a mobile application, a web application, and microservices. The technologies used

were selected to be compliant with the Federal Risk and Authorization Management Program (FedRAMP), which is a US government-wide program that delivers a standard approach to the security assessment, authorization, and continuous monitoring for cloud products and services.

Physical Data Model. The SPARTA platform's physical data model defines all of the logical database components and services that are required to create, manage, and execute Army testing events. The data model consists of the table's structure, column names and values, foreign and primary keys, and the relationships among the tables (known as 'objects'). The data model also has constraints such as size, configuration, and security. The data model is a relational model using a Relational Database Management System (RDBMS).

Security. The data model will help the SPARTA platform to implement security via User Roles. The data tables will also have the capability of encrypting personal identifiable information (PII) on a per column basis.

Technology. The data model was created using Microsoft's Structured Query Language™ (SQL) Server 2017 RDBMS. We used the following Database objects:

- Tables,
- Views,
- Stored Procedures, and
- Table-valued Functions.

Challenges and solutions. The biggest challenge we faced was to create a data model that could accommodate various types of Army testing with various rule sets; for example, the EIB 2018 two-strike rule versus the EIB 2016 three-strike rule. Our solution was to design the model to create "gold" copies of each type of test and then have an "instance" of a test for actual testing. The model supports both training and test modes to help further refine the process after analyzing results. We were also able to create a recursive "Test Result" structure that allows for infinite levels of steps for each or any task. This structure allows us to future proof the SPARTA platform if the test were to be revised (adding new steps or sub-steps) to existing or newly developed tasks.

Mobile Data Model. The mobile model is a sometimes-flattened sub-set of the large physical model. It was designed to pull only those fields required by the evaluator to assess a Soldier, e.g., Soldier identifying, current status and/or task information. It houses all the information necessary to conduct an event once the initial download is completed. In addition to providing real-time data sharing in a connected state, this design empowers the evaluator to assess the task with zero connectivity, store the results locally on the tablet, and push the information to the server once connectivity has been re-established.

Technology. We used the Core Data Framework for the local model. The Core Data Framework is an object graph and persistence framework provided by Apple™ in the iOS operating system. The Core Data Framework allows data organized by the relational entity-attribute model to be stored in various formats. The data can be manipulated using higher-level objects representing entities and their relationships. The Core Data Framework manages the serialized version, providing object lifecycle and object graph management, including

persistence. The Core Data Framework interfaces directly with SQLite™ but can be accessed without using SQL through the application programming interface (API).

Challenges and solutions. The challenge faced with the mobile data model was to ensure data synchronization from the mobile model to the physical data model. As modifications occur within the physical model (i.e., additional data points being added), we had to be diligent and ensure the changes were propagated to the mobile model. Additionally, we had to ensure that we had enough of the model in the mobile application so it could completely run in a disconnected mode without needing to retrieve data from the main server.

Our solution was to create a local model with entities and the corresponding relationships. We created the following data process: Controller-Data Store Service/Data REST Service. The Controller would request data from the Data Store Service to display on the UI of the SPARTA mobile application. The Data Store Service would determine if it had the data locally or if it had the ability to retrieve the data by calling the Data REST Service. This process ensures minimal trips to the database resulting in time and bandwidth savings, but also having the capability to refresh the data as needed.

Mobile Application. In response to the Army's move to Apple™ products, the SPARTA mobile application is an iOS based iPad application. The application design was driven by EIB evaluator workflow identified during our training observations. The key workflow items (see Appendix B for a visual sequencing of the workflow steps) are:

1. Select an installation,
2. Select an instance,
3. Select a Test evaluator,
4. Login as the Test evaluator,
5. Select a task to evaluate,
6. Select a Soldier,
7. Read the task instructions to the Soldier,
8. Evaluate the task (GO, NO-GO, and if appropriate, Protest or Elimination due to a Safety or Integrity Violation),
9. Transmit the results, and
10. Repeat for the next Soldier.

The SPARTA mobile application works in a connected (i.e., connected to the internet) and disconnected (i.e., not connected to the internet) state of operation. If in the disconnected state (red colored connectivity icon), the mobile application stores the task performance results locally (on the storage drive of the device). It enables the Test evaluator to have real-time information (when connected, green colored connectivity icon) about each Soldier's current test status, e.g., if the Soldier already has a NO-GO for the task(s). When connected, the SPARTA mobile application sends testing results to the server in real time that is then displayed on the web application at the TOC. The SPARTA mobile application was designed as an adaptive UI that would adjust the controls for the type of test—based on the test meta-data and test parameters. It presents the Test evaluator with the standardized instructions and workflow to test each Soldier and displays a digital stopwatch if the task requires it.

Technology. The SPARTA mobile application is a native iOS application built for the iPad using the following technologies and capabilities:

- Swift 5.0 programming language,
- Core Data for local data manipulation and offline capabilities,
- Native Encryption for data end points with PII,
- AlamoFire Communication libraries,
- Apple™ Design guidance for a familiar UI,
- 100% Apple™ Store compliant, and
- Web Token capabilities for communication authorization.

Challenges and solutions. The greatest challenge our researchers faced was the ability to have the application work in a low communication/no communication[5] environments. The SPARTA mobile application had to capture all the critical data while disconnected and allow the Test evaluator to continue testing. An additional challenge was to have one application have the functionality to adjust dynamically its UI and possible workflow based on the type of Army event. For example, the test results could be binary (GO or NO-GO) or points based (used for other competitions such as the Gainey Cup). The UI would have to adapt to handle either situation with the capability for future expandability.

Our solution for low communications/no communications environments was to use Core Data and a smart caching strategy to create the offline capabilities that enable the SPARTA mobile application to function while disconnected. Additionally, the SPARTA mobile application will automatically reconnect without any user action, and the user is directed to push "Unsent Results" to the server for storage and analysis.

In response to the requirement for a single solution for multiple and sometimes concurrent Army tests/competitions, our researchers designed and implemented a "Dynamic Responsive UI" for the SPARTA mobile application. The SPARTA mobile application has various types of UI elements that help the user capture the information. Based on the event and task type, the SPARTA mobile application will display the corresponding "UI Cell" at runtime, for example, if the task requires a GO/NO-GO response, the application will present the UI Cell with the typical toggle switch icon, set to a GO (green) default. If the Soldier fails any step, the Test evaluator would slide the toggle switch to red for a NO-GO. The SPARTA mobile application will also be able to handle hybrid situations where the task types are different (multiple choice, a timed or scored event) as the UI is driven by the task type metadata.

Web Application. The SPARTA web-based application has two main functions. The first function enables end-users (the event Test Manager) to administer a test. Controlled by user roles, a user can create a "Master Gold copy" of a test such as an EIB, from which other users (testing units) can create "Instances" based off the Master Gold copies. The Master Gold copy guarantees that all instances at any Army installation comes from the same blueprint with no unintentional variances. The second function contains simple analytical tools to help users quickly and accurately assess Soldier data. Some analyses are real-time, like those delivered to the TOC; whereas other, more complex analyses will be useful for comparing results across all

[5] Low and No Communication environments are related to the level of WiFi connectivity available.

test instances, such as identifying the exact sub-step that Soldiers are having difficulty with when tossing a grenade.

Technology. The SPARTA web-based application was created using the following software:

- .net core MVC framework,
- C# language,
- Entity Framework Object Relational mapper,
- Microsoft™ SQL Server,
- Microsoft™ Analysis Server,
- Bootstrap and various JavaScript™ frameworks such as React,
- HTML5, and
- Docker Technology.

Challenges and solutions. There were several challenges involved in making the SPARTA web-based application. The process of creating and managing an Event is simply complicated. The hierarchal structure of the tasks makes the UI especially challenging when entering multi-step tasks. The data analysis capabilities also present their own set of challenges. The real-time data analytics must be balanced carefully between complexity and performance to make the reports both useful and responsive.

Our solution involved an initial design that enabled one of our researchers to create a test and all its tasks. Future capabilities will incorporate a wizard-like interface to guide the user through the complex process while keeping the complexity behind the scenes. The wizard will enable a user to quickly understand the process with little training and make the process much less error prone.

Microservices. The microservices we used incorporated an architectural style that structured an application as a collection of loosely coupled services. In a microservices architecture, services are fine-grained and the protocols are lightweight (using Representational State Transfer [REST] that builds upon HTTP). The benefit of decomposing an application into different smaller services is that it improves modularity. Modularity makes the application easier to understand, develop, and test. The microservices also enabled us to scale up, or build upon existing services, by simply instantiating more containers of that microservices type. Since the microservices run in a Docker environment, they are cloud agnostic, meaning they can run on Amazon™ Web Services (AWS), Microsoft™ Azure, or any other cloud that can run Docker.

Technology. A number of microservices were used in the SPARTA platform. These microservices were:

- Written in C# in the .Net Core framework,
- Utilized Entity Framework Core,
- Containerized in Docker Containers, and
- Deployed using Elastic Container Services (ECS) on the Amazon™ Cloud (AWS).

Challenges and solutions. The main challenge identified will be porting the services over to the "Gov.Cloud" and maintaining a highly secure environment. "Gov.Cloud" is a more secure, hardened environment for highly secure government projects.

Our solution was to build services to be as agnostic as possible with high security capabilities built into the architecture. Because of its deployment within Docker, we were able to take advantage of any secure cloud provider such as the Amazon™ AWS or the Microsoft™ Azure. The security built into the services also protects against some of the more common types of cyber-attacks. Security was developed as follows. First, all communications between the SPARTA mobile application and the web services happen over an encrypted TLS tunnel (HTTPS). Second, the payloads with PII are also encrypted. This minimizes a 'man-in-the-middle' attack as the data are encrypted. This is an attack where the attacker secretly relays and possibly alters the communications between two parties who believe they are directly communicating with each other. Finally, communication tokens are used to authorize communications between the SPARTA mobile application and any web service: the web service will not respond without the corresponding token.

The SPARTA platform was developed in iterative versions, with each version building on the previous version based on end-user feedback and testing. End-users for the SPARTA mobile application were EIB evaluators; therefore, we traveled to multiple Army installations to put iterant versions of the SPARTA mobile application in Soldiers hands.

SPARTA User Experience Feedback

User Experience feedback is an essential element of application development. In this case, feedback from the EIB evaluators was deemed essential as we expect them to use the SPARTA mobile application during future EIB training and testing. The purpose of our UX feedback sessions was to validate the design of the SPARTA mobile application, ensuring that we had captured the evaluation workflow for each station appropriately. Therefore, we structured our UX feedback sessions as an exploratory, hands-on experience by putting the application in the evaluator's hands and allowing them to explore and provide feedback; this is known as Human-in-the-Loop testing. To this end, we coordinated with the training unit for our researchers to interact with EIB evaluators during their unit's EIB training week. The outcome of each interaction was documented by our researchers for consideration in the further development of the SPARTA mobile application.

Participants. Fourteen EIB evaluators participated in the feedback sessions. Evaluators ranged in rank from Sergeant to Sergeant First Class and were a combination of EIB station NCOICs and station evaluators. The EIB validation process, conducted by the USAIS EIB Test Manager prior to our arrival, certified each evaluator on their ability to evaluate each of the stations and tasks they are assigned. Hence, the feedback from these SMEs would help us validate our design approach.

Data collection instrument. A single data collection instrument was developed to facilitate the collection of feedback. The instrument consisted of three parts – Introduction, purpose of research and instructions, a notes section for documenting verbal feedback, and a UX

feedback form utilizing a 5-point Likert scale ranging from 'strongly disagree' to 'strongly agree'. The instructions were specific to the evaluation workflow design. We described the four paths that we wanted the participant to follow and provide feedback. The first path was to assign a "GO" for the station, the second path was to assign a "NO-GO" for the station, the third path was to process a "Protest," and the last path was to process a "Safety Violation." This instrument can be found in Appendix A.

Data collection procedures. Data were collected in a single day at the unit's EIB testing site. The testing site was located in a field environment and consisted of 30 separate EIB testing stations. Each station was manned by one station NCOIC and a minimum of two station evaluators. Each station NCOIC and evaluator had previously been certified as SMEs to evaluate their assigned task(s). Each data collector was equipped with an Apple™ iPad (9.7-inch, 6th Generation iPad) loaded with the SPARTA mobile application and the InvisionApp™ prototype, a Verizon™ MiFi Jetpack, and copies of the data collection instrument.

User Experience feedback sessions began with an introduction and purpose of the session, followed by distributing the iPads to the EIB evaluators. Minimal instructions were provided, e.g., "That is the SPARTA application icon, touch to launch," allowing the participant to explore and provide verbal feedback. Other instructions were provided to ensure that the participant completed all four paths and could provide feedback for each. These sessions lasted 30 – 45 minutes per participant.

Results. User Experience data collection results were aggregated across all participants. Our design approach for the User Experience evaluation workflow was confirmed by all participants with the exception of one function – capturing a NO-GO for exceeding the task time limits. The majority of the remaining UX feedback related to application improvements.

Qualitative (verbal comments) and quantitative (Likert-scale responses) data were collected from each participant. Qualitative data were entered into a Microsoft™ Excel spreadsheet and coded by categories related to navigation, workflow, display, and ergonomics. Comments that did not fall into these four categories were maintained but tabled for inclusion pending future data collections. In some cases, each category was further coded into themes based on a frequency count of the items. Similarly, the Likert-scale responses from the UX feedback instrument where compiled and analyzed. The data are presented by category in the Tables 2-5 below. The quantitative data were sort-ordered based on averages of ratings from high to low and coupled with qualitative data for a more complete picture.

Navigation. The participants verbal responses related to navigation were very positive and are reflected in the ratings in Table 2 below. Seventeen comments were noted that indicated that the complexity of the application was hidden by the simplicity of the navigational design. Comments ranging from "Very user friendly" to "Wonderfully designed" to "Love the layout and simple feel of the task page" were common.

Table 2

Participants' responses relating to navigation

Navigation Questions	Average	Minimum Rating	Maximum Rating
Navigating through the task instructions and performance measures was easy	5.0	5.0	5.0
Selecting an installation was easy	4.9	4.0	5.0
Selecting a test instance was easy	4.9	4.0	5.0
Logging-in was easy	4.9	4.0	5.0
Selecting a station was easy	4.9	4.0	5.0
Selecting a Soldier was easy	4.9	4.0	5.0
Logging-out of the application was easy	4.9	4.0	5.0
The touch target sizes (buttons, switches, etc.) were large enough to select	4.9	4.0	5.0
Overall the navigation was intuitive	4.9	4.0	5.0
Navigating to the task timer was easy	4.8	4.0	5.0
I was able to back-out of a wrong selection	4.5	1.0	5.0

Note: $N = 14$.

***Workflow*.** We noted 73 comments related to workflow in Table 3. Of these comments, we were able to identify seven themes for improving the application. In addition to these themes, 19 comments were related to confirming our evaluation workflow design. The themes, in descending order based on the frequency count of comments, included the following:

- Comment boxes and free text entries – 18 comments
  - Participants would like to see comment boxes in each of the NO-GO, Safety Violation, and Protest pop-up windows with the ability to enter textual comments associated with each.
- Performance Measures/Steps/Sub-steps – 7 comments
  - These comments related primarily to expanding the steps and sub-steps for certain tasks.
- Soldier information – 7 comments
  - Participants identified that the hierarchy of information for each Soldier should be Company, Battalion, and Brigade, with Company and Battalion as a minimum to identify the correct Soldier for testing.
- Timer functions – 6 comments
  - Participants would like audio and haptic feedback when task time ends.
  - Participants would like to have a pop-up reminder of how much time is left for a Soldier to re-test. This was primarily identified as a Station NCOIC function, as it could interfere with an evaluator when grading a Soldier.
- Soldier Selection – 6 comments
  - Participants indicated that they would like to have the Soldiers' names in the contact list grouped by Battalion to ease selection. Two participants indicated that they would be favorable to the use of QR codes as a method of selection.
- Pop-up confirmation – 6 comments

- Participants indicated that they would like to have pop-up messages to confirm an action. For example, logging a NO-GO would result in a pop-up message of "Are you sure?"
- NO-GO for time – 4 comments
  - While only four participants indicated that a NO-GO is needed for the exceeding time option, we recognized that this was a crucial missing piece of the workflow.

Table 3

Participants' responses relating to workflow

Workflow Questions	Average	Minimum Rating	Maximum Rating
Assigning a **GO** for the task was easy	4.9	4.0	5.0
Processing a **Protest after a NO-GO** is easy	4.8	3.0	5.0
Assigning a **NO-GO** for the task was easy	4.7	3.0	5.0
The "Strikes" in the Soldier information area are easy to understand (status of Soldier's NO-GOs)	4.6	4.0	5.0
The sequence of screens matched the way I would evaluate my task	4.6	2.0	5.0
The Soldier information area (name, unit, etc.) provides enough information to properly identify the correct Soldier for testing	4.4	1.0	5.0
Assigning a **Safety Violation** was easy	3.5	2.0	5.0

Note: N = 14.

***Display*.** We received 19 comments related to the display of information (see Table 4). Of these comments, we were able to identify three themes for improving the application. The themes, in descending order based on frequency count of comments, included the following:

- Unclear icon – 14 comments
  - The participants indicated that they had a hard time recognizing that the triangle (hazard) icon represented a Safety Violation. Most recommended a change to the color of the icon but did acknowledge that pre-training would eliminate the confusion.
  - More importantly, two participants identified that Soldiers could be eliminated from the EIB through an integrity violation (primarily during the Land Navigation event) and suggested that we consider providing the option of choosing either Safety or Integrity within the violation pop-up window.
- Button layout – 3 comments
  - Participants indicated that they would like to have a two-button layout in the Protest pop-up window, one button for Clear and Restart Task, and the other for NO-GO.
- Elimination message – 2 comments
  - These two participants indicated that they would like a clear indication when a Soldier is eliminated from the EIB regardless of the reason. Their recommendation was for this message to appear on the Soldier status information page.

Table 4

Participants' responses related to display

Display Questions	Average	Minimum Rating	Maximum Rating
The font type is easy to read	4.9	4.0	5.0
The font size is easy to read	4.9	4.0	5.0
The colors made reading the text easy	4.9	4.0	5.0
The contrast made reading the text easy	4.9	4.0	5.0
The text was easy to read under all light conditions	4.9	4.0	5.0
The screen protector did not hamper my ability to read the information	4.9	4.0	5.0
The icons were easy to recognize	4.7	3.0	5.0

Note: N = 14.

Ergonomics. We noted six comments related to ergonomics (see Table 5). Two participants referenced the possibility of using the application on a smartphone. One participant stated the desire to use a stylus to write comments. One participant recommended using an aviation-grade glare protector on the screen, and one indicated simplifying the options for brightness themes to only Shade and No-shade and removing all other icons from the home screen to reduce information or distractions.

Table 5

Participants' responses related to ergonomics

Ergonomics Questions	Average	Minimum Rating	Maximum Rating
If you used the case handle, was it comfortable	4.6	3.0	5.0
The iPad is the right size	4.5	2.0	5.0
I would use the kick stand in most cases	3.7	1.0	5.0

Note: N = 14.

Other comments. The other comments provided by the participants that did not fall into the above categories but are of equal importance are primarily related to the SPARTA mobile application revisions. We noted six comments where the participants asked for a way for evaluators to send a message to Station NCOICs and Lane NCOICs when there is a Protest on a station. This request, while valid, is predicated on each Station NCOIC and each Lane NCOIC having an iPad linked to the SPARTA platform.

One additional grouping of comments was related to the iPad battery power. Participants voiced concerns that the iPad battery power would not accommodate continuous use over a 10-12 hour testing day. One participant recommended using an Army battery coupled with an adapter to periodically recharge the iPad. We tracked the power usage of two iPads used to collect data; one iPad had to be replaced with a back-up after nine hours of continuous use, and the other iPad had 64% of its battery power available after seven hours and 45 minutes of continuous use. This variance may be due to actual usage and setting options.

SPARTA Modifications

Our primary focus for the UX data collection effort was to validate our assessment workflow design. The results of this data collection confirmed the accuracy of our design with one exception – capturing a NO-GO for exceeding the task time limits. Based on this feedback, we established this functionality as a priority development requirement. The remainder of the feedback was captured and prioritized for future development cycles as we shifted focus to develop the task results web-based application for SPARTA in preparation for the next data collection effort.

Limited User Test

Our second data collection was with an Infantry Brigade Combat Team (IBCT). Prior coordination was again conducted to allow us to interact with EIB evaluators and collect validation data. However, this data collection was also a test of the SPARTA microservices that send the test results to the server and then displays them on the SPARTA web-based application on a laptop located in the TOC.

Web application. Our researchers created a test results web application based on the EIB Standard Unit Tracker.[6] The EIB Unit Tracker is a pre-formatted Microsoft™ Excel spreadsheet that requires manually inputting results at the completion of the day's training or testing. The SPARTA web-based application, as depicted in Figure 4, was designed to replicate the levels of information in a format that was recognizable to Army leaders. The purpose of the SPARTA web-based application is to display the EIB test results captured by the SPARTA mobile application in real-time (connected) or near real-time (disconnected) states of operation.

[6] The EIB Standard Unit Tracker is a required document that each testing unit must complete daily and send to the EIB Test Manager at Fort Benning, GA. The tracker is accessible from the EIB website.

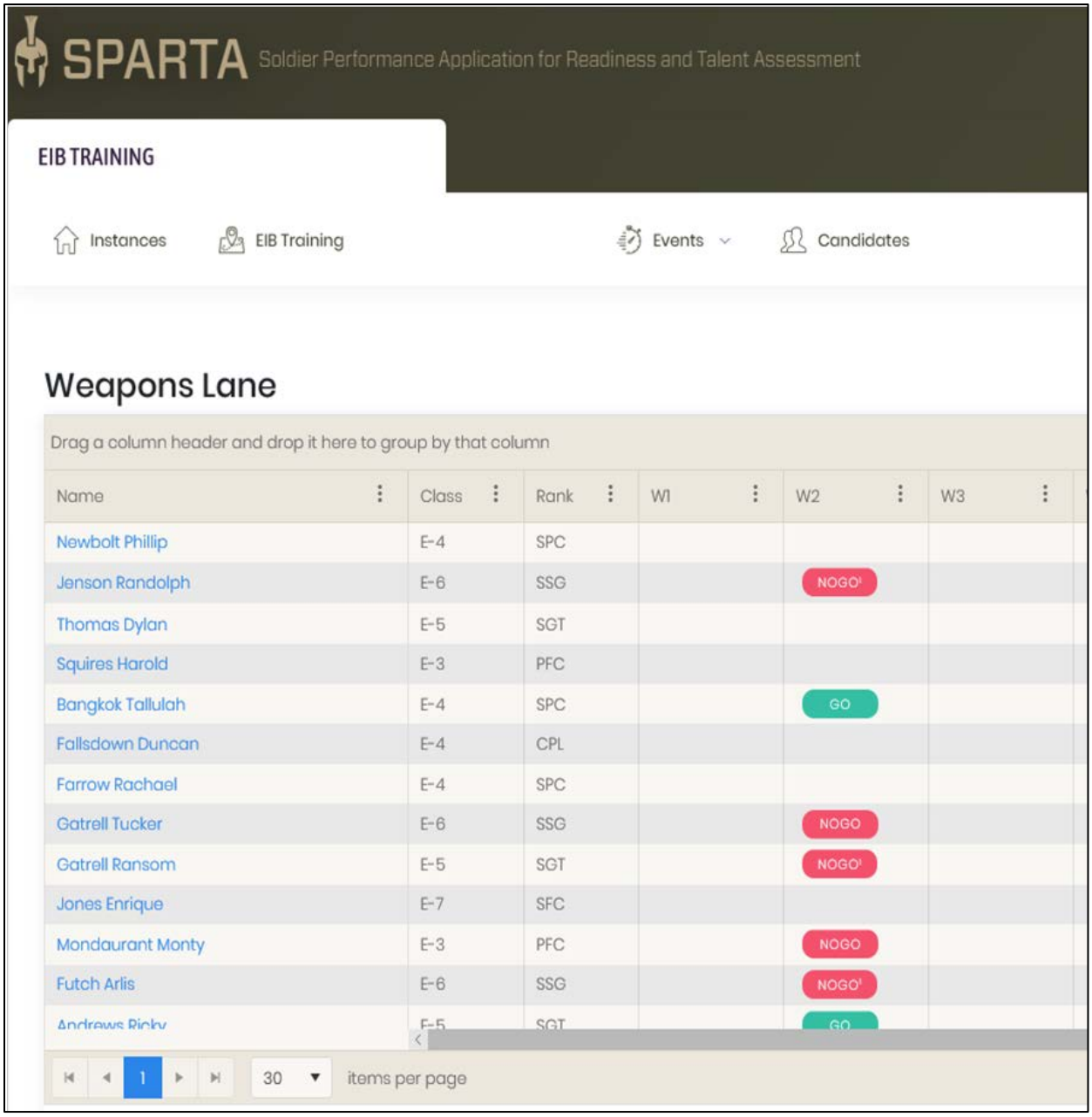


Figure 4. SPARTA test results web application. The application was designed to replicate the Microsoft™ Excel spreadsheet that each testing unit was required to maintain. Soldier information is fictional.

Participants. EIB evaluators from 15 EIB stations participated in the data collection. Evaluators ranged in rank from Sergeant to Sergeant First Class and were again a combination of EIB station NCOICs and station evaluators.

Data collection instrument. The data collection instruments were the same 9.7-inch 6th Generation iPads loaded with the SPARTA mobile application used previously. A test instance was created based upon the tasks the IBCT would evaluate and populated with fictitious names for evaluators and Soldiers (see Figure 5).

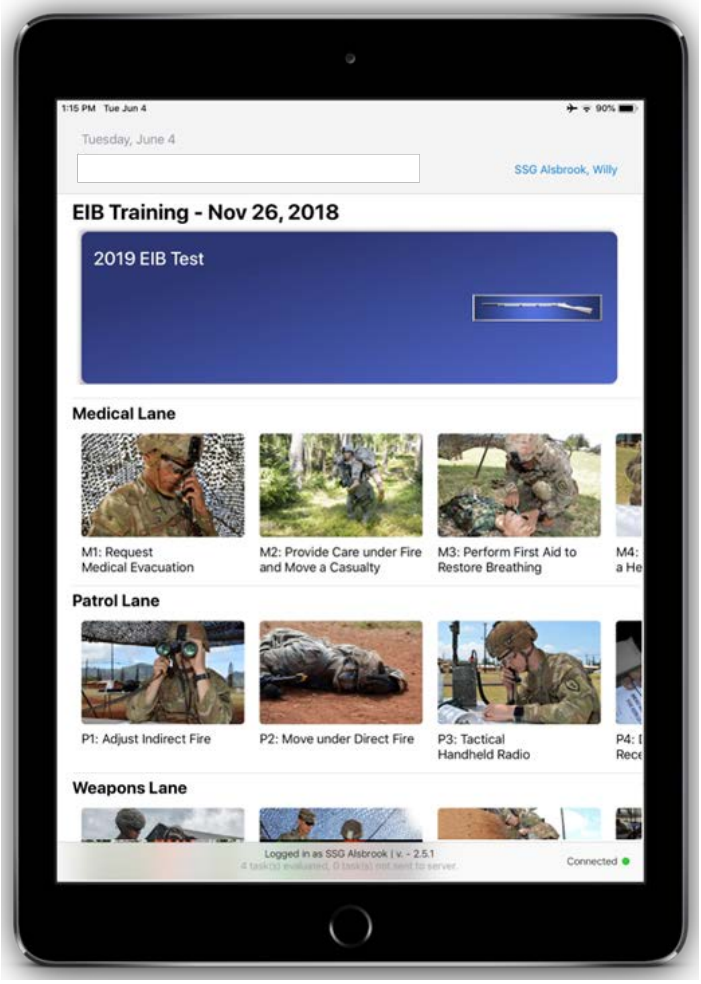


Figure 5. Example test instance. This test instance was specifically created for this unit's actual testing. All names are fictional.

Data Collection Procedures. Data were collected in a single day at the training unit's EIB testing site. The testing site was located in a remote field environment and consisted of 30 separate EIB testing stations. Each station was manned by one station NCOIC and a minimum of two station evaluators. Each station NCOIC and evaluator had previously been certified as SMEs to evaluate their assigned task(s). Our researchers comprised four members – the Principal Investigator, who coordinated each session, and three data collectors who conducted the sessions. Each data collector was equipped with an Apple™ iPad loaded with the SPARTA mobile application and a Verizon™ MiFi JetPack (also known as a 'puck'), as depicted in Figure 6, for connectivity.

Figure 6. Verizon™ MiFi puck. The puck was deployed to provide connectivity in an austere field environment.

Data were collected from EIB evaluators as each had the opportunity to select their specific task, evaluate a Soldier, and send the results to the server. Minimal instructions were provided, e.g., "You are Willy Alsbrook, and your login passcode is XXXX." This allowed participants to explore the SPARTA mobile application, select their task, select a Soldier, evaluate the task, and send the results to the server.

During the data collection, a researcher was co-located next to the TOC and equipped with a Dell™ Inspiron 15-inch laptop and a Verizon™ puck. Displayed on the laptop was the SPARTA web-based application used to monitor the real time delivery of the results and associated data packages sent by the EIB evaluators in the field.

Results. Our primary objective for this data collection was to test the send-receive-display microservices with the SPARTA web application. We accomplished our primary objective with 123 data packages received; however, we did identify the need to revisit the ability to push unsent/disconnected results that were stored in the iPad's local memory.

Integral to our success was the level of connectivity provided by the pucks in a remote location of an Army installation. We had conducted bench testing at locations with established constant connectivity but had yet to test the SPARTA platform in a Low Communication/No Communication connectivity environment.

Data were collected from a subset of the 30 stations with researchers sending additional data from throughout the test site to test levels of connectivity. Results are presented in two sections. First, results related to our primary task, testing the send-receive-display microservices of the SPARTA web application, and second, supplementary UX feedback provided by EIB evaluators.

***Microservices*.** As discussed in the SPARTA Development section, our researchers created microservices to send and receive data packages from the iPad to the server, and from the server to the web application. The data packages, approximately 100 kilobytes or less, contain the task evaluation results associated with a Soldier, the task, and the result (GO, NO-GO, and if

appropriate, Protest or Elimination due to a Safety or Integrity Violation). Packages vary in size based on the result: a GO result aggregates the data at the performance measure level, while a NO-GO result is captured down to the failed performance step or sub-step level. This process enables a finer level of analysis related to task failure. The data is further aggregated up to the task level when received within the relational database, i.e., a NO-GO for a Soldier on task W2 – Pistol and Shotgun. This task-aggregated data is then pulled by the web application for display in the test results table. No Safety or Integrity violations occurred during this data collection. Our data collection resulted in a combined total of 123 GO and NO-GO data packages received and displayed via microservices (Table 6).

Table 6

Example Data Packages

Event	1st Time GO	1st Time NO-GO	2nd Time GO	2nd Time NO-GO
EPFA	2	--	--	--
Land Navigation	1	--	--	--
Weapons Lane	12	8	6	4
Medical Lane	12	9	4	2
Patrol Lane	30	16	3	10
12-Mile Forced March	1	2	--	--
Final Event	1	--	--	--
Total	59	35	13	16
Grand Total		123		

Based on the most recently established rules for EIB testing (the 2-strike rule), Soldiers can be evaluated no more than twice for the same task and receive a GO or NO-GO depending on task proficiency. For example, as presented in Table 6, Soldiers who received a 1st time NO-GO on a task within the Weapons Lane would re-test and either receive a 2nd time GO or a 2nd time NO-GO (resulting in elimination). The SPARTA web-based application was designed to capture and display both attempts if necessary.

The test results table was designed to display the GO/NO-GO status of each Soldier based on the evaluated result. Figure 7 depicts the design incorporated to display these results. The use of superscripts indicates the GO/NO-GO status for each Soldier for each task.

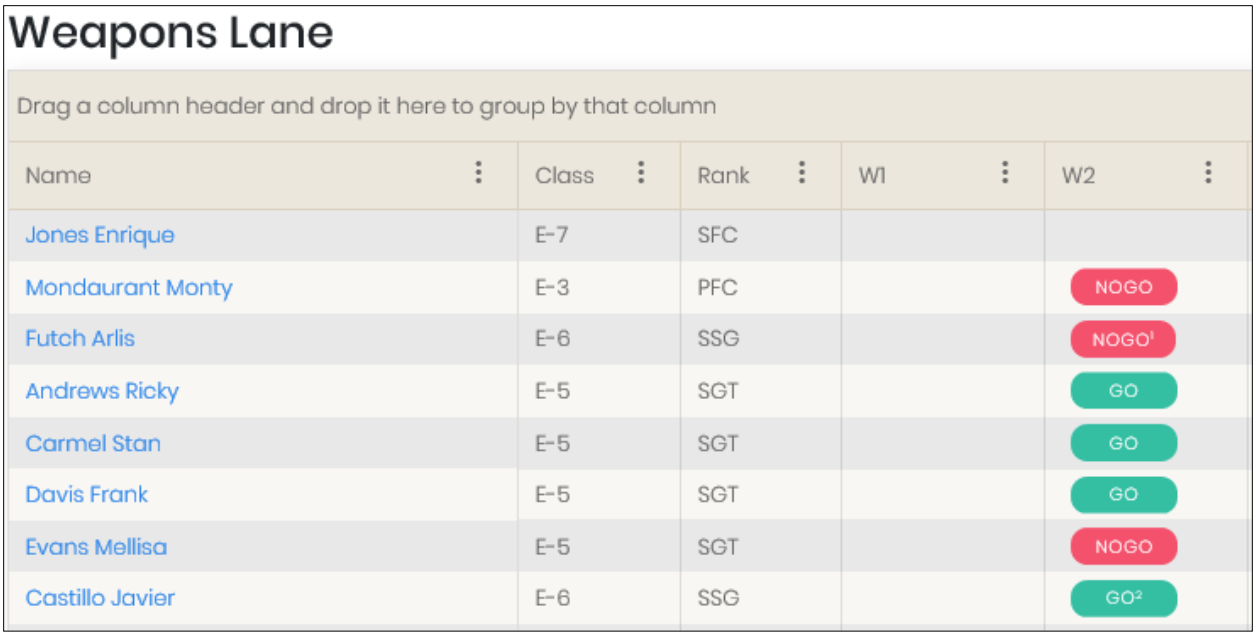


Figure 7. Example of task results table. The task results table was designed to display Soldier GO/NO-GO status. The use of superscripts indicated either first or second attempts. All names are fictional.

The results with superscripts are differentiated between GO and NO-GO as the status of the Soldier is impacted by both outcomes. For example, if a Soldier passes the task on the 1st attempt a green pill with 'GO' (no superscript) is displayed and the Soldier continues on to the next task. If a Soldier fails their first attempt, then a red pill with 'NO-GO[1]' is displayed and the Soldier must re-test on that task within one hour. The re-test result is displayed either as a green pill 'GO[2]' (passed on the second attempt, bottom row of Figure 7) or red pill 'NO-GO' (no superscript) indicating a double NO-GO, and the Soldier is eliminated from testing.

Overall, we found that connectivity varied across the EIB test site from a signal strength of two bars for 4G LTE to no connectivity and variations between 2-3G speeds interspersed throughout. However, even under very low-communication conditions (1X), the SPARTA platform was still able to send the data packages from the iPad to the server. On the receiving end, the SPARTA web-based application, we had a constant 4G LTE signal, and all sent data were retrieved from the server and displayed in the test results table. The implications here are that the WiFi infrastructure impacts the real-time data transfer and the situational awareness to the TOC.

The one area that we had accounted for but needed to revisit was the functionality for pushing unsent results when the iPad transitions from a disconnected (no communications) to a connected (with WiFi) state. The iPad works in a disconnected state by saving each Soldier's task result to the local memory of each iPad. We needed to confirm the process for pushing unsent results upon reconnecting and validating the data result on the receiving end to ensure that a complete package was received.

***EIB evaluator feedback*.** User Experience feedback from the EIB evaluators was primarily positive; we documented comments related to the "Simple interface" that would require very little training to use. One evaluator stated "I could take this [SPARTA mobile application], and the two minutes I spent with you [one of our researchers], and evaluate everybody today." Another evaluator noted that "When you get an app on a phone, no one is there to train you on how to use it, you don't need training for this [SPARTA mobile application]." Other positive comments heard were "Pretty user friendly," "Super cool," and "This is awesome."

Evaluators also identified additional features that could improve the SPARTA mobile application. One such feature was the ability to evaluate Soldiers multiple times for a single task during training. Our initial development focused on the EIB ***test*** mode, where established rules limit the Soldier to two tries per task; however, this comment highlighted the need to have a ***training*** mode, where the Soldier is afforded the opportunity to try as many times as they want for each task during the train-up. Additionally, two other revisions were suggested, (a) reduce the size of the timer and have it permanently on the screen, and (b) provide an elimination message in the test mode that identifies all Soldiers who are eliminated from that test instance.

SPARTA Updates

The outcome of the multiple UX and workflow data collection efforts culminated with the testing and release of five subsequent revisions of the SPARTA mobile application and associated SPARTA platform microservices. All UX revisions were created in the InvisionApp™ prototyping application and demonstrated to the EIB Test Managers for approval. Upon approval, changes were made to the SPARTA mobile application and the SPARTA platform microservices in parallel. The following list of revisions were incorporated into the SPARTA mobile application and the SPARTA platform microservices:

- Resized the timer and enabled persistent view,
- Created microservices to retrieve and store performance time data,
- Created a training mode or testing mode option,
- Enabled multiple evaluations per task, per Soldier in the training mode,
- Created microservices to push and receive unsent results,
- Created UI to push unsent results,
- Created UI for text comments within tasks, and
- Created a 'Soldier eliminated from test' message.

These revisions were readied in anticipation for a large scale deployment of the SPARTA platform with an Infantry Battalion.

SPARTA Deployment

The initial smaller scale user tests were followed by a larger scale deployment test. The intent of this test was to deploy the SPARTA platform on a large scale with over 250 Soldiers being evaluated by 40 EIB evaluators. The researchers coordinated with the leadership of an Infantry Battalion to deploy the SPARTA platform during one day of their EIB train-up. We created a test instance specific to the stations and equipment used by the Infantry Battalion and deployed the SPARTA platform on the second to last day of their EIB training, which is typically used as a "simulated" testing day (no coaching; evaluations are conducted under test conditions).

EIB Test Instance

The SPARTA platform is capable of hosting multiple instances of an EIB test, each instance customized to the unit conducting the test. The USAIS PAM 350-6 enables each unit conducting an EIB to select from options available for certain individual testing stations (Weapons [W] 1, 2, and 10, and Patrol [P] 9). To illustrate, depending on the weapon systems issued to a unit, units can choose from either the M203 or the M320 grenade launcher for station W1. Similarly, the unit can choose from the M9 pistol or the M17/18 pistol for station W2. During initial coordination with the Battalion leadership, we identified the options that this unit had chosen for their EIB test in order to create a test instance specific to that unit; the unit chose the following options:

- W1 - M320 Grenade Launcher (vice the M203),
- W2 - M17/18 Pistol (vice the M9),
- W2 - M500 Shotgun (vice the M870),
- W10 - AT4 Anti-tank Weapon (vice the LAW or Carl Gustaf), and
- P9 - M18A1 Claymore Mine (non-electric initiation) (vice the electric initiation).

Coincidentally, during initial coordination with this testing unit, we found out that the EIB test had been modified by the USAIS, the proponent for the EIB, and they had completely removed W6 – AK 47 Rifle from all EIB testing. Therefore, we made the necessary changes to the testing unit's EIB instance resulting in 29 tasks available for evaluation.

Additionally, we created unique EIB evaluator identifiers for each station, as well as unique Soldier identifiers for this evaluation. The SPARTA platform, in its present state, does not use Soldier PII. Therefore, we created the EIB evaluator's login information by using the station that the EIB evaluators were assigned (W1 evaluators logged in by selecting "W1_evaluator 1" or "W1_evaluator 2," W2 evaluators used "W2_evaluator 1" or "W2_evaluator 2," and so on). We created 60 evaluator login names, 20 for W1-W10, 20 for P1-P10, and 20 for M1-M10, but only used 58 as the W6 station was removed. Further, we developed a method by which each Soldier testing could be identified to a corresponding Soldier number in the SPARTA mobile application for the evaluation during the training day. We created 541 numerical Soldier identifiers, labeled 001-541, and then printed each number on a business-sized card for the Soldiers to use as their identification prior to being evaluated at any station during the training day. Neither the EIB evaluator identifiers nor Soldier numbers were matched in any way to an evaluator's or Soldier's identity.

SPARTA as Deployed

The deployed SPARTA platform comprised of 60 Apple™ 9.7-inch, 6th generation iPads protected by NewTrent™ cases, seven Verizon™ MiFi pucks for connectivity, and one Dell™ laptop to display the web application. Prior to use, all hardware was fully charged, iPad settings were optimized for minimal power consumption, and each iPad was loaded with the SPARTA mobile application displaying the test instance created for the unit's EIB.

Participants

Approximately 38 EIB evaluators participated in this larger-scale user test. Evaluators ranged in rank from Specialist to Sergeant First Class and were a combination of EIB station NCOICs and station evaluators. The EIB validation process certified each evaluator on their ability to evaluate the stations and tasks they were assigned. Therefore, as the SPARTA mobile application was designed to replicate the evaluation workflow and grading sheets for each task, the NCOICs and evaluators were familiar with the content contained within the SPARTA mobile application. Additionally, approximately 280 Soldiers were evaluated using the SPARTA mobile application. Soldiers ranged in rank from Private (E-1) through Captain (O-3).

Data Collection

Through prior coordination by the Principal Investigator, our researchers were granted access to the EIB training site and EIB evaluators. Scheduled training for the day of data collection was restricted to the Weapons and Patrol lanes only, with the Medical lane closed as the unit split training between Land Navigation and Individual Task Testing. Our data collection spanned over 10 hours and involved deploying, monitoring and troubleshooting the SPARTA platform, and gathering user feedback.

Deploying SPARTA. We deployed the SPARTA platform across 19 individual testing stations, P1 – P10 and W1 – W10 minus W6. The individual testing stations were located in an area approximately $0.23km^2/0.14m^2$ on a remote training site. Stations were spread linearly in a heavily wooded environment with a TOC [command post] centrally located in an adjacent open field. MiFi 4G LTE connectivity was provided by a commercial cell tower located approximately 8.67km/5.39m from the training site. Environmental conditions on the day of this evaluation were approximately 60 degrees Fahrenheit with intermittent showers.

Upon arrival at the EIB training site, our researchers issued 42 iPads to evaluators of each EIB individual testing station at a rate of two iPads per station with one exception: one station requested six iPads as they planned to evaluate six Soldiers at a time. We positioned six MiFi pucks throughout the training site to provide a connectivity footprint to cover all 19 stations. A seventh MiFi puck was co-located in the TOC with the laptop for access to the SPARTA web application and test results table.

We conducted a brief user training session for the evaluators after issuing the iPads that was focused on how to login to the test instance, select their assigned stations, and select a Soldier to evaluate. We also briefed the evaluators on the numbering solution to identify each

Soldier for evaluation. We answered all questions and released the evaluators to occupy their testing stations.

When the Soldiers arrived, they were evenly distributed and staged behind station placards (groups of Soldiers were lined up behind the placards listing stations W1, W2, W3, and so on). We issued each Soldier a numbered business card and provided them with a brief of how the day's training would be conducted; when the Soldier was ready to test they would provide the station evaluator with their number. The evaluator would then select that number in the SPARTA mobile application to begin the evaluation process for that station. This data collection procedure was in addition to the usual process of the Soldiers handing their EIB tracking sheets to the EIB evaluator. For this, and for all data collections mentioned in this report, the typical procedure was followed, and then the EIB evaluators subsequently added the evaluation decision into the SPARTA mobile application, adding just a few seconds to their procedure. The Soldiers were then released and proceeded to their first test station of the day.

Monitoring and troubleshooting SPARTA. Our four researchers rotated between the test stations and the TOC to observe the EIB evaluators using the SPARTA mobile application for the primary purpose to address and troubleshoot any issues that arose with using the SPARTA mobile application, while also documenting the EIB evaluators' initial impressions. Throughout the course of the day, our researchers observed each individual testing station as Soldiers were evaluated, and elicited feedback from each EIB evaluator. Troubleshooting was limited to swapping iPads as power consumption varied, and moving/elevating the MiFi pucks to re-establish connectivity. At the end of the day, we recorded the power levels for all the iPads.

Results

Results from this larger scale data collection are presented in two sections. The first section describes the results related to the SPARTA hardware tools and the effects of using iPads in an austere field environment. The second section describes the results of the use of the SPARTA mobile application to assess EIB performance.

Hardware results. We identified that power consumption and connectivity are two crucial factors when using iPads in an austere field environment. What we, as smartphone users, take for granted in our daily lives is a logistical challenge when moving technology into an austere field environment that has limited opportunities for recharging and sporadic interruptions to connectivity.

***Power consumption*.** Battery life was identified as a concern from users during previous data collection efforts. The iPad specifications state that under normal use the battery should last approximately 10 hours; this was our benchmark. To reach this benchmark, we optimized power consumption settings on each iPad using the following steps:

1. We updated each iPad with the most current iOS version,
2. We adjusted Auto-Lock settings to 2-minutes,
3. We turned on Auto-Brightness,
4. We disabled auto-updates, and
5. We put each iPad in Guided Access mode to disable users from changing these settings.

Power consumption varied by tablet and was impacted by proximity to the MiFi puck. Power levels for each iPad were documented at the end of the 10+ hour evaluation and are presented in Table 7.

Table 7

iPad power consumption

Remaining Power	Number of iPads
0-25%	24
26-50%	17
51-75%	4
76-100%	2

Of the 47 iPads used, one iPad was replaced within 5-hours with 19% power remaining due to the auto-brightness setting mistakenly set to "never"[7]; three iPads were replaced around the 7-hour mark with 0% power remaining; and one iPad had 0% power remaining at the completion of training. Average power consumption of the original 42 iPads issued was 77% of available power (23% remaining). Additionally, the average power use of the seven Verizon™ MiFi pucks consumed 75% of available power (25% remaining) over the same timeframe.

Connectivity. We know that WiFi infrastructure does not typically exist in field environments on most Army installations, the exceptions being the Combat Training Centers at Forts Irwin and Polk. Therefore, one of our aims was to demonstrate the use of local hotspots as a means of providing the connectivity required to capture and display real-time data. We deployed six MiFi pucks across the 19 stations. We located the pucks in areas that would provide overlapping coverage to the stations and elevated them as much as possible for better connectivity.

WiFi connectivity across all the EIB stations was spotty at best due to the distance between stations, the number of MiFi pucks deployed, and the heavily forested landscape. Those stations that were nearest the MiFi pucks had no issues pushing data in real-time. Stations on the periphery of connectivity resulted in greater power consumption of the iPads and had some unsent data stored on the iPads. These data were pushed as soon as connectivity was re-established. This observation leads to the recommendation that, at a minimum, there should be at least one MiFi puck per two EIB stations if no overarching infrastructure is available.

Performance results. The SPARTA mobile application was designed to collect data for each evaluated task at the performance measure, step, and sub-step(s) level. For example, P1 – Adjust Indirect Fire is a single task comprised of three performance measures, nine performance steps, and eight sub-steps. Based on the task assessment results of an EIB evaluation performed by a Candidate (GO or NO-GO), the SPARTA platform will collect a subset of data points (GO results) or all data points (NO-GO results) for all tasks. During this 10-hour data collection, the SPARTA platform collected 33,391 rows of data. Of those, the Patrol Lane data is presented below as an example of the level of analysis fidelity achievable.

[7] Auto-brightness settings dim the screen and then shuts it off completely. This iPad was set at 'never,' which caused the screen to remain on, consuming battery power quicker.

Patrol lane results. The SPARTA platform was used to capture performance data as Soldiers were evaluated on the following stations:

- P1: Adjust Indirect Fire,
- P2: Move under Direct Fire,
- P3: Tactical Handheld Radio,
- P4: Defense Advanced GPS Receiver (DAGR) Operations,
- P5: Camouflage and Visual Signaling Techniques,
- P6: Range Card,
- P7: Chemical and Biological Operations,
- P8: Resection and Military Maps,
- P9: M18A1 Claymore Mine (Non-electric Initiation for this instance), and
- P10: Transmit a Spot Report with a Tactical Man Pack Radio.

This data (Figure 8) is available within the TOC in real time when the iPads are in the connected state, and in near real time when in the disconnected state. In an analog TOC (paper and pencil, which is the current mode of operation), the results are not available until the end of the training day (oftentimes going into the night) when all paper copies have been turned in and collated.

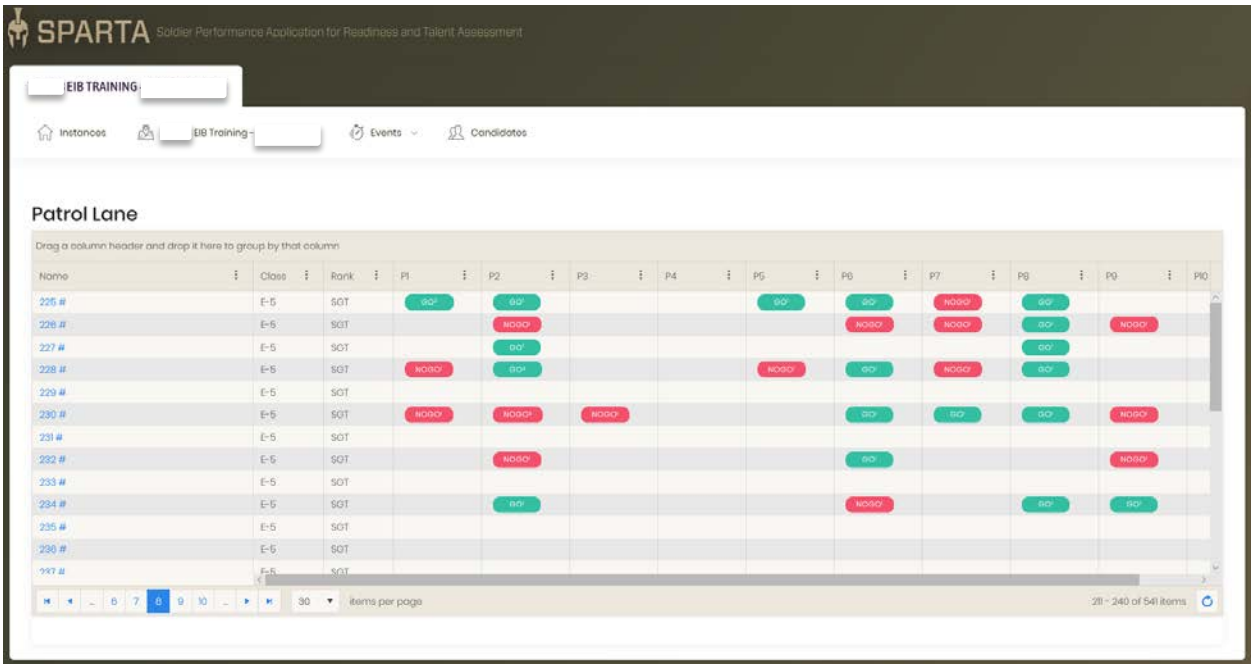


Figure 8. Example task results table. The SPARTA web application (located in a TOC) enables leaders to view results in real time. Refresh rate can be set from minutes to seconds.

At the completion of the data collection, we had amassed over 16,000 lines of data across the 10 Patrol tasks. When aggregated at the task level, 516 records were available with respect to GOs and NO-GOs by task as displayed in Figure 9.

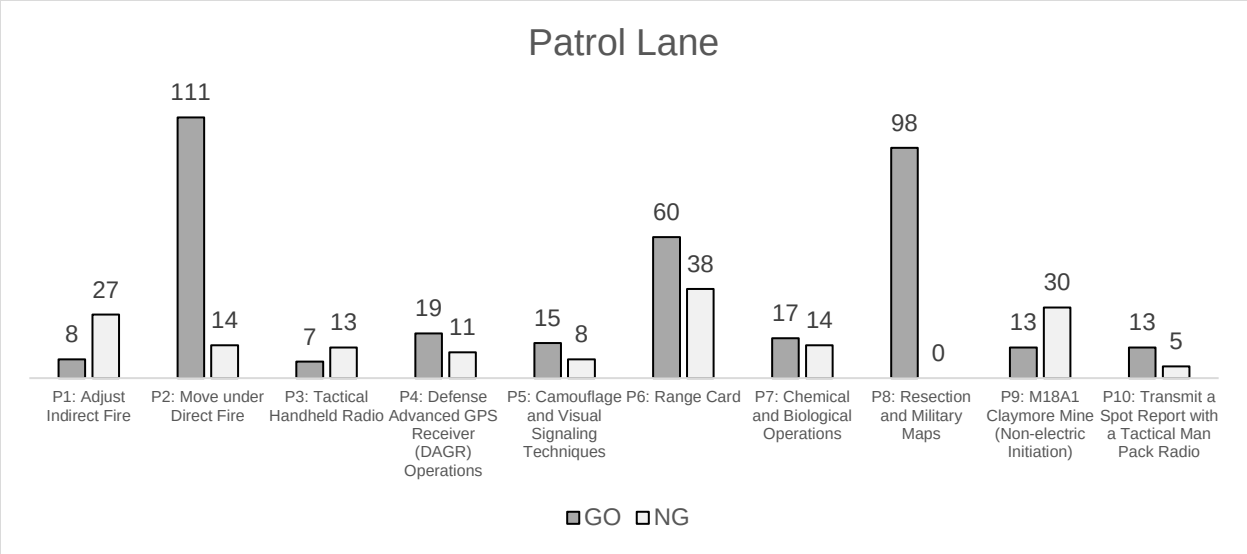


Figure 9. Patrol lane results. The SPARTA platform provides macro-level results in an easily digestible format.

The SPARTA platform then further enables drill down analysis within tasks, unlike the current paper and pencil process. For example, drilling down into the task P1 – Adjust Indirect Fire, the micro-level results by performance measure, step, or sub-step for the 27 Soldiers who received a NO-GO are depicted in Figure 10.

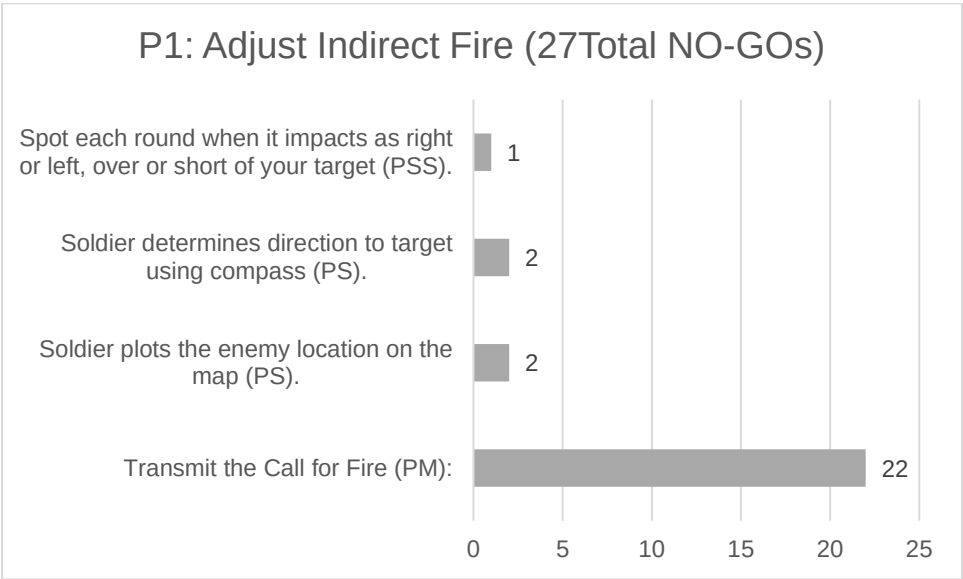


Figure 10. Station P1 – Adjust Indirect Fire micro-level results. The data captured by the SPARTA platform enables leaders to determine the specific step that a Soldier failed a task.

In addition, the SPARTA platform collects time data. The time data is associated with certain performance measures. For the task Adjust Indirect Fire, the first performance measure, Transmit the Call for Fire, must be accomplished within 3 minutes, the second, Adjust Fire within 45 seconds, and the third, Fire for Effect within 30 seconds. For example, the NO-GOs associated with the performance measure (PM), Transmit the Call for Fire (Figure 10) are due to

Soldiers exceeding the time standard for the task. Of the 27 Soldiers who received a NO-GO at P1, 22 Soldiers exceeded the 3 minutes allowed for Transmit the Call for Fire. Usually, during training days, the evaluators allow the Soldiers to continue the task until complete and then record the time to completion. Table 8 displays the actual time collected per Soldier.

Table 8

Times recorded for Soldiers receiving a NO-GO for exceeding the Transmit the Call for Fire time standard

Soldier ID Number	Time in seconds	Converted Time
210	189	3 minutes, 9 seconds
215	202	3 minutes, 22 seconds
238	384	6 minutes, 24 seconds
241	191	3 minutes, 11 seconds
245	189	3 minutes, 9 seconds
286	184	3 minutes, 4 seconds
289	187	3 minutes, 7 seconds
296	188	3 minutes, 8 seconds
299	219	3 minutes, 39 seconds
302	182	3 minutes, 2 seconds
304	222	3 minutes, 42 seconds
312	265	4 minutes, 25 seconds
322	193	3 minutes, 13 seconds
327	184	3 minutes, 4 seconds
332	222	3 minutes, 42 seconds
335	194	3 minutes, 14 seconds
339	182	3 minutes, 2 seconds
356	233	3 minutes, 53 seconds
363	182	3 minutes, 2 seconds
366	191	3 minutes, 11 seconds
404	181	3 minutes, 1 seconds
575	182	3 minutes, 2 seconds

Similar analyses are available for each task assessed using the SPARTA platform. The fidelity of the results could enable Army leaders not only to conduct focused analysis within tasks but also to look at trends over time. Providing this level of analysis will provide leaders with actionable information on which to make informed decisions for future training, ultimately impacting unit readiness.

Discussion

Whether assessing an Infantry Battalion conducting a Combined Arms Breach, or a Soldier employing a M18A1 Claymore Mine, the assessment format is the same, the T&EO. Through this research, we determined that the T&EO is the model for a sustainable digital assessment tool, if and only if the tool is easily editable in order to address any future changes.

This determination was the basis for which the SPARTA platform was designed; a data-driven, event agnostic tool that would enable Army leaders and evaluators to assess proficiency and determine readiness across all individual tasks (and subsequently collective tasks) conducted by all branches and military occupational specialties. The challenges were many, and, as such, the SPARTA platform had to be:

- flexible to host multiple tests and competitions (EIB, Expert Field Medical Badge (EFMB), Expert Soldier Badge (ESB), Best Warrior, Sullivan and Gainey Cups, etc.);
- flexible to host multiple instances of a single or multiple test and/or competition (simultaneous EIBs while the Gainey Cup is in progress);
- editable to accommodate any future changes to events or tasks;
- able to accommodate differing rule sets for tests and competitions (binary, points-based, or weighting schemes, etc.);
- scalable and able to provide data storage for, at last count, an estimated 1million person Army (USAR/ARNG/Active; Department of Defense, 2018); and
- functional in connected and disconnected states of operation.

Employing an end-user workflow-focused iterative Agile process, our researchers were able to design, development, test, revise, and deploy the SPARTA platform as a proof of concept with Soldiers at all levels. Feedback was positive; Soldiers appreciated the complexity hidden under the veneer of simplicity in a discoverable mobile application.

Inherent in the successful deployment of any tablet-based assessment tool is a connectivity infrastructure with a footprint that encompasses the field environment where Army task training is conducted. The Army is moving in the direction of providing this capability to deploying units; however, we have not identified a concrete move to instrument the installations where home station training is conducted. We demonstrated the capability using commercial WiFi hotspots. In the long run, these could be cost prohibitive when equipping all training and testing units. Further exploration of alternate solutions within the Army IT community is warranted.

Additionally, there are accompanying implications with the use of a tablet with an application for assessing tasks and recording information – tablet maintenance and application updates. The tablets used for this research were the newest 9.7-inch 6th generation iPads running on iOS 12.1. Over the period of performance for this research, the iOS version was updated multiple times ending at version 12.3.1 in June 2019 with a projected release of iPadOS in September 2019. Updating 60 iPads each time requires access to either a WiFi node or an understanding of tethered updates.

Timely tablet maintenance and application maintenance were critical to this research effort. Our researchers, experienced software architects and programmers intimately familiar with the SPARTA platform, were able to push 21 updates during development. We did not use a mobile device management (MDM) solution as this was strictly a proof of concept. However, transitioning from a production application to a fully deployed application will require an Army MDM process that would enable updating the application in similar fashion to how it is currently accomplished commercially (Apple™ App Store or Google™ Play Store).

Recommendations

While the SPARTA platform is capable of hosting multiple instances of differing tests/competitions, Soldiers (ultimately, the end-users) provided feedback on additional applications that could benefit from this product and they offered suggestions for future research. Soldier recommendations related to two categories, the EIB and collective task training.

EIB Recommendations

Soldiers indicated that an additional application that could benefit Soldiers who are testing during EIB train-up might be an EIB Training application (currently known as the Candidate Book). Testing units have historically reproduced the task performance measures, steps, and sub-steps from USAIS PAM 350-6 in pocket-sized booklets and handed them out to each Soldier as a training guide. It was recommended that we create an application that uses the data contained within the SPARTA platform to develop a training application that Soldiers could download to their personal phones. The benefit of this application would be that when changes are made to the "gold" copy of the test, the changes would automatically propagate to the end user tablets through an application update, ensuring all Soldiers are referencing the appropriate information and yielding a potential cost and time savings resulting from not having to (re)print the booklets.

Lane and Station NCOICs indicated that they would like to have a NCOIC application that provides them with summary information related to the number of Soldiers who completed testing on their station or lane, the assessment results, and the remaining Soldiers needing to test. Currently, the NCOICs use Soldiers as runners to gather this information, which is reported periodically to the TOC to enhance the situational awareness of unit leaders. A mobile application that leverages the task data inputted by the evaluators would provide that situational awareness automatically.

Additionally, the NCOICs identified the capability of using an application for notifications and instant messaging. Currently, if there is a protest at one station, the station NCOIC must communicate with the lane NCOIC in order for them to adjudicate the protest. We witnessed multiple protests during our data collection that had the lane NCOIC moving from one station to another, but only if they were aware of the protest. The station NCOICs stated that if they had a summary application, then they would have the capability to quickly communicate with the lane NCOIC about a protest distribute messages across all tablets within the lane, or even across the test site.

Collective Task Training Recommendations

Army leaders have suggested that the SPARTA platform should be used to assess collective task training. Soldiers train, and are assessed, on individual tasks that support the unit collective tasks. Units, on the other hand, train and fight collectively. Collective task assessment follows the same T&EO format; however, the evaluation rule set is significantly more complex. The Army has established Objective T as the rule set for collective task evaluation that results in ratings of Fully Trained (T), Trained (T-), Practiced (P), Marginally Practiced (P-), and Untrained (U) (DA, 2016, p.1). Figure 11 depicts an objective task evaluation matrix for an Infantry Company executing the task *Conduct an Attack.*

Plan and Prepare			Execute						Assess
Operational Environment		Training Environment (L/V/C)	% of Leaders Present at Training/Authorized	% of Soldiers Present at	External Eval	% Performance Measures 'GO'	% Critical Performance Measures 'GO'	% Leader Performance Measures 'GO'	Task Assessment
CO & BN									
Dynamic and Complex (4+ OE Variables and Hybrid Threat)	Night	A live training environment is required for external evaluation of this task and to achieve a T or T- task assessment. The virtual training environment can be used during crawl and walk training events to enhance follow-on live training.	>=85%	>=80%	Yes	>=91%	All	>=90%	T
			75-84%			80-90%		80-89%	T-
Dynamic (Single Threat)			65-74%	75-79%	No	65-79%			P
	Day		60-64%	60-74%		51-64%	<All	<=79%	P-
Static (Single Threat)			<=59%	<=59%		<=50%			U

Figure 11. Example of an objective task evaluation matrix. The evaluation rule set for collective tasks contain numerous variables that are assessed before determining the overall rating.

Additional variables that impact the overall rating are related to the percentage of Soldiers and leaders present for the training. More importantly, as it relates to the SPARTA platform, is the ability to differentiate between the coding of leader, critical, and standard performance measures. The recommendation is to include Objective T as a rule set providing additional capabilities to the SPARTA platform.

We also recommend co-locating all of the SPARTA platform hardware/kit in a single location with employees who have the skills required to maintain the equipment. Possible locations are the Training Audiovisual Support Centers (TASCs). The organizations that provide training support and training devices may be the appropriate locations for the SPARTA platform/kit to reside.

References

Department of the Army (2016). *Train to win in a complex world* (FM 7-0). Washington D.C.: Author.

Department of the Army (2019). *Expert Infantryman Badge*. (United States Army Infantry School [USAIS] Pamphlet 350-6). Fort Benning, GA: U.S. Army Infantry School.

Department of Defense (2018). *Defense manpower requirements report (FY 2019)*. Washington D.C.: Office of the Assistant Secretary of Defense for Manpower & Reserve Affairs.

Milley, M.A. (2016). *Army Readiness Guidance, Calendar Year 2016-17*. Washington, DC: Department of the Army.

U.S. Army Combined Arms Center (2015). *Talent Management Concept for Force 2025 and Beyond*. Retrieved from http://usacac.army.mil/pubs/Force-2025-and-Beyond-Human-Dimension.

U.S. Army Infantry School (Dec 3, 2018). *History of the Expert Infantryman Badge*. Retrieved from https://www.benning.army.mil/infantry/eib/History.html.

U.S. Army Infantry School (July 15, 2019). *Expert Infantry Community*. Retrieved from https://www.milsuite.mil/book/community/spaces/apf/maneuvernet/infantry_profession/expert-infantry-community/pages/statisticstasks.

U.S. Army Maneuver Center of Excellence (July 9, 2019). *Board Analysis.* Retrieved from https://www.benning.army.mil/Infantry/OCOI/Board-Analysis.html.

Acronyms and Abbreviations

DA	Department of the Army
EFMB	Expert Field Medical Badge
EIB	Expert Infantryman Badge
EPFA	EIB Physical Fitness Assessment
ESB	Expert Soldier Badge
MDM	Mobile Device Management
NCO	Noncommissioned Officer
NCOIC	Noncommissioned Officer in Charge
RDBMS	Relational Database Management System
SFC	Sergeant First Class
SPARTA	Soldier Performance Application for Readiness and Talent Assessment
SSG	Staff Sergeant
TOC	Tactical Operations Center
T&EO	Training and Evaluation Outline
USAIS	United States Army Infantry School
UI	User Interface
UX	User Experience

APPENDIX A

UX FEEDBACK DATA COLLECTION PROTOCOL

SUBJECT: Feedback on the Soldier Performance Application for Readiness and Talent Assessment (SPARTA)

Introduction:

Hello, my name is _____ and this is my colleague _____. We work with the U.S Army Research Institute, which is a unit under the Army G1. Our unit focuses on the assessment of Soldiers for the Army. We have been working with the USAIS and EIB managers to digitize some of the EIB process. Today, we are conducting a design validation of a prototype mobile application for grading the EIB. To validate the application we are asking you to try using the application we made so we can see whether it's designed as intended. The first thing I want to make clear right away is that we're testing the design of the application, not you. You can't do anything wrong here. Also, please don't worry that you're going to hurt our feelings. We're doing this to improve the application, so we need to hear your honest reactions.

As you go through the application, I'm going to ask you to try to think out loud as much as possible: say what you're looking at, what you're trying to do, and what you're thinking. This will be a big help to us.

Not everything is selectable yet, but if you try to select something that will be selectable but isn't I'll explain what would happen.

*We want you to use the application as if you were grading an EIB Soldier for your station. First we would like you to perform the log-in steps (**provide an evaluator Name to select and passcode of 1111**), select your station, and select a Soldier (**provide a Soldier name to select)** from the list; then, we would like you to go down four paths.*

The first path is when a Soldier completes the task to standard and receives a GO.

The second path is when a Soldier does not complete the task to standard and receives a NO GO and does not elect to protest

The third path is when a Soldier protests a NO GO.

The last path is logging a Safety Violation when you witness a Soldier performing a safety violation associated with the task.

We will pause between paths to write down your feedback.

After completing all paths we will ask you to complete a feedback form about specific aspects and functions of the application. Any feedback you provide today will be consolidated with feedback from the other evaluators and will help us to improve the application. We really appreciated you taking the time today to assist us.

Comments: (Use this area to capture any evaluator feedback while they are using the application)

SPARTA User Interface Feedback Form

	Strongly Disagree	Disagree	Neither agree or disagree	Agree	Strongly Agree
Navigation					
Selecting an installation was easy	□	□	□	□	□
Selecting a test instance was easy	□	□	□	□	□
Logging-in was easy	□	□	□	□	□
Selecting a station was easy	□	□	□	□	□
Selecting a Soldier was easy	□	□	□	□	□
Navigating through the task instructions and performance measures was easy	□	□	□	□	□
Navigating to the task timer was easy	□	□	□	□	□
Assigning a **GO** for the task was easy	□	□	□	□	□
Assigning a **NO GO** for the task was easy	□	□	□	□	□
Processing a **Protest** after a NO GO is easy	□	□	□	□	□
Assigning a **Safety Violation** was easy	□	□	□	□	□
I was able to back-out of a wrong selection	□	□	□	□	□
Logging-out of the application was easy	□	□	□	□	□
The touch target sizes (buttons, switches, etc.) were large enough to select	□	□	□	□	□
Overall the navigation was intuitive	□	□	□	□	□
Display					
The font type is easy to read	□	□	□	□	□
The font size is easy to read	□	□	□	□	□

The colors made reading the text easy	□	□	□	□	□
The contrast made reading the text easy	□	□	□	□	□
The icons were easy to recognize	□	□	□	□	□
The text was easy to read under all light conditions	□	□	□	□	□
The screen protector did not hamper my ability to read the information	□	□	□	□	□
Workflow					
The Soldier information area (name, unit, etc.) provides enough information to properly identify the correct Soldier for testing	□	□	□	□	□
The "Strikes" in the Soldier information area are easy to understand	□	□	□	□	□
The sequence of screens matched the way I would evaluate my task	□	□	□	□	□
Ergonomics					
If you used the case handle, was it comfortable	□	□	□	□	□
I would use the kick stand in most cases	□	□	□	□	□
The iPad is the right size	□	□	□	□	□

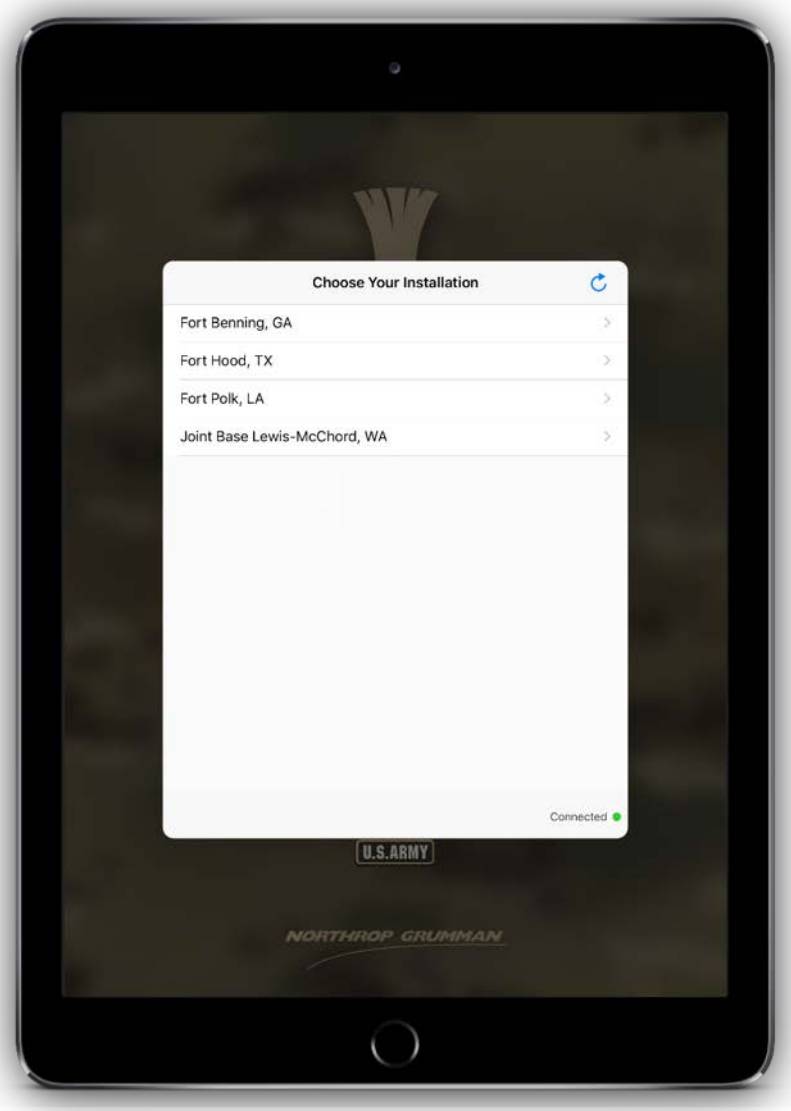
Thank you, this has been very helpful!

APPENDIX B

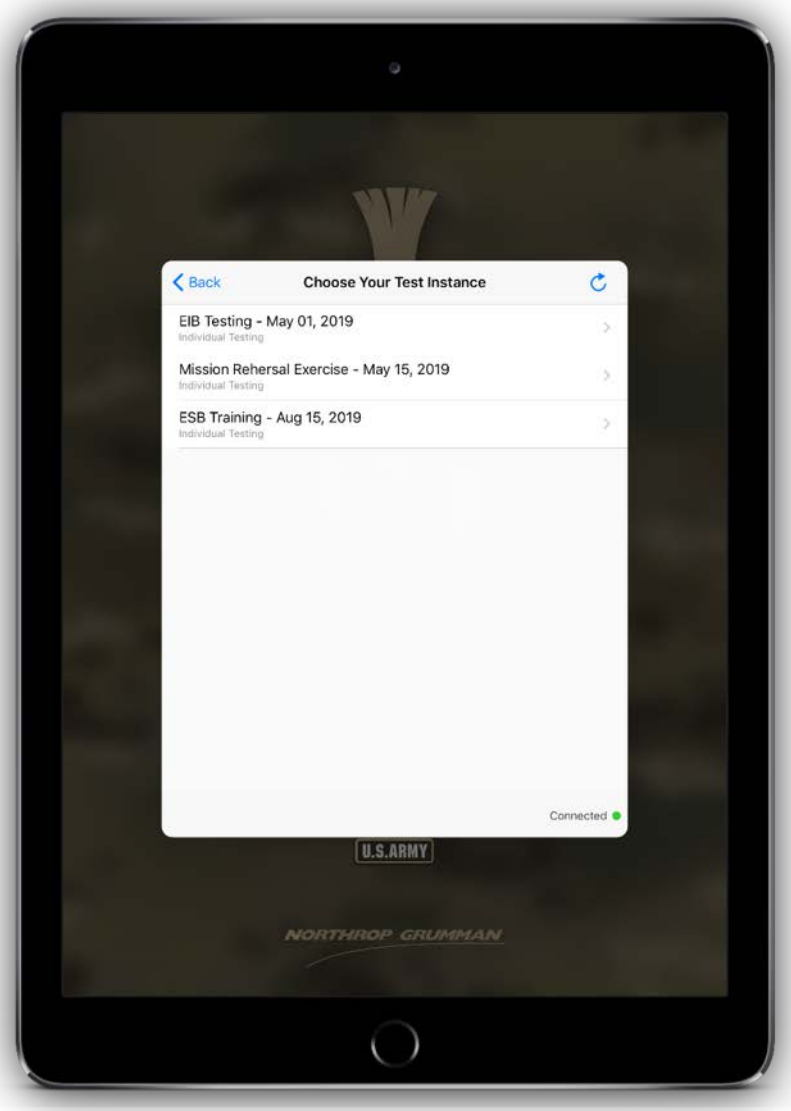
SPARTA MOBILE APPLICATION SCREENS

SPARTA MOBILE APPLICATION SCREENS

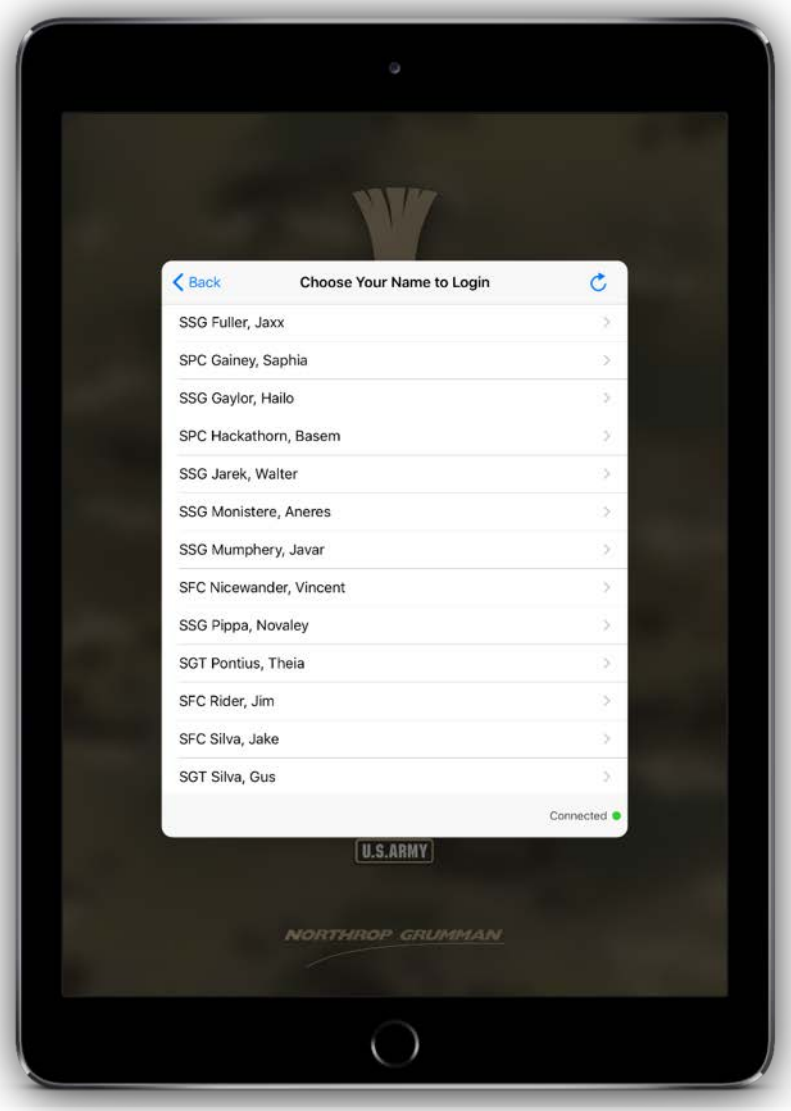
SPLASH SCREEN



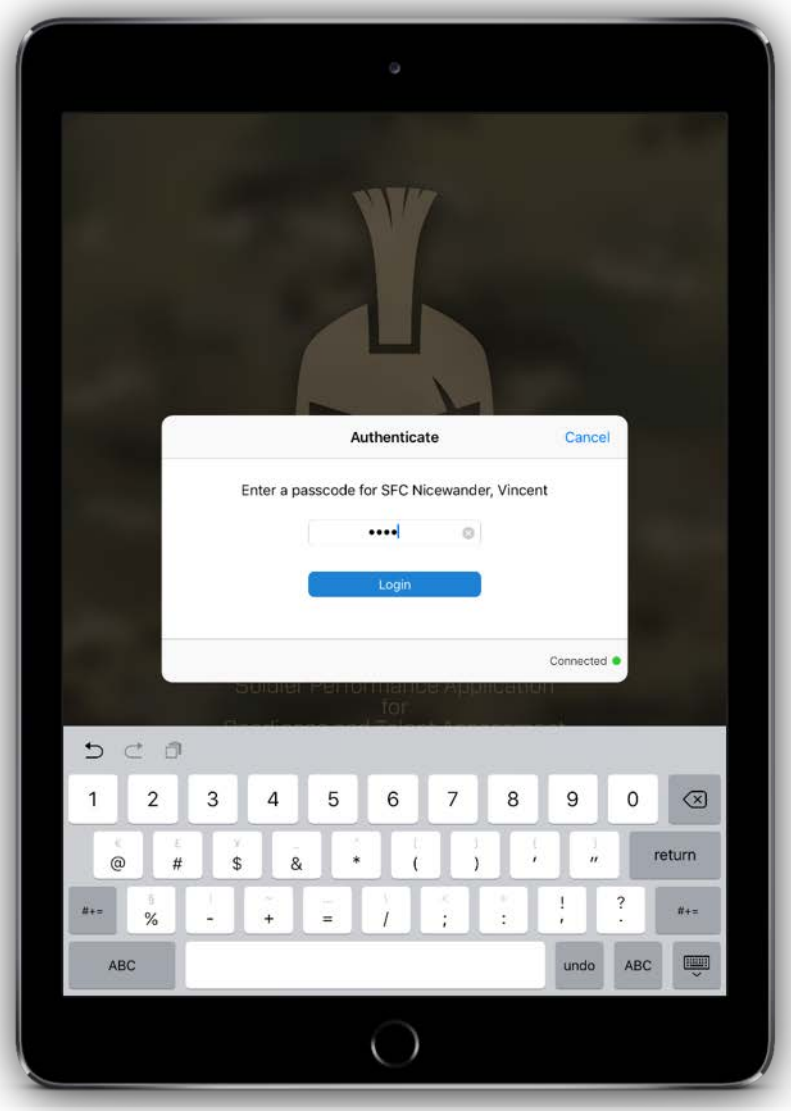
PICKING AN INSTALLATION



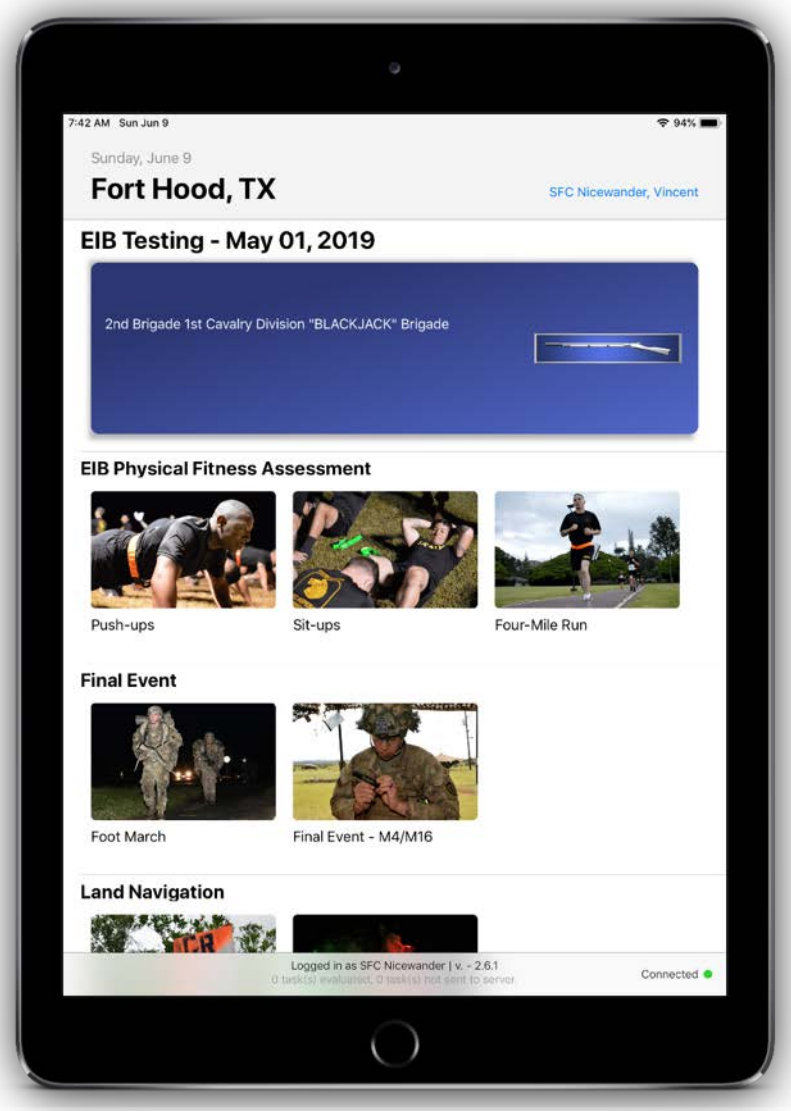
PICKING AN INSTANCE


Back
Choose Your Name to Login
SSG Fuller, Jaxx
SPC Gainey, Saphia
SSG Gaylor, Hailo
SPC Hackathorn, Basem
SSG Jarek, Walter
SSG Monistere, Aneres
SSG Mumphery, Javar
SFC Nicewander, Vincent
SSG Pippa, Novaley
SGT Pontius, Theia
SFC Rider, Jim
SFC Silva, Jake
SGT Silva, Gus
Connected
U.S.ARMY
NORTHROP GRUMMAN

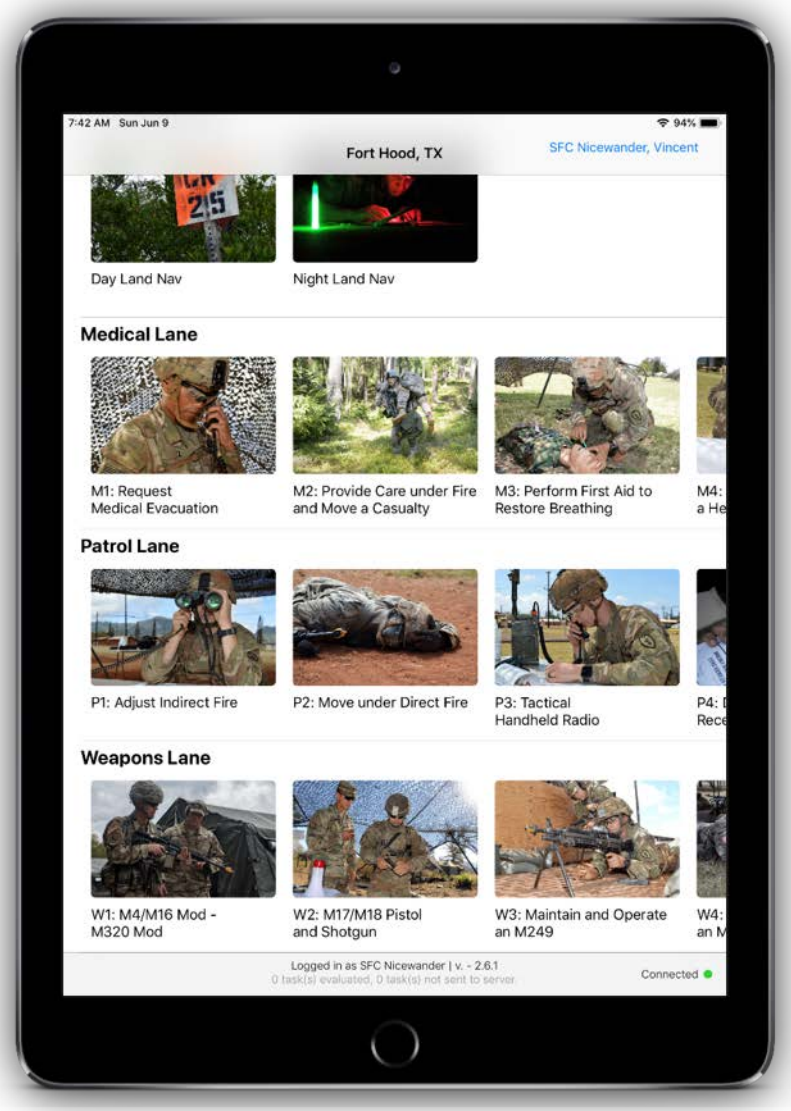
LOGGING IN



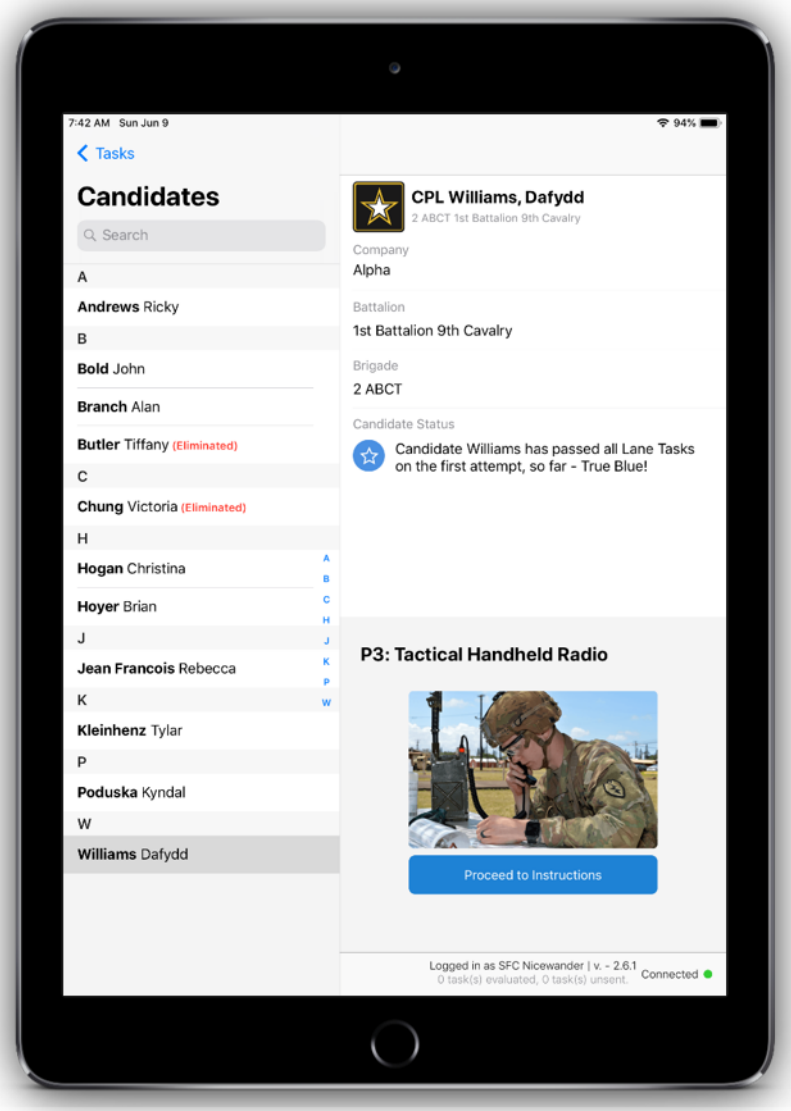
ENTERING A PASSWORD



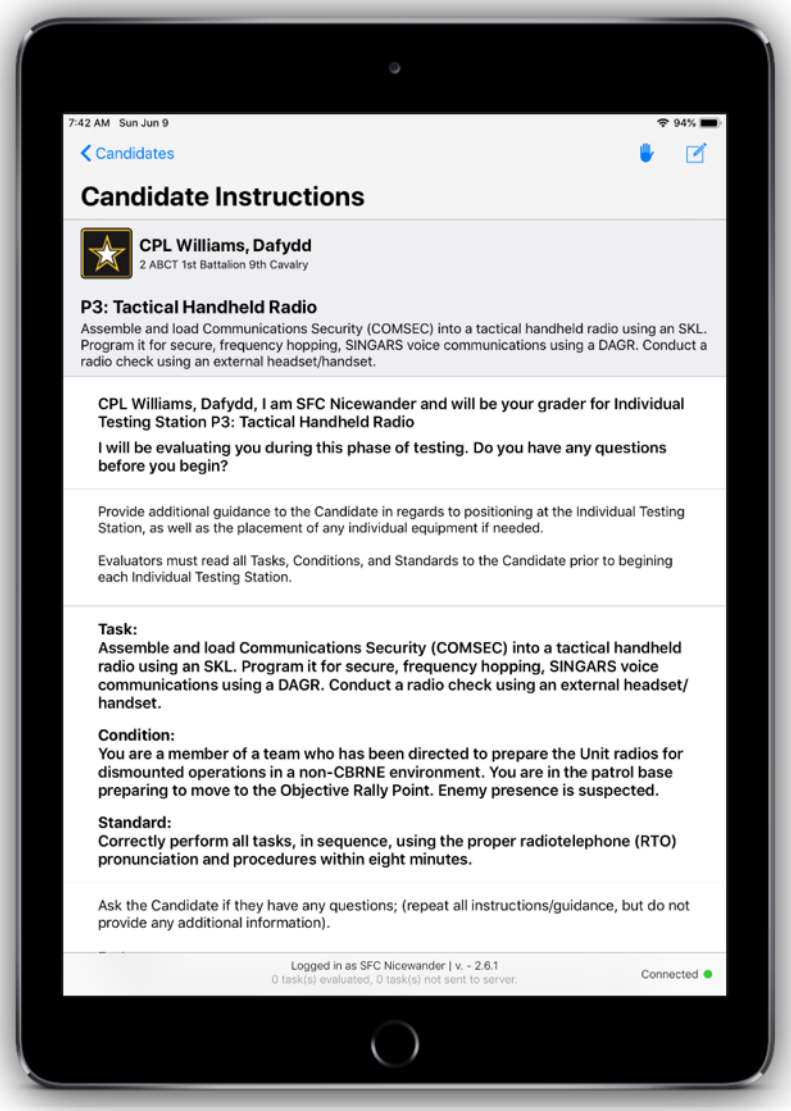
INSTANCE FRONT PAGE



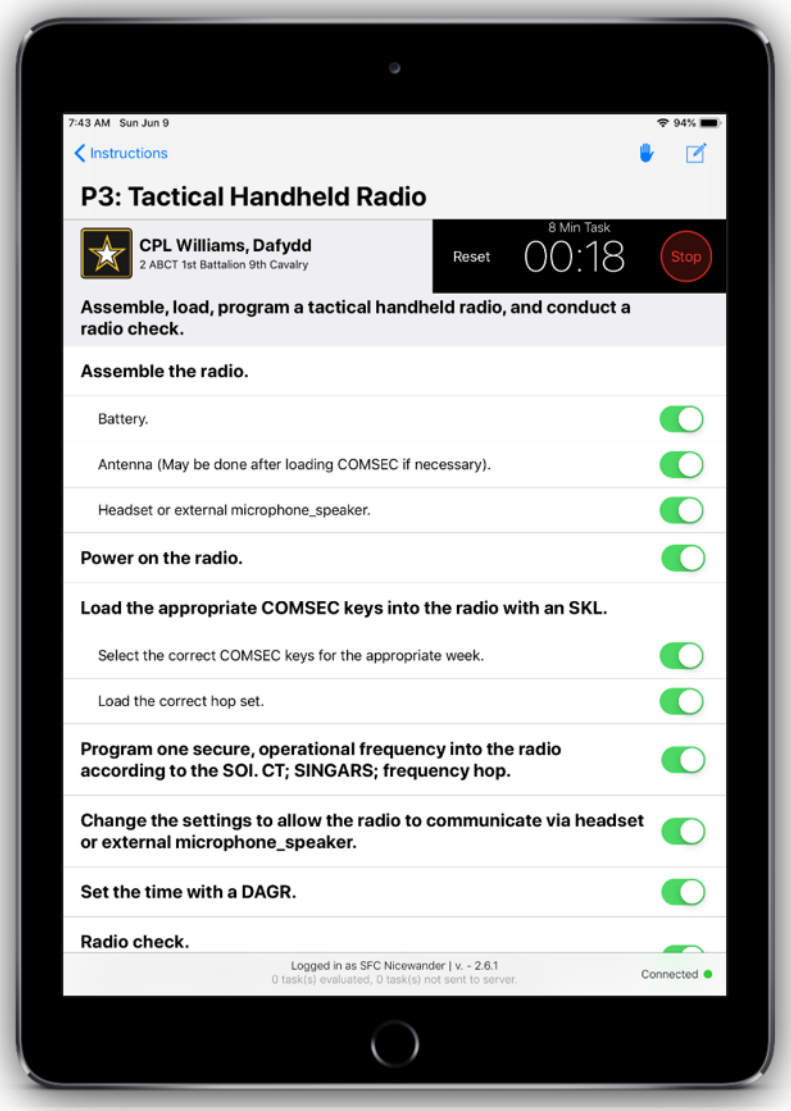
ASSESSMENT OPTIONS



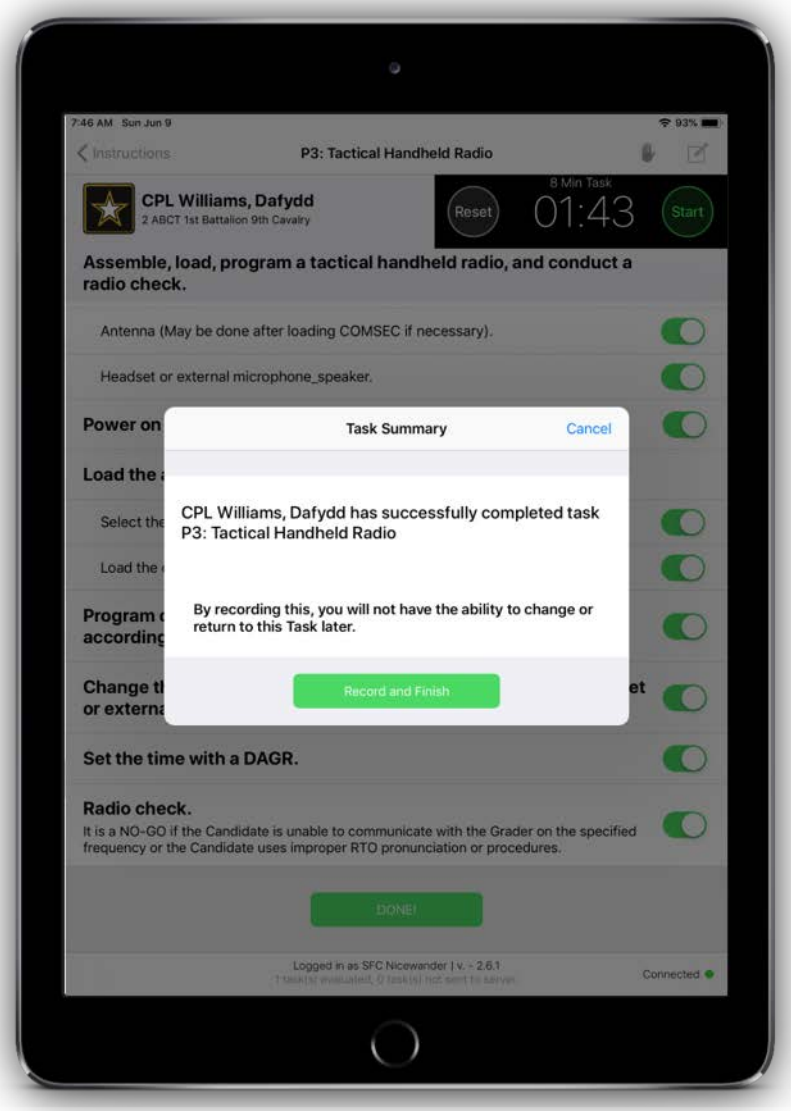
CANDIDATES LIST



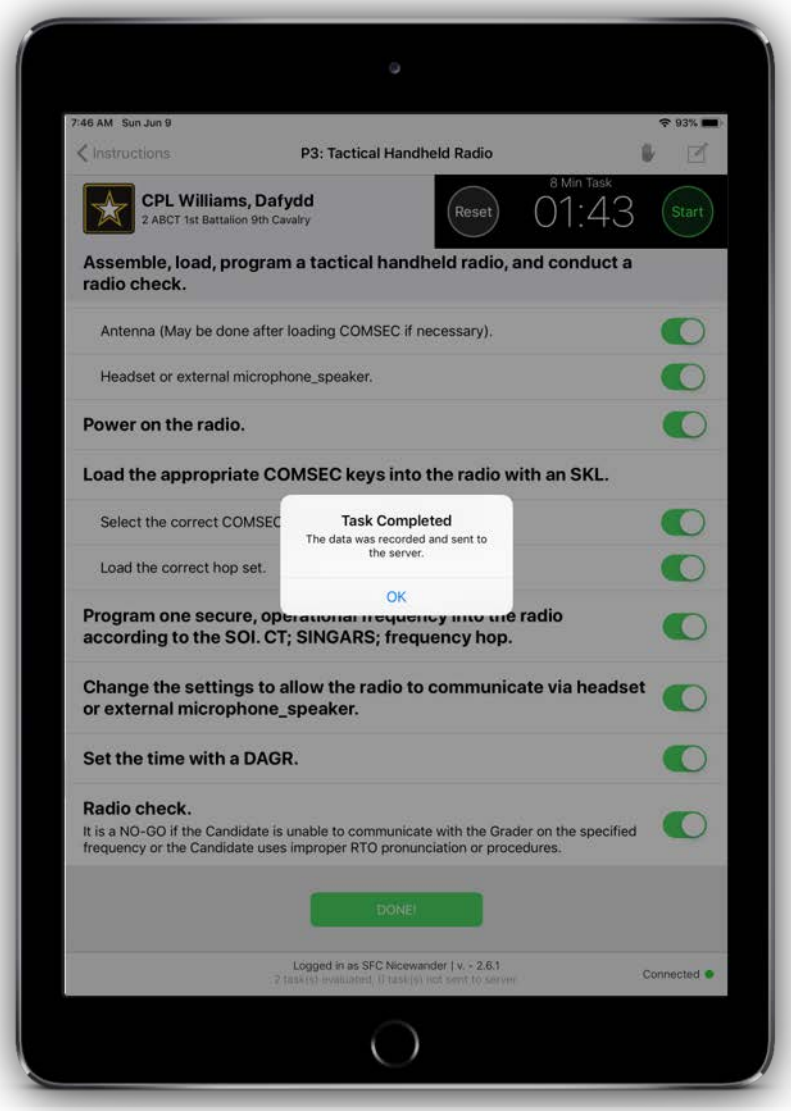
STANDARDIZED CANDIDATE INSTRUCTIONS



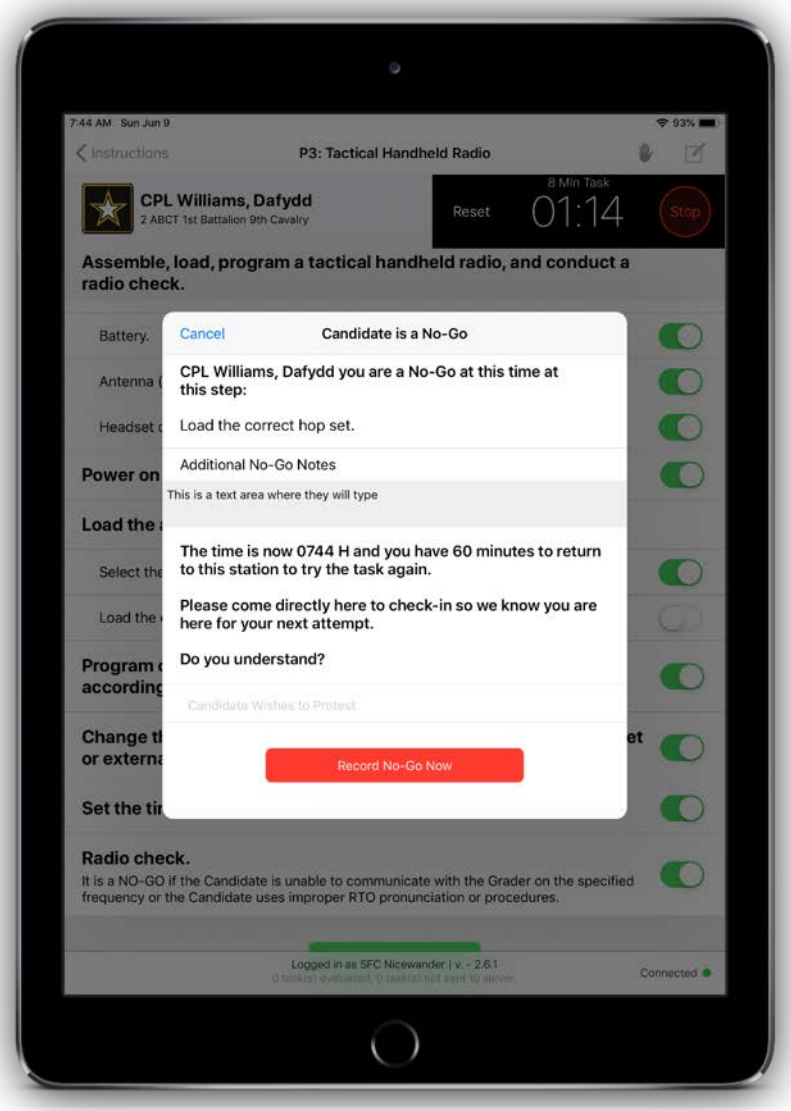
TASKS AND TIMER



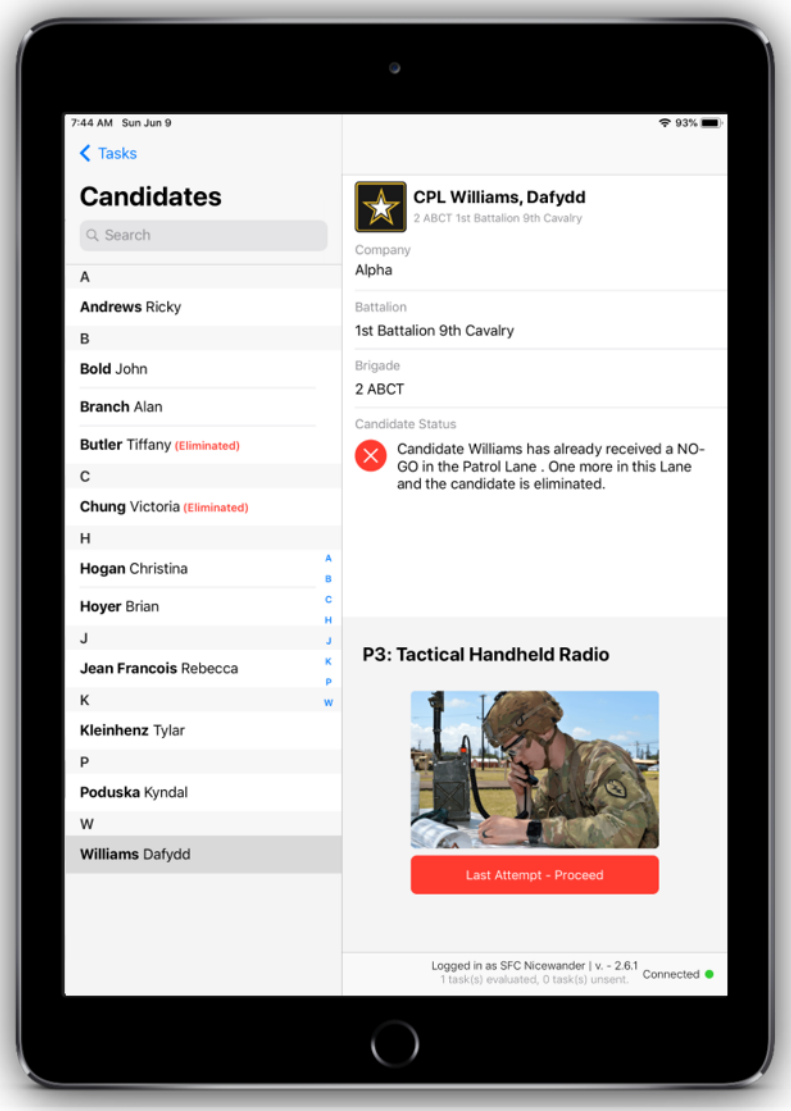
RECORDING A GO



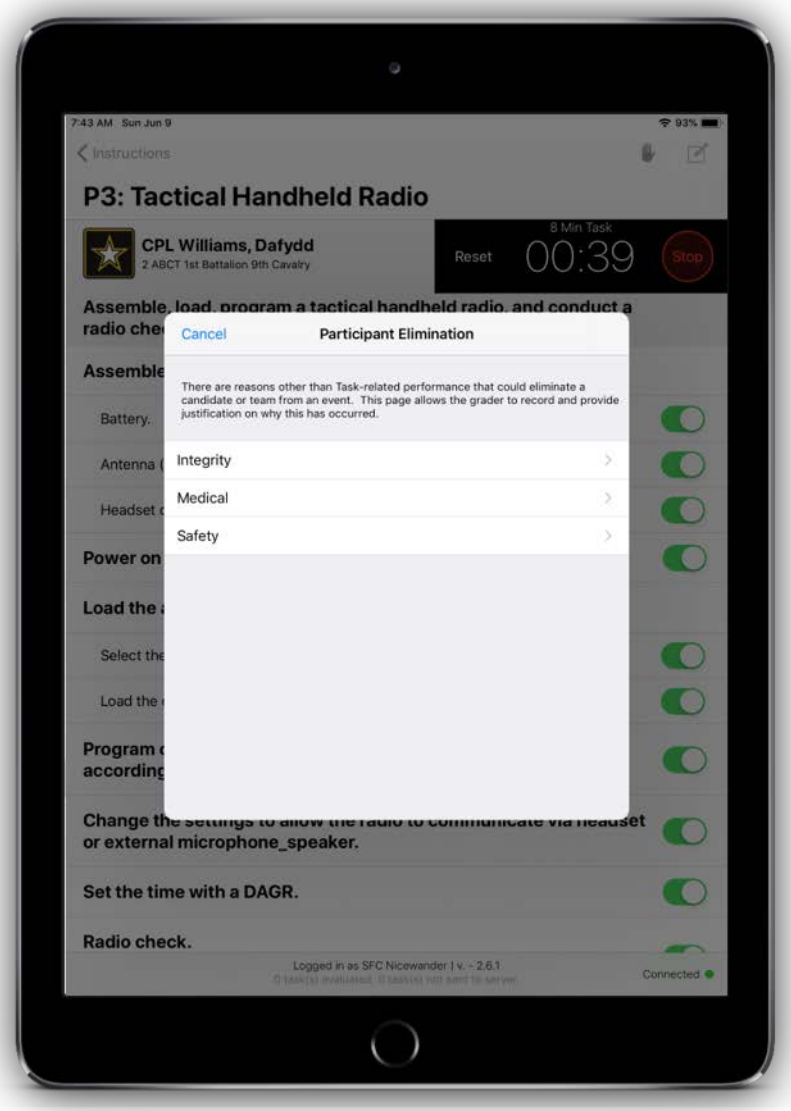
VERIFYING THE GO



RECORDING A NO-GO



UPDATED CANDIDATE INFORMATION



ELIMINATION OPTIONS

APPENDIX C

EXAMPLE OF THE TIME DATA COLLECTION SHEET

P4: Defense Advanced GPS Receiver (DAGR) Operations

Candidate (C)		Duration (D) [mm:ss]	2nd Attempt (2nd) [X if Yes]	Reason for NOGO (NG) [Task step]
e.g.	13	3:27	X	2.a.2

C	D	2nd	NG	(15 minutes)
				1. Load current month's and next month's crypto keys into the DAGR. 2. Restrict DAGR to use only secure satellites. 3. Enter mission duration. 4. Mark present position as a waypoint. 5. Enter three waypoints given. 6. Create a route using all four waypoints in the correct order. 7. Place DAGR in Average Mode.