Research Report 1983

Developing Performance Measures for Manned-Unmanned Teaming Skills

John E. Stewart and Martin L. Bink
U.S. Army Research Institute

Courtney R. Dean
Aptima, Incorporated

Troy Zeidman
Imprimis Incorporated

February 2015

United States Army Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

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

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved:

MICHELLE SAMS, Ph.D.
Director

Research accomplished under contract
for the Department of the Army

Aptima, Incorporated

Technical Review by

Peter Schaefer, U.S. Army Research Institute
Shala Blue, U.S. Army Research Institute

NOTICES

DISTRIBUTION: This Research Report has been submitted to the Defense Information Technical Center (DTIC). Address correspondence concerning ARI reports 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 his 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

1. REPORT DATE *(DD-MM-YYYY)* February 2015	**2. REPORT TYPE** Final	**3. DATES COVERED** *(From - To)* November 2012 - April 2014

4. TITLE AND SUBTITLE Developing Performance Measures for Manned-Unmanned Teaming Skills	**5a. CONTRACT NUMBER** W5J9CQ-11-D-0004 **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER** 622785
6. AUTHOR(S) John E. Stewart and Martin L. Bink; Courtney R. Dean; Troy Zeidman	**5d. PROJECT NUMBER** A790 **5e. TASK NUMBER** 225 **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Aptima, Incorporated 12 Gill Street, Suite 1400 Woburn, MA 01801	**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 / Mail Stop: 5610) Fort Belvoir, VA 22060-5610	**10. SPONSOR/MONITOR'S ACRONYM(S)** ARI **11. SPONSOR/MONITOR'S REPORT NUMBER(S)** Research report 1983

12. DISTRIBUTION / AVAILABILITY STATEMENT; Distribution Statement A.
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Contracting Officer's Representative and Subject Matter POC: Dr. Martin L. Bink

14. ABSTRACT
Manned-unmanned teaming (MUM-T) combines armed helicopters and unmanned aircraft system (UAS) in integrated tactical teams. Previous research identified critical skills required for successful MUM-T that many UAS operators lacked the training to execute. The current research effort was intended to develop performance measures for training-critical MUM-T skills that can support effective training. The refined performance measures were developed through interviews with senior UAS operators and at workshops consisting of scout-attack pilots and UAS operators with MUM-T experience. The product was 36 behavior-based measures that indicated good, average, and poor performance on MUM-T skills. The majority of measures (i.e., 26) were determined to have high validity and high utility, and only two measures were determined to have low validity and low utility. The measure determined to have low utility may reflect uncertainty as to the current role of UAS aircrews in MUM-T operations. The MUM-T measures can be used to effectively assess training outcomes in live training and virtual training.

15. SUBJECT TERMS
Unmanned Aircraft, Performance Measurement, Teams - Manned-Unmanned, Training-Live-Virtual

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dorothy Young
Unlimited	Unlimited	Unlimited	Unlimited Unclassified	57	**19b. TELEPHONE NUMBER** *(include area code)* 703-545-2316

Research Report 1983

Developing Performance Measures for Manned-Unmanned Teaming Skills

John E. Stewart and Martin L. Bink
U.S. Army Research Institute

Courtney R. Dean
Aptima, Incorporated

Troy Zeidman
Imprimis, Inc.

Fort Benning Research Unit
Scott E. Graham, Chief

United States Army Research Institute
for the Behavioral and Social Sciences

February 2015

Approved for public release; distribution is unlimited.

ACKNOWLEDGMENT

We are indebted to the support we received from the staff of Training Capability Manager, Unmanned Aircraft Systems (UAS) at Fort Rucker, AL. Experienced UAS operators provided invaluable input to us in one on one interviews as to the observation and measurement of critical skills for manned-unmanned aviation teaming. We also would like to acknowledge the dedicated personnel from an air reconnaissance squadron consisting of both OH-58D helicopters and RQ-7B unmanned aircraft, for their participation in two sets of focus groups where our draft performance measures were meticulously reviewed, refined, and then validated as to measurable content. Finally, we would like to thank Dr. Peter Schaefer and Dr. Shala Blue for critical comments on earlier drafts of this report.

DEVELOPING PERFORMANCE MEASURES FOR MANNED-UNMANNED TEAMING SKILLS

EXECUTIVE SUMMARY

Research Requirement:

Accurate measurement of human performance is a prerequisite for the development of an effective training program. Decades of previous research has demonstrated the importance and usefulness of observer-based measures for assessing both individual and team performance. Manned-unmanned teaming (MUM-T) is an aviation collective activity that requires close communication and coordination between scout-attack helicopters and unmanned aircraft systems (UAS). As the role of UAS in combat missions continues to evolve, new tools and techniques will be required to prepare UAS operators to assume more tactical responsibility and to directly interact with manned assets. In order to measure these MUM-T skills, there must be clear consensus among experts that the measurement scales describe accurately the content of the behavior in question, and that distinguish teams performing well from those performing poorly. This report describes the development and validation of a set of observable MUM-T performance measures.

Procedure:

The methodology for both the development and validation of performance measures relied on input from the experiential knowledge base of subject matter experts (SMEs). The research utilized an iterative series of workshops with SMEs consisting of active duty UAS operators and scout-attack helicopter pilots. In order to develop candidate measures, senior UAS operators were interviewed to identify the range of likely and desired behaviors that defined good and poor performance for a set of pre-defined performance indicators. Next, a group of UAS operators and scout-attack helicopter pilots reviewed and refined these candidate performance measures and their associated rating scales for appropriateness of measure type, measure wording, scale type, and scale wording. Finally, content validity and measure utility was assessed with ratings of relevance and observability provided by a separate group of UAS operators and scout-attack helicopter pilots from an air cavalry squadron specialized for MUM-T operations.

Findings:

The present research successfully demonstrated the development process of moving from skill definitions, to indicators of skill performance, and to behaviorally-anchored quantitative skill measures. A total of 36 performance measures were produced. The majority of measures (i.e., 26) were determined to have high validity and high utility, and only two measures were determined to have low validity and low utility.

Utilization and Dissemination of Findings:

The MUM-T measures developed in this effort could provide feedback for training in any UAS simulator (e.g., Universal Mission Simulator). Likewise, the measures could be used in live training exercises such as the culminating training exercise at the UAS schoolhouse at Fort Huachuca or the force-on-force exercises at the National Training Center at Fort Irwin. The results of the MUM-T measure development were briefed to U.S. Army Training and Doctrine Command Capabilities Manager for Unmanned Aircraft Systems and to U.S. Army Aviation Center of Excellence Directorate of Training and Doctrine. The results were also presented at the 2014 Interservice-Industry Simulation and Training Conference.

DEVELOPING PERFORMANCE MEASURES FOR MANNED-UNMANNED TEAMING SKILLS

CONTENTS

Page

Developing Performance Measures for Manned-Unmanned Teaming Skills

Introduction

Challenges to Training Manned-Unmanned Teaming Skills

Manned-unmanned teaming (MUM-T) exploits the advantages of the unmanned aircraft system (UAS) and the armed scout-attack helicopter. In order to be effective members of the Army Aviation manned-unmanned team, UAS aircrews must master many of the reconnaissance, surveillance, and target acquisition (RSTA) skills that are performed by pilots of reconnaissance-attack helicopters. Advanced individual training (AIT) for UAS operators focuses primarily on operation of the UAS system, and few RSTA skills are trained. Rather, RSTA skills are acquired while assigned to the operational unit. Previous research (Sticha, Howse, Stewart, Conzelman, & Thibodeaux, 2012; Stewart, Bink, Barker, Tremlett & Price, 2011) has shown that UAS operators graduating AIT lack many of the critical RSTA skills necessary for MUM-T operations. Sticha et al. were able to identify 20 training-critical skills and corresponding performance indicators for MUM-T skills that many UAS operators graduating AIT could not execute. These skills were prioritized in terms of being important to mission execution yet poorly trained or not trained at all.

MUM-T doctrine and tactics are still evolving and as of this date there has been only one unit that executed organic MUM-T missions. This Attack Reconnaissance Squadron consisted of 21 OH-58D armed scout helicopters and 8 unarmed RQ-7B Shadow UAS. As part of the Army's Aviation Restructuring Initiative, other combat aviation brigades (CABs) are establishing manned-unmanned Attack Reconnaissance Battalions. The role of the UAS operator is to perform as an integrated member of this manned-unmanned tactical team. With the increased reliance on MUM-T, there is an increased need to define, measure, and train MUM-T skills.

Aviation Performance Measurement

The development of objective measures of performance for aviation training is not a new concept (Stewart, 1985; Stewart, Dohme, & Nullmeyer, 1999). For the past two decades, performance measures have been within the capabilities of modern flight simulators. Stewart (1994) demonstrated how automated performance measures could be predefined and captured directly from the data recording system of a high-fidelity AH-64A research simulator. The system-based performance measures were comprehensive (e.g., control input, pilot's visual orientation, and aircraft state). The measures were validated by senior instructor pilots and found to significantly correlate with real-time performance ratings of a set of standard maneuver tasks. *Post-hoc* blind rankings of output graphs of several of these maneuver tasks by the same instructors showed very high concordance.

A substantial body of research has shown that both automated and precisely scaled observer-based measures can provide more objective benchmarks for assessing not only the effectiveness of simulators but also effectiveness training programs as well (Benton, Corriveau, & Koonce, 1993; Nullmeyer & Rockway, 1984). Benchmarked observer-based measures have been shown to be superior to the current subjective criteria used to train student pilots (e.g., daily flight grades and checkride put-up grades). Empirical evidence further suggests that flight training grades alone may not be a valid predictor of future aviator performance in the unit. For instance, Bales, Rickus, and Ambler (1973) followed U.S. Navy aviators to their fleet air groups after graduation, and confirmed that performance in flight school did not predict performance in mission-oriented skills nor the abilities required to perform successfully in the field.

Stewart, et al. (1999) saw the development of performance measures keyed to mission-relevant skills for Army aviation as a critical need to determine if skills trained in the schoolhouse transferred to performance in the field. At that time, the authors concluded that developing benchmarks for measuring unit-level performance would be very costly and difficult. Developing benchmarks would involve assessing such tasks as gunnery, troop insertions, lift operations, and coordination of battle plans with other units. This conclusion, however, was made prior to the advent of shared virtual environments. More recently, a set of measures was developed for aviation collective-mission skills that were based on both trainers' observations and simulator system data (Seibert, Diedrich, Stewart, Bink, & Zeidman, 2011).

Aviation collective performance measures for manned attack-weapons teams and scout-weapons teams were developed and validated to assist observers and trainers during collective training in virtual simulation (Bink, Dean, Ayers, & Zeidman, 2014; Seibert, et al., 2011). The measures were based on behaviorally-anchored rated scales that indicate poor, average, and good performance and were implemented in two modalities. First, all 109 empirically-validated measures can be assessed as observer-based ratings with behaviorally-anchored scales. These observer-based measures can be recorded using a mobile-device application. Trainers can use the application to record performance as simulation training occurs. Second, a select set of 13 measures can be assessed using the simulator data stream. That is, data packets published from the simulators (i.e., Distributed Interactive Simulation) are processed to automatically compute values on the measures. For example, an automated performance measure can be used to determine whether an aircrew violated restricted airspace. This "system-based" tool displays and saves performance measures (including audio) in near-real time. The trainer can monitor performance as simulation training occurs or can use the measures and displays as feedback during an after-action review.

Thus, the technology for objective performance measurement for Army aviation, both manned and unmanned, is maturing and can address the challenges of MUM-T training. In MUM-T training, many cognitive and procedural skills must be mastered, team performance must be assessed, and accurate feedback must be provided so that trainees can improve their performance. The instructor must likewise have knowledge of which skills deserve the most attention, which have been well-learned, and which seem to need frequent refreshment. Trends over time are also critical to determine how long it takes the teams to become proficient and how often they should practice to sustain proficiency. Subjective ratings and *post hoc* evaluations would not be sufficient to track the status and progress of complex cognitive and procedural skills. A toolset consisting of behaviorally-anchored performance measures of known content validity would be required to evaluate and track MUM-T performance, such as the measures designed for manned aviation collective performance. These real-time metrics must have performance descriptors relevant to the task at hand, must be commonly understood by leaders and trainers, and must be based upon overt behavior that can be observed in the appropriate setting.

Technical Objectives

The purpose and technical objectives of the current research are as follows:

- Identify candidate performance indicators for MUM-T.
- From these, develop prototype performance measures that are both relevant to specific critical skills, and tied to overt behaviors that can be observed.
- Benchmark prototype performance measures to rating scales.
- Determine content validity of these measures.
- Refine performance measures and format for implementation in a mobile application.

Determining the utility and content validity of these metrics will rely heavily upon experienced subject-matter experts (SMEs), at least some of whom must have had MUM-T experience in combat. To accomplish these objectives, both UAS operators and pilots of Army attack and scout helicopters must be utilized.

Measure Definition and Development

The methodology for measure development combined the knowledge base of SMEs with performance-measure products from previous research efforts (Seibert, et al., 2011; Sticha, et al., 2012) and established psychometric scaling practices using behaviorally-anchored rating scales (BARS). Leveraging these research products, an iterative series of workshops with SMEs was used to develop performance measures. The process ensured that SMEs would work collaboratively with scientists to reveal insights and drive the creation of the measures (e.g., Seibert, Diedrich, MacMillan, & Riccio, 2010).

Performance-measure development used in this effort was adapted from an established methodology that blends review of background and published materials with a series of SME workshops featuring facilitators with backgrounds in psychometric theory (MacMillan, Entin, Morley, & Bennett, 2013). The first step in this effort was the identification of behavior-based performance indicators for critical skills that support objective assessment of MUM-T performance in pilots and operators. The performance indicators were analyzed and refined by researchers from those indicators identified in previous ARI research (Sticha, et al., 2012). The second step was to develop candidate measures from the performance indicators based on SME input. Performance indicators were addressed one at a time in a series of one-on-one SME interviews. During each interview, the SME was asked to provide specific behavioral examples of performance relevant to each performance indicator. The third step was to refine the candidate measures using SME and researcher review. This step used a round table discussion to produce consensus. The combination of individual and group discussions provided more opportunity for individual perspective because opinions were offered in a less competitive environment.

The results of this process were measures and descriptors (i.e., anchors) for each performance measure that were easily understood and used by trainers, unit leaders, and other SMEs who may not be familiar with the theory upon which the measures were based. In addition, development focused on establishing an assessment system which maintains sensitivity to the evolution of team processes as teams become more efficient, coordinated, and cohesive over time. The ability to track performance over time would allow for more complete feedback by the instructors and the ability to identify training gaps by training developers.

Method

Participants. Active duty UAS operators and scout helicopter pilots, as well as manned and unmanned aircraft instructors, were utilized in the development of the MUM-T performance measures. The 18 SMEs were a mix of seven UAS Instructor-Operators (six active-duty senior noncommissioned officers and one civilian), eight UAS operators (enlisted and junior noncommissioned officers) and three scout-attack helicopter pilots, all warrant officers (CW3 & CW4). Participants' most recent platform operation experience included Medium UAS (e.g., RQ-7B), OH-58D and AH-64D helicopters. All had recently been deployed in combat to Afghanistan. Most had had MUM-T experience during deployment. This mix of SMEs helped ensure that measure development would be based on a variety of viewpoints, which was an important consideration given the newness of MUM-T operations.

Procedure.

Performance indicator identification. Materials from the previous ARI research efforts were used to develop initial performance indicators. A recent ARI analysis of MUM-T training needs (Sticha, et al., 2012) yielded a set of 20 MUM-T skills and 140 performance indicators critical for training UAS operators. Performance measure development efforts leveraged these indicators as well as pre-existing measures relevant to scout-reconnaissance missions for aviation collective training (Seibert, et al., 2011). As a result, the measure-development process began with the research team identifying a logical mission event flow to order the performance indicators. This mission flow was applied to the performance indicators to help workshop SMEs maintain contextual awareness during knowledge elicitation and draft measure review. Thus, the selected performance indicators served as stimulus materials for SMEs to provide feedback from which performance measures could be developed.

Candidate measure definition. Following production of a mission-ordered performance indicator list, more detailed information was sought from the group of seven UAS Instructor-Operators to create behaviorally-anchored performance measures. For each performance indicator, the questions involved specific behaviors related to quality of performance. These needed to be determined in order to create benchmarks with appropriate rating scales. In a series of four mostly one-on-one interviews (one to two hours each), the performance indicators were discussed and SMEs identified explicit behaviors that are representative of good or poor

`

performance for each indicator. Using individual interviews was thought to be a more thorough and efficient method, compared to group sessions, for obtaining detailed information required for the development of behaviorally-anchored measures and scales (Macmillan, et al., 2013).

During the interviews, a variety of questions were asked to obtain information describing personnel most responsible for each performance indicator, to elicit behavioral anchors relevant to each of the performance indicators, and to determine, from the perspective of the SMEs, the appropriate type of measures to develop, based on each performance indicator. The following example represents a small set of the types of questions asked during the individual interviews:

- What might a member of the flight say or do to indicate good/average/poor performance for this performance indicator?
- What would cause an operator or flight team to do well or poorly at this performance indicator?
- Does the operator interact with other crewmembers, the ground, or their tactical operations center for this performance indicator?
- In what situations during this step of the mission could a person be observed performing well or poorly for this performance indicator?
- What specific tools/systems help the flight accomplish this performance indicator?

In the course of the interviews, detailed notes were taken and efforts made to log direct quotes as often as possible. Just as it is essential to have multiple note takers in a single interview, it is essential to obtain multiple perspectives on each performance indicator. A single SME may only be able to provide a partial description of the situation, or may provide a perspective not shared by others. By recording notes from several members of the research team on perspectives and descriptions provided by multiple SMEs on each performance indicator, it was more likely that the resulting performance measures were conceptually accurate.

The information gathered during the interviews was used in post-workshop analysis to develop tentative sets of behaviorally-anchored performance measures. This process involved taking each performance indicator and the associated notes to create measures with behavioral anchors that defined degrees of good and poor performance. BARS scales were based on the Critical Incident Technique developed by Flanagan (1954). Thus, one performance indicator could have one or more measures associated with it, and these measures could describe observable behaviors for either individual roles or the entire flight team. Ultimately, this process provided a set of candidate measures that could be used together or in separate elements depending on the specific evaluation criteria. These candidate measures were then presented to a second group of SMEs for the group-based workshop that sought to refine the draft performance measures.

Candidate measure refinement. The candidate measures derived from individual interviews were reviewed by 11 SMEs from a recently-deployed aviation unit that specifically executed MUM-T missions. The goal of the four-hour workshop was to conduct a detailed review and to modify the set of draft performance measures. SMEs were asked to evaluate the material to ensure that performance indicators and performance measures were operationally

`

relevant, as thorough as possible given the mission context, and appropriately worded using the experts' language and terminology.

During the four-hour measure-review workshop, each performance measure was reviewed with respect to the following criteria:

- Measure type (e.g., scale, yes/no, checkboxes; system-based vs. observer)
- Measure wording
- Scale type
- Scale wording

SMEs were also asked if there were additional measures that needed to be developed to fill any gaps in the measurement framework or if there were measures that needed to be completely removed. All measures were then revised according to SME input. The measures were also formatted with the appropriate BARS.

Results and Discussion

The measure-development process resulted in a set of 36 measures that represented the viewpoints of a range of MUM-T SMEs and that were formatted for actual use. The measures were anchored with verbal descriptors of good, average, and poor performance. The refined set of measures then served as the basis for validation.

Performance indicator identification. The identified performance indicators were structure in a way that facilitated focus group discussion through a hypothetical mission timeline. In general, the indicators represent critical tasks and interactions among team members that require proper execution for successful mission completion. Most importantly, these performance indicators represented specific opportunities to observe measureable behavior during the course of a mission. Performance indicators, then, represent both task outcomes and the *processes* used to achieve a given outcome. The assessment of process is important because the efficiency of team interaction is a hallmark of team performance (e.g., Ilgen, 1999).

The general format of a performance indicator is a phrase or sentence that begins with an action verb that focuses on an observable behavior. For example, one indicator from this effort read: *Transmit a SPOT report IAW SALT-W format (IAW TC 1-248)*. This indicator was extracted from SME descriptions of two Skill Statements: *Utilize standard execution commands to initiate attack*; and *Transmit information about the location of threat forces, terrain, and obstacles that influence operations*) from Sticha, et al. (2012).

The performance indicators were formatted in a spreadsheet to show their hierarchical dependencies. The spreadsheet numbered each indicator and identified the personnel most likely to exhibit the performance indicator. This list was also used to organize the development of the measures derived from the indicators. The list was organized according to an operational timeline with mission phases serving as major segments. A total of 84 performance indicators were developed for 16 mission phases. The full listing of indicators appears in Appendix A.

Candidate measure definition. At the conclusion of the individual interviews, notes were compiled and organized in order of the performance indicator list. Once the full set of notes

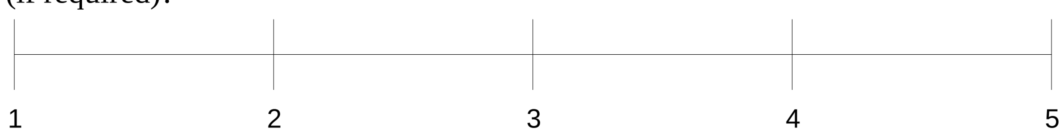
`

was organized, each performance indicator was reviewed and modified to form a measure stem that referred to the task or objective characterized. For instance, using the performance indicator, *Transmit a SPOT report IAW SALT-W format (IAW TC 1-248)*, the measure stem was written as *Does the aircrew send a SPOT report to the supported ground unit upon target detection (if required)?* Stems, scale types, and scale anchors for each performance measure were then developed. Scale types were determined based on the nature of the question and the available information to assess it. In the case of tasks or procedures where completing the action is the only observable behavior and there were no degrees of differential performance, a dichotomous (i.e., “yes”/”no”) measure was used. Other measures assessed tasks that follow a set of regimented procedures or verbal messages that must be followed the same way every time. These tasks were assessed with a checklist of all appropriate steps completed.

Most candidate performance measures derived from the interviews used a Likert-type BARS format. The BARS incorporated three verbal anchors representing good, average, and poor performance. The anchors ranged from 1 (indicating poor performance) to 5 (indicating exceptional performance). The identified key descriptors were formatted into a draft performance measure (see Figure). The three anchors depict varying levels of quality or completeness, which, by design, is meant to achieve higher levels of inter-rater reliability and reduce subjectivity in making ratings.

Performance Indicator: Transmit a SPOT report in accordance with SALT-W (size, activity, location, time, what I will do next) format.

Does the aircrew send a SPOT report to the supported ground unit upon target detection (if required)?



Aircrew does not communicate detection of target

Aircrew announces detection supported ground unit; does not transmit target information

Aircrew sends proper SPOT report to supported ground unit

Figure. Example of candidate performance measure derived from a performance indicator.

It is important to note that when behaviors could not be distinguished among several performance indicators, multiple performance indicators were covered in one measure. No assumptions were made regarding the intent of SME descriptions. In instances where confusion about SME intent occurred, references such as Field Manuals were reviewed for clarification. If no resolution could be found from published documents, notes were attached to the indicator specifying what clarifications to request during the forthcoming measure review workshop. At the end of this step of the development process, there were 45 candidate measures.

Candidate measure refinement. During the SME review workshop, modifications to the measures were made in real time. While many measures were modified during the workshop, some were deemed satisfactory with no edits and remained unchanged. The original set of 45 candidate measures was reduced to 36 measures. Some measures were culled because they were redundant or reflected behaviors that were not part of a typical mission. Several of the measures were merged with others where tasks and behaviors were deemed by the SMEs to be complementary or indistinguishable. The final set of refined performance measures is presented in Appendix B.

Content Validation of Measures

To ensure that the measures developed during the multi-step procedure were useful for assessing MUM-T performance, the content validity of each performance measure was assessed using expert ratings from the manned and unmanned aviation communities. Using a pencil-and-paper questionnaire, each rater was asked to provide input as to whether each performance measure was *observable* during a MUM-T training event and was *relevant* to the MUM-T mission. These two ratings were important dimensions for the determination of both utility and validity of the measures. Obviously, the measures must be both observable and relevant to MUM-T missions in order to be valid. Also, any measure deemed "not observable" by the raters would be unusable by an observer in a training event, i.e., low utility. Likewise, any measure deemed "not relevant" would similarly serve only as a distraction to a trainer trying to gather meaningful performance data.

Method

Participants. Nine OH-58 pilots and 10 RQ-7B operators participated in one of two workshops. All participants were from two active-duty Air Cavalry Troops of an Attack Reconnaissance Squadrons located within the United States, and each troop participated in separate workshops. One RQ-7B instructor-operator who participated in the first workshop was also present for the second but did not complete the questionnaire twice. Between the two troops, experience with MUM-T integration varied. One troop operated with UAS platforms as organic to the scout company, whereas the other troop was less integrated.

Materials and Procedure. A questionnaire was developed to gather utility data for each measure. Respondents were asked to rate each performance measure on two dimensions: relevance to MUM-T operations and degree of observability by an instructor or trainer during MUM-T exercises. Each response was given on a four-point Likert-type scale (i.e., *Strongly Disagree*, *Disagree, Agree,* and *Strongly Disagree)*, and respondents simply circled the preferred option on the questionnaire form. The questionnaire can be found in Appendix C. Prior to questionnaire distribution, participants were given a brief explanation of the effort and the objectives of the questionnaire. Questions and clarification were encouraged before and during the questionnaire. Each participant completed the questionnaire individually and returned the form to the researchers when finished.

Results and Discussion

For each measure, separate chi-square analyses were used for each rating (i.e., relevance and observability) to determine the "level" of each dimension. For example, if the statistically greatest proportion of responses for the relevancy rating was "Agree," then the measure would be deemed "relevant." The alpha-error for the statistical tests was set at 5% with 3 degrees-of-freedom. The measures were then categorized based on their levels of relevance and observability. One category was composed of measures that were determined to be both relevant and observable. A second category contained those measures that were determined to be relevant but not observable. The final two categories contained (a) measures that were *not* relevant but were observable and (b) measures that were *not* relevant and not observable. These four categories indicated the degree of content validity (i.e., relevance) and utility (i.e., observability).

The analyses of response frequencies for the relevancy ratings showed that 30 of the 36 measures determined to be relevant for UAS operators in MUM-T missions. The ratings on the remaining six measures failed to reach statistical significance, which indicated lack of agreement among the raters. As a consequence, these measures could not be categorized as "relevant" or "not relevant." The response frequencies and resulting chi-square statistics are presented in Appendix D. An inspection of Appendix D clearly shows that the numerical majority of ratings on these six measures were "relevant" (i.e., Agree or Strongly Agree). However, these six measures were assigned a separate "relevancy" category from the other measures to distinguish the lack of rater consensus.

The six measures that could not be assigned a relevance category (i.e., non-statistically-significant ratings) were closely related to the engagement of targets during fires missions. The implications for these measures were that the UAS aircrew would have the lead role in a decisional process. It is unlikely that all UAS operators would be trained to execute the lead role. These measures were previously identified as skills UAS operators graduating AIT were poorly prepared to perform (Sticha et al., 2012).

The analyses of response frequencies for the observability ratings were quite similar to recency ratings (average $r = .73$). In fact, the six measures that failed to reach statistical significance on relevancy ratings were the same measures that failed to reach statistical significance on observability ratings. One additional measure failed to reach statistical significance on observability ratings (i.e., *Flight recommends course of action to ground commander*). The remaining 29 measures were clearly rated as "observable. Appendix E presents the frequencies and resulting chi-square statistics for observability ratings.

Taken at face value, the results indicated that nearly all of the measures captured an observable element of MUM-T performance. However, as previously noted, the observability ratings practically mirrored the relevance rating. This dependency was especially evident in the measures *UAS acknowledges receipt from fire direction center* and *UAS relays target direction & range to other aircraft*. Each of these measures would have overt observable communication behaviors associated with them (i.e., radio calls with the correct information). However, the observability ratings were distributed evenly across responses and nearly mirrored the relevance

`

rating. As such, it could be inferred that raters were simply rating relevancy twice as opposed to making independent ratings on observability. It could also be the case that because the ratings were made in sequence (i.e, relevance rating was made followed by observable rating for each measure), there was a carryover effect of relevance ratings on observability ratings. Another contributing factor to the dependency of relevance ratings and observability ratings could be perception that some skills are *only* relevant to manned helicopter pilots. For example, *deconfliction of airspace* was not rated as relevant to UAS and would only be considered the province of the manned aircrews. As a consequence, if UAS are not or should not be performing the skill, then there would be no opportunity to *observe* the skill.

The next step in the analysis of content validity and of utility was to categorize each measure based on the relevance ratings and observability ratings. The measures were hierarchically ranked on each rating with relevance ratings being primary. The rankings were based on the proportion of Strongly Agree responses followed, in turn, by Agree responses, Disagree responses, and Strongly Disagree responses. The intent was to assign the measures to one of four categories based on whether the measure was determined to be both relevant and observable, to be relevant but not observable, to be *not* relevant but were observable, or to be not relevant and not observable. However, because all the measures were ultimately deemed to be relevant and observable (or at least not irrelevant and not unobservable), the distinctions among categories had to be made on whether the measures were clearly relevant and observable (i.e., statistically significant), marginally relevant and observable (i.e., not statistically significant), or some combination of clarity. The results of the categorization are presented in the Table with overall rankings and simple rankings retained to illustrate the relative validity and utility of each measure.

The largest category contained 26 measures that were rated as highly relevant (i.e., valid) and highly observable (i.e., useful). The second category contained eight measures and was more difficult to define. Three of the measures (i.e., "UAS shares sensor feed with ground unit," UAS uses proper format for indirect fire mission," and UAS proactive in executing call for indirect fire") were clearly rated as "relevant" (see Appendix D) and were mostly rated as observable (see Appendix E). However, the ratings on both dimensions for these measures were distributed across the four response categories. The other measures in this category were also rated as "relevant," but observability ratings were almost evenly distributed across responses. The final category contained the two measures for which there was no clear pattern of responses for either relevance ratings or observability ratings.

Table.
Categorization and Rankings of Manned-Unmanned Teaming Performance Measures.

Category	**Performance Measure**	**Overall Rank**	**Relevance Rank**	**Observability Rank**
High validity and High Utility				
	UAS verifies location of friendly forces	1	1	2
	UAS uses appropriate sensors	2	2	7
	UAS provides continuous reconnaissance	3	3	4
	Flight conducts collateral damage assessment	4	4	14
	UAS updates target behavior	5	5	11
	Flight actively searches for target	6	6	10
	UAS uses correct procedure for remote Hellfire launch	7	7	1
	Flight confirms hostile intent before lethal force	8	8	6
	UAS maintains positive identification of target.	9	9	8
	Flight coordinates duties after target acquired	10	10	9
	UAS conducts standardized Battle Damage Assessment	11	12	5
	UAS announces target acquisition	12	11	19
	UAS reports when contact lost	13	15	3
	UAS shares sensor feed with tactical operations center	14	13	13
	UAS sends SPOT report	15	14	22
	UAS shares sensor feed with flight	16	17	12
	UAS recognizes threats during mission	17	18	15
	Flight incorporates ISR plan	18	16	17
	UAS follows correct procedures target handover	19	20	20
	UAS provides early warnings, overwatch	20	19	26

`

Category	Performance Measure	Overall Rank	Relevance Rank	Observability Rank
	UAS correctly identifies targets during handover	21	21	16
	UAS updates flight on changes in Common operating picture	22	22	18
	UAS coordinates manned aircraft laser code	23	23	21
	UAS provides updates to ground unit	24	24	23
	Flight uses appropriate engagement scheme	25	25	27
	Flight selects appropriate weapon	26	26	25
Marginal Validity and Marginal Utility				
	UAS shares sensor feed with ground unit	27	31	36
	UAS uses proper format for indirect fire mission	28	32	28
	UAS proactive in executing call for indirect fire	29	34	33
	Flight recommends course of action to ground commander	30	27	24
	UAS proactive in airspace deconfliction	31	29	29
	UAS acknowledges receipt from fire direct center	32	28	31
	UAS deconflicts airspace before missile launch	33	30	30
	UAS relays target direction & range to other aircraft	34	33	32
Low Validity and Low Utility				
	UAS prioritizes engagement of targets	35	35	34
	UAS updates engagement priority	36	36	35

`

General Discussion

The present research successfully demonstrated the development process of moving from skill definitions, to indicators of skill performance, and to behaviorally-anchored quantitative skill measures. The final step in development of the 36 performance measures was to determine their content validity and utility. All of these steps were the result of input from experienced MUM-T SMEs. This does not mean that all of these performance measures are necessarily expected to have predictive validity. Field testing of the measures during live exercises or virtual exercise was outside the scope of the present research but should be done to obtain additional evidence of validity. All in all, the measure-development efforts were successful insofar as specific MUM-T measures were produced that were considered by experts to be most relevant to the mission and most likely to be observed in the course of training or mission execution.

The patterns of ratings and the apparent dependency of the ratings provided in the validation phase of measure development indicated that the perceived role of UAS in MUM-T missions was still support for the manned component. More specifically, the ratings suggested that the role of the UAS aircrew in MUM-T combines both traditional ISR skills, such as aerial observation, reconnaissance, target detection and identification, and reporting, with tactical RSTA activities such as overwatch, providing information on changes in the common operating picture, and handing-over and designating targets for armed helicopters. This perceived role was more active than a traditional ISR role (Stewart Roberts, & Bink, 2012). However, it was clear that SMEs did not perceive the UAS role as being proactive and cognitive (e.g., decision making in direct and indirect fire situations). Perhaps the perceived role of UAS in MUM-T missions is based on the fact that UAS operators currently do not have the ability to execute some skills required for MUM-T. For example, both *UAS deconflicts airspace before missile launch* and *UAS prioritizes engagement of targets* were two measures that are objectively relevant to the MUM-T mission, can technically be accomplished with a UAS, and have observable components yet both measures yielded low relevance ratings and observability ratings. However, UAS operators are not often trained on these skills and currently do not execute these skills in regular combat operations (Sticha, et al., 2012).

Consequently, the measures with low relevancy ratings may be needed in MUM-T operations. However, it was not the perceived role of UAS aircrews to plan and initiate these skills at this moment in time. MUM-T doctrine and tactics are still evolving, and it is premature to project what the future UAS operator's role in MUM-T will look be. The rapid growth of applications and roles for unmanned aircraft systems in the last ten years (U.S. Army UAS Center of Excellence, 2010) makes it virtually certain that changes in manpower, personnel, and training requirements will follow. Thus, the UAS operator's role in the next five or ten years may have far different skill and training requirements than is currently the case. The state of flux of UAS organization in the Army should provide ample opportunities to explore the impact of the increasing responsibilities and rising expectations for UAS aircrews in MUM-T. The future trend should be toward increasing integration of roles with manned platforms, which should expand the responsibilities of UAS operators.

`

Regardless of the future roles of UAS in MUM-T, the measures developed in the present effort could be used to enhance training. The measures could be used in either live training or simulation training. In fact, a similar set of measures was previously developed by ARI for use in aviation collective simulation training (Bink, et al., 2014; Seibert et al., 2011). The Aviation Collective Performance Assessment Tool (AC-PAT) provides observers and trainers a means to quantify flight-team performance during collective-training exercises using similar BARS measures as were developed here. The measures are recorded in an Android tablet application. The measures are aggregated across critical elements (e.g., mission phase) to provide visual feedback to the pilots during training hot-wash or after-action review.

A similar approach could be used for the MUM-T measures. The rating scales developed here lend themselves to a measurement tool like AC-PAT. If the MUM-T measures were implemented in the AC-PAT architecture, then the measures could provide feedback for training in any UAS simulator (e.g., Universal Mission Simulator). Likewise, the measures could be used in live training exercises such as the culminating training exercise at the UAS schoolhouse at Fort Huachuca or the force-on-force exercises at the National Training Center at Fort Irwin.

As the role of UAS in combat missions continues to evolve, new tools and techniques will be required to prepare UAS operators to assume more tactical responsibility and to directly interact with manned assets. The measures developed in this effort provide an initial step to that evolution by identifying critical MUM-T skills for UAS operators and providing a means to measure and provide feedback on those skills. The results of the MUM-T measure development were briefed to U.S. Army Training and Doctrine Command Capabilities Manager for Unmanned Aircraft Systems and to U.S. Army Aviation Center of Excellence Directorate of Training and Doctrine. The results were also presented at the 2014 Interservice-Industry Simulation and Training Conference.

References

Bale, R. M., Rickus, G. M., & Ambler, R. K. (1973). Prediction of advanced level aviation performance criteria from early training and selection variables. *Journal of Applied Psychology, 58*, 347-350.

Benton, C. J., Corriveau, P., & Koonce, J. M. (1993*). Concept development and design of a semi-automated flight evaluation system (SAFES).* (Technical Report AL/HR-TR-1993-1024) Brooks AFB, TX: Armstrong Laboratory Human Resources Directorate Manpower and Personnel Research Division. (DTIC No. ADB176885).

Bink, M. L., Dean, C., Ayers, J., & Zeidman, T. (2014). *Validation and Evaluation of Army Aviation Collective Performance Measures*. (Research Report 1972). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA596907).

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*, 327-358.

Ilgen, D. R. (1999). Teams imbedded in organizations: some implications. *American Psychologist*, *54*, 129-139.

MacMillan, J., Entin, E. B., Morley, R., & Bennett, W. (2013). Measuring team performance and complex and dynamic military environments: The SPOTLITE method. *Military Psychology, 25,* 266-279.

Nullmeyer, R .T, & Rockway, M. R. (1984). Effectiveness of the C-130 weapon system trainer for tactical aircrew training. *Proceedings of the Interservice/Industry Training Systems and Education Conference, 6*. Washington, D.C. : American Defense Preparedness Association.

Seibert, M. K., Diedrich, F. J., MacMillan, J., & Riccio, G. E. (2010). Training research in the wild. *Proceedings of the Interservice/Industry Training, Simulation and Education Conference, 2010*, Arlington, VA: National Training Systems Association.

Seibert, M.K., Diedrich, F. J., Stewart, J. E., Bink, M. L., & Zeidman, T. (2011). *Developing performance measures for Army aviation collective training*. (Research Report 1943). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA544425).

Stewart, J. E. (1985, October). Learning and performance in an air-refueling part-task trainer: preliminary data analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *29*, 408-411.

`

Stewart, J. E. (1994). *Using the backward transfer paradigm to validate the AH-64 simulator training research advanced testbed for aviation*. (Research Report 1666). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA285758).

Stewart, J. E., Bink, M.L., Barker, W.C., Tremlett, M.L., & Price, D. (2011). *Training needs for the RQ-7B operators in the scout-reconnaissance role*. (Research Report 1940). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADB367652)

Stewart, J. E., Dohme, J.A., & Nullmeyer, R.T. (1999). *Optimizing simulator-aircraft mix for U.S. Army initial entry rotary wing training*. (ARI Technical Report 1092). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA361814).

Stewart, J. E., Roberts, K.R., & Bink, M.L. (2012). *Unmanned aircraft systems in the scout-reconnaissance role: perceptions of the manned and unmanned aircraft* communities. (Research Report 1956). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA563620).

Sticha, P. J., Howse, W. R., Stewart. J. E., Conzelman, C E., & Thibodeaux, C. (2012). *Identifying critical manned-unmanned teaming skills for unmanned aircraft system operators*. (ARI Research Report 1962). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA565510).

U.S. Army UAS Center of Excellence (2010). "*Eyes of the Army": U.S. Army roadmap for unmanned aircraft systems 2010-2035*. Fort Rucker, AL. Author.

`

APPENDIX A

MUM-T Performance Indicators

`

MUM-T Performance Indicators
Find/track targets (e.g., HPTs, HVTs)
Gain and Maintain Enemy Contact
ID the target
Have a positive ID on the Target
Observe the target behavior
Accurately determine target location
Assess terrain for most likely route of target
Develop and Send Common Operation Picture information to air-ground team
Transmit information about location of threat forces, terrain, and obstacles that influence operations
Transmit a SPOT report IAW SALT-W format (IAW TC 1-248
Utilize standardized radio communication and signal operating procedures
Use clear, concise verbiage, correct terminology
Relay information to all parties involved (TOC, Ground Commander)
Terminology and sequence IAW current JFIRE publications
Updates COP promptly
Report when contact is lost
Transmit imagery, sensor data, tactical situation maps, overlays, reports (e.g. SPOT reports)
Provide early warning, ambush detection, overwatch, threat identification
Identify out of the ordinary behavior/placement of vehicles
Utilize standardized radio communication and signal operating procedures
Tailor appropriate information to appropriate recipient
Provide target description information
Provide Accurate description of target to support target selection
Utilize standardized radio communication and signal operating procedures
Provide target location (i.e., direction of target in degrees and range from battle position)
Transmit information about the location and direction of the UA as it relates to the target

`

MUM-T Performance Indicators
Prioritize Engagement of Targets
Prioritize based on CDRs intent
Prioritize based on threat systems' capabilities
Inform ME of new, more dangerous targets
Select best weapon system to engage target (e.g., lethal/nonlethal, munitions effect, collateral damage assessment)
Matches weapon system to desired effects
Considers collateral damage, 2nd and 3rd order effects
Follows ROE
Provide confirmation of target prior to engagement
Applies ROE
Utilize standardized radio communication and signal operating procedures
Transmit information about method of attack (scheme of manuever, fire distribution, maneuver for attack)
Knowledge on current TSOP and CARPUT
Chooses appropriate scheme of maneuver
Positions UAS appropriately to facilitate manned platform attack
Utilize standardized radio communication and signal operating procedures
Conduct Target Handover
Utilize Standard report formats
IAW TC 1-248 task
Utilize standardized radio communication and signal operating procedures
Switch roles of laser designator and missile launch platform
Deconflicts airspace/scheme of maneuver
Clears area of fire, UA and friendly ground forces
Communicates transfer of responsibilities
Transmit target location, description, laser code, laser target line information to shooter
Confirms laser codes is same as munitions code
Effects transfer in a timely manner

`

MUM-T Performance Indicators
Utilize standard execution commands to initiate attack
Utilize Standard report formats
Terminology and sequence IAW current JFIRE publications
Ensure positive ID of target
Ensure clearance of fires
Utilize standardized radio communication and signal operating procedures
Conduct Call for Direct Fires
Confirm location of friendlies
Meet safety and employment guidelines
Meets Rules of Engagement
Utilize Standard report formats
Provide close air support 9-line briefing information to FW or RW attack aircraft including post-attack instructions
Utilize standardized radio communication and signal operating procedures
Conduct Call for Indirect Fires
Provide direction of target in degrees and range from battle position
Utilize Standard report formats
Conduct call for fire with supporting FSE by using terminology and sequence IAW JFIRE and FM 6-30, IAW TC 1-248
Utilize standardized radio communication and signal operating procedures
Demonstrate understanding of SOP for information needed
Acknowledge fire mission receipt
Deconflict Munition Trajectories (Deconflict Airspace - Big Sky, Little Bullet)
Maintain situational awareness of positions of friendly aviation assets
Identify conflicting munition trajectories and recommend lateral, vertical, or sequential contra measures
Responds to instructions to deconflict
Moves aircraft to safe area
Utilize standardized radio communication and signal operating procedures
Perform BDA
Utilize Standard report formats
Utilize standardized radio communication and signal operating procedures
Utilize joint, Army and civilian personnel recovery terminology

`

APPENDIX B

Final MUM-T Performance Measures

MUM-T Performance Measures

Definitions

Flight – references several airframes working together to accomplish mission objectives. This can be an AWT/SRT or an AH/OH and UAS or an AWT/SRT and a UAS. Flight measures assess the interaction of these airframes.

Aircrew – references the pilots or operators for a specific airframe. This can be the pilots in a manned aircraft or the aircraft operator/pilot, payload operator and mission commander. Aircrew measures assess the interaction of the "crew" manning the specific airframe.

UAS – means UAS aircraft only. This includes the interactions of the entire crew (aircraft operator/pilot, payload operator and mission commander)

Reconnaissance

Does the flight incorporate an ISR plan?

1	2	3	4	5
No deliberate pre-planned ISR plan; does not develop a hasty ISR plan prior to beginning search		Flight utilizes deliberate or hasty ISR plan; coordinates use of sensors and assets; distributes areas of observation		Flight utilizes deliberate or hasty ISR plan; coordinates use of sensors and assets; distributes areas of observation; adjusts to situation throughout

Does the flight actively search for the target (i.e. consider METT-TC factors such as "dead" spaces, avenues of approach, and key terrain)?

1	2	3	4	5
No plan for scanning the area; does not distribute areas of observation to cover tactical METT-TC considerations		Flight distributes areas of observation; does not consider all tactical METT-TC considerations		Flight uses a pre-determined search plan; maximizes distribution of areas of observation; considers all tactical METT-TC considerations

Does the UAS use appropriate sensors to search for targets (i.e. FLIR, IR, TVS, etc.)?

1	2	3	4	5
UAS does not use available sensors		UAS uses some of the available sensors		UAS maximizes use of all available sensors

Does the UAS share sensor feeds among the flight and communicate throughout reconnaissance?

1	2	3	4	5
UAS does not share sensor feeds; no cross-talk among aircraft		UAS shares sensor feeds when requested; does not demonstrate cross-talk among aircraft		UAS proactively shares sensor feeds with team; cross-talk focuses on specific observations

Does the UAS share sensor feeds with TOC (if available/required)?

1	2	3	4	5
UAS does not share sensor feeds until prompted by TOC		UAS shares sensor feed and ensures TOC can see imagery when requested		UAS proactively shares sensor feed and ensures TOC has imagery

Does the UAS share sensor feeds with Ground Unit (if available)?

1	2	3	4	5
UAS does not share sensor feeds with Ground Unit		UAS shares sensor feed with Ground Unit when requested		UAS proactively shares sensor feed and ensures Ground Unit has imagery

Does the UAS effectively recognize threats during mission execution?

1	2	3	4	5
UAS does not detect threats		UAS detects threats; classifies it as friendly or enemy; determines threat's exact location		UAS detects, classifies, and determines exact location of threatevaluates potential impact on mission; adapts accordingly

Target Detection, Identification, and Surveillance

Does the UAS announce target acquisition?

- ☐ Yes ☐ N/A
- ☐ No ☐ N/O

Does the UAS send a complete observation report (SPOT) report upon target detection (to other aircraft or to ground unit per msn requirements)?

1	2	3	4	5
UAS does not communicate detection of target		UAS announces detection; sends incomplete SPOT report		UAS sends complete SPOT report

Does the UAS maintain positive identification (PID) of the target after acquisition?

1	2	3	4	5
UAS does not maintain PID on the target		UAS maintains PID but minimal scanning around target		UAS maintains PID on the target; expands field of view to increase SA; coordinates with other aircraft to maintain PID when necessary

Does the UAS report when contact is lost (e.g., last known location, last known traveling direction, course of action)?

1	2	3	4	5
UAS does not report when contact is lost		UAS delays in reporting contact lost; includes some of the relevant information in report		UAS immediately reports contact lost; includes all relevant information; tries to regain contact

Does the flight coordinate duties after target acquisition?

1	2	3	4	5
Flight does not discuss duties upon target acquisition; UAS continues current actions		Flight discusses duties after target acquisition; UAS follows the plan		Flight discusses duties after target acquisition; UAS discusses its utilization to best accomplish mission

Does the UAS provide continuous reconnaissance?

1	2	3	4	5
UAS focuses only on the target's location		UAS verifies target composition, disposition, and intentions		UAS proactively adjust reconnaissance plan in anticipation of future actions

Does the UAS provide updates on target behavior (changes in size, composition, disposition, activities, and movement)?

1	2	3	4	5
UAS does not provide any updates after target acquisition		UAS provides target updates when prompted by teammate or supported ground unit		UAS provides updates on target as they occur or that have a potential impact on the mission

Does UAS continuously update the Flight on changes to the common operational picture (COP)?

1	2	3	4	5
UAS does not communicate changes to COP		UAS communicates changes in target but not friendly troops or effects of terrain or obstacles		UAS continuously updates COP for changes in target, friendly troops, terrain, and obstacles

Does UAS update the ground unit they are supporting with information that could influence operations?

1	2	3	4	5
UAS does not communicate changes		UAS communicates changes in target but not friendly troops or effects of terrain or obstacles		UAS continuously updates for changes in target, friendly troops, terrain, and obstacles

Does the UAS provide early warning and threat detection to supported unit?

1	2	3	4	5
UAS provides delayed, inaccurate, or misdirected information		UAS reports accurate and timely information about threat activity		UAS reports accurate and timely information on threat, and METT-TC

Does the UAS prioritize engagement of targets?

1	2	3	4	5
UAS does not discuss priority of targets		UAS discusses priority but decision does not reflect Cdr's intent or threat(s)		UAS prioritizes engagement based on Cdr's Intent and threat(s)

Does the UAS update the engagement priority as it changes?

- ☐ Yes
- ☐ No

Pre-Engagement

Does the UAS follow the correct procedures and format for a Target Handover to wingman (Voice/Laser)?

- ☐ Yes
- ☐ No

a. If applicable, which required elements were missed?

- ☐ Laser spot/code
- ☐ Target Description
- ☐ Target Location
- ☐ Attack Instructions
- ☐ Additional information as required

Did the UAS correctly identify the target(s) during target handover?

- ☐ Yes
- ☐ No

Does the flight recommend lethal and nonlethal COAs to Ground Commander?

1	2	3	4	5
Flight provides no COA to Ground Commander		Flight provides an (one) acceptable COA		Flight provides multiple, prioritized COAs based on aerial perspective

Does the flight select the appropriate weapon for desired effect on target?

- ☐ Yes
- ☐ No

Does the flight conduct a Collateral Damage Assessment (CDA)?

- ☐ Yes
- ☐ No

Does the flight confirm hostile intent prior to applying lethal force?

- ☐ Yes
- ☐ No

a. If no, why not?

☐ Flight never discusses hostile intent

☐ Flight assumes hostile intent; relies on other reports

☐ Flight determines possible hostile intent; talks themselves into it

Does the UAS verify location of friendly forces near the target to prevent fratricide?

1	2	3	4	5
UAS does not verify friendly locations in relation to the target		UAS obtains location for friendly forces in the area and has visual of friendly locations		UAS obtains locations, has visual, and confirms outside safety fan/surface danger zone

Does the UAS deconflict the airspace in preparation for missile launch?

1	2	3	4	5
UAS does not use information sources to anticipate events; does not make radio calls		UAS uses information sources to anticipate events but does not push information to rest of team		UAS uses information sources to anticipate events and proactively pushes information to rest of team

Direct Fire Engagement (Remote Hellfire Launch)

Does the flight choose and brief the appropriate engagement scheme of maneuver?

☐ Yes

☐ No

a. If applicable, which required elements were missed?

☐ Techniques.

☐ Patterns

☐ Munitions

☐ Range

Does the UAS coordinate with the manned aircraft to ensure launcher designation angles (max offset), safety fan, laser code and laser-on time requirements can be met?

☐ Yes

☐ No

a. If applicable, which required elements were missed?

☐ LDA

☐ Safety Fan

☐ Laser Code

☐ Laser-on Time

Does the UAS follow the correct procedures and format for a remote Hellfire launch (Voice/Digital/Laser)?

☐ Yes

☐ No

b. If applicable, which required elements were missed?

☐ Laser spot/code

☐ Target Description

☐ Target Location

☐ Attack Instructions

☐ Additional information as required

Indirect Fire Engagement

Was the UAS proactive in executing call for indirect fire?

1	2	3	4	5
UAS not ready for indirect fire call		UAS prepared to make indirect fire call		UAS foresaw need for indirect fires; prepared to make indirect fires call

Does the UAS use the proper format for indirect fire missions?

☐ Yes

☐ No

a. If applicable, which required elements were missed?

☐ Observer Identification

☐ Warning Order

☐ Target Location

☐ Target Description

☐ Method of Engagement

☐ Method of Fire and Control

Does the UAS acknowledge receipt, or any changes, from Fire Direct Center (FDC)?

- ☐ Yes
- ☐ No

Does the UAS relay direction of target in degrees and range from the firing unit's position to other aircraft in order to clear the airspace?

- ☐ Yes
- ☐ No

Is the UAS proactive in airspace deconfliction throughout execution of indirect fires?

1	2	3	4	5
UAS does not have SA of battlespace during the call for indirect fire		UAS de-conflicted battlespace after making the call for indirect fires		UAS knows location of all assets and de-conflicts airspace prior to the call for indirect fire

Battle Damage Assessment

Does the UAS conduct a standardized battle damage assessment?

1	2	3	4	5
UAS does not conduct BDA; assumes target is destroyed without verification		UAS evaluates target; reports BDA to engaging aircraft (or ground unit) after prompting		UAS evaluates target; proactively reports BDA to engaging aircraft (or ground unit)

APPENDIX C

Utility Analysis Questionnaire

MUM-T Utility Analysis Questionnaire
U.S. Army Research Institute

Rank: __________ Unit Role: ________________________ Platform: ______________

Dates of last deployment: __________________ Did you perform MUM-T operations? Yes No

Instructions: The following pages contain a number of Performance Measures designed to help trainers assess UAS operator performance during Manned-Unmanned Teaming operations including scout-reconnaissance and close combat attack. This questionnaire is designed to determine if these Performance Measures are indeed useful for assessing UAS operators and providing meaningful feedback. For each Performance Measure, there are two statements: "This measure is **relevant** for MUM-T training" and "This measure is **observable** during MUM-T training." For each statement, please check the box that indicates the extent to which you agree with the statement. It is OK to agree with one statement and disagree with the other. In addition to making the ratings, please feel free to write comments about the Performance Measures or response anchors.

Performance Measure	This measure is **relevant.**	This measure is **observable.**
Does the flight incorporate an ISR plan? 1 2 3 4 5 1 – No deliberate pre-planned ISR plan; does not develop a hasty ISR plan prior to beginning search 3 – Flight utilizes deliberate or hasty ISR plan; coordinates use of sensors and assets; distributes areas of observation 5 – Flight utilizes deliberate or hasty ISR plan; coordinates use of sensors and assets; distributes areas of observation; adjusts to situation throughout	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the flight actively search for the target (i.e. consider METT-TC factors such as "dead" spaces, avenues of approach, and key terrain)? 1 2 3 4 5 1 – No plan for scanning the area; does not distribute areas of observation to cover tactical METT-TC considerations 3 – Flight distributes areas of observation; does not consider all tactical METT-TC considerations 5 – Flight uses a pre-determined search plan; maximizes distribution of areas of observation; considers all tactical METT-TC considerations	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the UAS use appropriate sensors to search for targets (i.e. FLIR, IR, TVS, etc.)? 1 2 3 4 5 1 – UAS does not use available sensors 3 – UAS uses some of the available sensors 5 – UAS maximizes use of all available sensors	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐

Does the UAS share sensor feeds among the flight and communicate throughout reconnaissance? 1 — UAS does not share sensor feeds; no cross-talk among aircraft 2 3 — UAS shares sensor feeds when requested; does not demonstrate cross-talk among aircraft 4 5 — UAS proactively shares sensor feeds with team; cross-talk focuses on specific observations	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree
Does the UAS share sensor feeds with TOC (if available/required)? 1 — UAS does not share sensor feeds until prompted by TOC 2 3 — UAS shares sensor feed and ensures TOC can see imagery when requested 4 5 — UAS proactively shares sensor feed and ensures TOC has imagery	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree
Does the UAS share sensor feeds with Ground Unit (if available)? 1 — UAS does not share sensor feeds with Ground Unit 2 3 — UAS shares sensor feed with Ground Unit when requested 4 — UAS proactively shares sensor feed and ensures Ground Unit has imagery	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree

Does the UAS effectively recognize threats during mission execution? 1 — UAS does not detect threats 2 3 — UAS detects threats; classifies it as friendly or enemy; determines threat's exact location 4 5 — UAS detects, classifies, and determines exact location of threat evaluates potential impact on mission; adapts accordingly	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree
Does the UAS announce target acquisition? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree
Does the UAS send a complete observation report (SPOT) report upon target detection (to other aircraft or to ground unit per msn requirements)? 1 — UAS does not communicate detection of target 2 3 — UAS announces detection; sends incomplete SPOT report 4 5 — UAS sends complete SPOT report	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree

Measure	Relevance	Observability
Does the UAS maintain positive identification (PID) of the target after acquisition? 1 — UAS does not maintain PID on the target 2 3 — UAS maintains PID but minimal scanning around target 4 5 — UAS maintains PID on the target; expands field of view to increase SA; coordinates with other aircraft to maintain PID when necessary	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the UAS report when contact is lost (e.g., last known location, last known traveling direction, course of action)? 1 — UAS does not report when contact is lost 2 3 — UAS delays in reporting contact lost; includes some of the relevant information in report 4 5 — UAS immediately reports contact lost; includes all relevant information; tries to regain contact	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the flight coordinate duties after target acquisition? 1 — Flight does not discuss duties upon target acquisition; UAS continues current actions 2 3 — Flight discusses duties after target acquisition; UAS follows the plan 4 5 — Flight discusses duties after target acquisition; UAS discusses its utilization to best accomplish mission	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐

Does the UAS provide continuous reconnaissance? 1 2 3 4 5 1: UAS focuses only on the target's location 3: UAS verifies target composition, disposition, and intentions 5: UAS proactively adjust reconnaissance plan in anticipation of future actions	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the UAS provide updates on target behavior (changes in size, composition, disposition, activities, and movement)? 1 2 3 4 5 1: UAS does not provide any updates after target acquisition 3: UAS provides target updates when prompted by teammate or supported ground unit 5: UAS provides updates on target as they occur or that have a potential impact on the mission	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does UAS continuously update the Flight on changes to the common operational picture (COP)? 1 2 3 4 5 1: UAS does not communicate changes to COP 3: UAS communicates changes in target but not friendly troops or effects of terrain or obstacles 5: UAS continuously updates COP for changes in target, friendly troops, terrain, and obstacles	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐

Does UAS update the ground unit they are supporting with information that could influence operations? 1 — UAS does not communicate changes 2 3 — UAS communicates changes in target but not friendly troops or effects of terrain or obstacles 4 5 — UAS continuously updates for changes in target, friendly troops, terrain, and obstacles	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree
Does the UAS provide early warning and threat detection to supported unit? 1 — UAS provides delayed, inaccurate, or misdirected information 2 3 — UAS reports accurate and timely information about threat activity 4 5 — UAS reports accurate and timely information on threat, and METT-TC	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree
Does the UAS prioritize engagement of targets? 1 — UAS does not discuss priority of targets 2 3 — UAS discusses priority but decision does not reflect Cdr's intent or threat(s) 4 5 — UAS prioritizes engagement based on Cdr's Intent and threat(s)	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree
Does the UAS update the engagement priority as it changes? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree	This measure is **observable** during MUM-T training. ☐ Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree

Does the UAS follow the correct procedures and format for a Target Handover to wingman (Voice/Laser)? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
If applicable, which required elements were missed? ☐ Laser spot/code ☐ Target Description ☐ Target Location ☐ Attack Instructions ☐ Additional information as required	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Did the UAS correctly identify the target(s) during target handover? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the flight recommend lethal and nonlethal COAs to Ground Commander? 1 2 3 4 5 1: Flight provides no COA to Ground Commander 3: Flight provides an (one) acceptable COA 5: Flight provides multiple, prioritized COAs based on aerial perspective	This measure is **relevant** for MUM-T training. Strongly Disagree Disagree Agree Strongly Agree	This measure is **observable** during MUM-T training. Strongly Disagree Disagree Agree Strongly Agree
Does the flight select the appropriate weapon for desired effect on target? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree Disagree Agree Strongly Agree	This measure is **observable** during MUM-T training. Strongly Disagree Disagree Agree Strongly Agree

Does the flight conduct a Collateral Damage Assessment (CDA)? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree / Disagree / Agree / Strongly Agree	This measure is **observable** during MUM-T training. Strongly Disagree / Disagree / Agree / Strongly Agree
Does the flight confirm hostile intent prior to applying lethal force? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ / Disagree ☐ / Agree ☐ / Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ / Disagree ☐ / Agree ☐ / Strongly Agree ☐
If no, why not? Flight never discusses hostile intent ☐ Flight assumes hostile intent; relies on other reports ☐ Flight determines possible hostile intent; talks themselves into it	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ / Disagree ☐ / Agree ☐ / Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ / Disagree ☐ / Agree ☐ / Strongly Agree ☐
Does the UAS verify location of friendly forces near the target to prevent fratricide? 1 — 2 — 3 — 4 — 5 1: UAS does not verify friendly locations in relation to the target 3: UAS obtains location for friendly forces in the area and has visual of friendly locations 5: UAS obtains locations, has visual, and confirms outside safety fan/surface danger zone	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ / Disagree ☐ / Agree ☐ / Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ / Disagree ☐ / Agree ☐ / Strongly Agree ☐
Does the UAS deconflict the airspace in preparation for missile launch? 1 — 2 — 3 — 4 — 5 1: UAS does not use information sources to anticipate events; does not make radio calls 3: UAS uses information sources to anticipate events but does not push information to rest of team 5: UAS uses information sources to anticipate events and proactively pushes information to rest of team	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ / Disagree ☐ / Agree ☐ / Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ / Disagree ☐ / Agree ☐ / Strongly Agree ☐

Does the flight choose and brief the appropriate engagement scheme of maneuver? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
If applicable, which required elements were missed? ☐ Techniques. ☐ Patterns ☐ Munitions ☐ Range	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the UAS coordinate with the manned aircraft to ensure launcher designation angles (max offset), safety fan, laser code and laser-on time requirements can be met? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
If applicable, which required elements were missed? LDA ☐ Safety Fan ☐ Laser Code ☐ Laser-on Time	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the UAS follow the correct procedures and format for a remote Hellfire launch (Voice/Digital/Laser)? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Was the UAS proactive in executing call for indirect fire? 1 — 2 — 3 — 4 — 5 1: UAS not ready for indirect fire call 3: UAS prepared to make indirect fire call 5: UAS foresaw need for indirect fires; prepared to make indirect fires call	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐

Does the UAS use the proper format for indirect fire missions? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
If applicable, which required elements were missed? ☐ Observer Identification ☐ Warning Order ☐ Target Location ☐ Target Description ☐ Method of Engagement ☐ Method of Fire and Control	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the UAS acknowledge receipt, or any changes, from Fire Direct Center (FDC)? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Does the UAS relay direction of target in degrees and range from the firing unit's position to other aircraft in order to clear the airspace? ☐ Yes ☐ No	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐
Is the UAS proactive in airspace deconfliction throughout execution of indirect fires? 1 — 2 — 3 — 4 — 5 1: UAS does not have SA of battlespace during the call for indirect fire 3: UAS de-conflicted battlespace after making the call for indirect fires 5: UAS knows location of all assets and deconflicts airspace prior to call for indirect fire	This measure is **relevant** for MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐	This measure is **observable** during MUM-T training. Strongly Disagree ☐ Disagree ☐ Agree ☐ Strongly Agree ☐

APPENDIX D

Relevancy Ratings for Each Manned-Unmanned Teaming Performance Measure.

Performance Measure	Rating				X^2
	Strongly Disagree	Disagree	Agree	Strongly Agree	
Flight incorporates ISR plan	0	0	12	7	21.63
Flight actively searches for target	0	0	9	10	19.11
UAS uses appropriate sensors	0	0	7	12	21.63
UAS shares sensor feed with flight	0	0	12	7	21.63
UAS shares sensor feed with tactical operations center	0	0	11	8	19.95
UAS shares sensor feed with ground unit	0	1	15	3	30.47
UAS recognizes threats during mission	0	0	12	7	21.63
UAS announces target acquisition	0	1	9	9	15.32
UAS sends SPOT report	0	0	11	8	19.95
UAS maintains positive identification of target.	0	0	10	9	19.11
UAS reports when contact lost	0	0	11	8	19.95
Flight coordinates duties after target acquired	0	0	10	9	19.11
UAS provides continuous reconnaissance	0	0	8	11	19.95
UAS updates target behavior	0	0	9	10	19.11
UAS updates flight on changes in Common operating picture	0	1	11	7	17.00
UAS provides updates to ground unit	0	2	11	6	14.89
UAS provides early warnings, overwatch	0	0	12	7	21.63
UAS prioritizes engagement of targets	4	7	7	1	5.21*

Performance Measure	Rating				X^2
	Strongly Disagree	Disagree	Agree	Strongly Agree	
UAS updates engagement priority	4	7	7	1	5.21*
UAS follows correct procedures target handover	0	0	12	7	21.63
UAS correctly identifies targets during handover	0	0	11	7	17.00
Flight recommends course of action to ground commander	1	2	9	5	9.12
Flight selects appropriate weapon	0	3	10	5	11.78
Flight conducts collateral damage assessment	0	1	7	11	17.00
Flight confirms hostile intent before lethal force	0	1	8	10	15.74
UAS verifies location of friendly forces	0	0	6	13	24.16
UAS deconflicts airspace before missile launch	1	6	8	4	5.63*
Flight uses appropriate engagement scheme	0	2	11	6	14.89
UAS coordinates manned aircraft laser code	0	0	13	6	24.16
UAS uses correct procedure for remote Hellfire launch	0	0	8	10	18.44
UAS proactive in executing call for indirect fire	1	2	15	1	29.63
UAS uses proper format for indirect fire mission	1	0	14	3	27.78
UAS acknowledges receipt from fire direct center	2	3	9	4	6.44*
UAS relays target direction & range to other aircraft	2	3	8	3	5.50*
UAS proactive in airspace deconfliction	2	3	9	4	6.44*
UAS conducts standardized Battle Damage Assessment	0	1	9	9	15.32

Note: Asterisks (*) indicate non-statistically-significant chi-square

APPENDIX E

Observability Ratings for Each Manned-Unmanned Teaming Performance Measure.

Performance Measure	Rating: Strongly Disagree	Disagree	Agree	Strongly Agree	X^2
Flight incorporates ISR plan	2	0	10	7	13.21
Flight actively searches for target	0	1	9	8	14.44
UAS uses appropriate sensors	0	0	10	8	48.44
UAS shares sensor feed with flight	0	1	9	8	14.44
UAS shares sensor feed with tactical operations center	1	1	9	8	11.95
UAS shares sensor feed with ground unit	0	5	11	1	17.59
UAS recognizes threats during mission	0	0	11	7	19.78
UAS announces target acquisition	0	1	9	7	13.82
UAS sends SPOT report	0	0	11	6	19.94
UAS maintains positive identification of target.	0	0	10	8	18.44
UAS reports when contact lost	0	0	9	9	15.32
Flight coordinates duties after target acquired	0	0	10	8	18.44
UAS provides continuous reconnaissance	0	0	9	9	15.32
UAS updates target behavior	0	0	10	8	18.44
UAS updates flight on changes in common operating picture	0	1	10	7	15.33
UAS provides updates to ground unit	0	2	10	6	13.11
UAS provides early warnings, overwatch	0	0	13	4	26.53
UAS prioritizes engagement of targets	3	7	6	2	3.78*

Performance Measure	Rating				X^2
	Strongly Disagree	Disagree	Agree	Strongly Agree	
UAS updates engagement priority	2	9	5	2	7.33*
UAS follows correct procedures target handover	0	0	12	6	22.00
UAS correctly identifies targets during handover	0	0	11	7	19.78
Flight recommends course of action to ground commander	1	4	5	6	3.50*
Flight selects appropriate weapon	0	3	8	5	8.50
Flight conducts collateral damage assessment	0	1	8	8	13.35
Flight confirms hostile intent before lethal force	0	1	7	9	13.82
UAS verifies location of friendly forces	0	0	7	10	18.06
UAS deconflicts airspace before missile launch	1	6	5	4	3.50*
Flight uses appropriate engagement scheme	0	2	11	4	16.18
UAS coordinates manned aircraft laser code	0	0	12	6	22.00
UAS uses correct procedure for remote Hellfire launch	0	0	8	10	18.44
UAS proactive in executing call for indirect fire	1	2	11	2	16.50
UAS uses proper format for indirect fire mission	1	1	9	4	11.40
UAS acknowledges receipt from fire direct center	3	4	4	4	<1.00*
UAS relays target direction & range to other aircraft	3	4	4	3	<1.00*
UAS proactive in airspace deconfliction	2	3	6	4	2.33*
UAS conducts standardized Battle Damage Assessment	0	0	8	9	17.12

Note: Asterisks (*) indicate non-statistically-significant chi-square