

ARI Research Note 2019-01

**Training and Evaluation Outlines (T&EO): Usage and
Scoring Method Preference for Task Steps and Sub-
steps**

Hayley S. Foo
U.S. Army Research Institute



Fort Hood Research Unit
Brian T. Crabb, Chief

June 2019

**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**

**A Directorate of the Department of the Army
Deputy Chief of Staff, G1**

Authorized and approved for distribution:

**MICHELLE L. ZBYLUT, Ph.D.
Director**

Technical Review by

Dr. Jeffrey E. Fite, U.S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Research Note has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: DAPE-ARI-ZXM, 6000 6th Street, Bldg 464/Mail Stop 5610, Fort Belvoir, VA 22060-5610

FINAL DISPOSITION: This document may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this note 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-yy)		2. REPORT TYPE Final		3. DATES COVERED (from. . . to) January 2015 – January 2016	
4. TITLE AND SUBTITLE Training and Evaluation Outlines (T&EO): Usage and Scoring Method Preference for Task Steps and Sub-steps				5a. CONTRACT OR GRANT NUMBER	
				5b. PROGRAM ELEMENT NUMBER 662785	
6. AUTHOR(S) Hayley S. Foo (U.S. Army Research Institute)				5c. PROJECT NUMBER A790	
				5d. TASK NUMBER 215	
				5e. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences Fort Hood Research Unit Bldg 36000, Darnall Loop Fort Hood, TX 76544-0993				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street, Bldg 1464/Mail Stop 5610 Fort Belvoir, VA 22060-5610				10. MONITOR ACRONYM ARI	
				11. MONITOR REPORT NUMBER	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES Subject Matter POC: Hayley S. Foo					
14. ABSTRACT (<i>Maximum 200 words</i>): A Training and Evaluation Outline (T&EO) describes the task, conditions, and standards for training and is the Army standard for training and evaluating individual and collective tasks. There is conflicting evidence on the extent to which T&EOs are used. This effort examined the frequency of T&EO use according to commission type (officers and noncommissioned officers), training environment (Combat Training Center versus Home Station), and Career Management Fields (CMF). Performance on T&EO task steps and performance measures (sub-steps) are evaluated using GO/NOGO, a method that may not be sufficiently discriminative. In order to determine the acceptability of using a numeric scale, preferences for GO/NOGO versus a numeric scale were examined. The results indicate that frequency of T&EO use was bimodal: falling in either < 40% or ≥ 60%. There was no strong preference for using GO/NOGO or a numeric scale for the task steps but a strong preference for using GO/NOGO for the sub-steps. The findings were the same regardless of commission type, training environment, and CMF.					
15. SUBJECT TERMS T&EO GO/NOGO Numeric Performance Measures					
SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT Unlimited	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON Brian T. Crabb 254-288-3833
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified			

ARI Research Note 2019-01

Training and Evaluation Outlines (T&EO): Usage and Scoring Method Preference for Task Steps and Sub-steps

Hayley S. Foo
U.S. Army Research Institute

Fort Hood Research Unit
Brian T. Crabb, Chief

U.S. Army Research Institute
6000 6th Street, Bldg 1464
Fort Belvoir, VA 22060-5610

June 2019

Army Project Number
20262785A790

Personnel Performance
and Training Technology

Approved for public release; distribution is unlimited.

TRAINING AND EVALUATION OUTLINES (T&EO): USAGE AND SCORING METHOD PREFERENCE FOR TASK STEPS AND SUB-STEPS

CONTENTS

	Page
Method	2
Participants and Procedure.....	2
Data Analyses	2
Results.....	2
Demographics	2
T&EO Usage: Overall Frequency.....	3
T&EO Usage: Frequency by Commission Type, Training Environment, and CMF	3
T&EO Task Steps: Overall Scoring Method Preference	4
T&EO Task Steps: Scoring Method Preference by Commission Type, Training Environment, and CMF	5
Performance Measures Task Sub-steps: Overall Scoring Method Preferences.....	6
Task Sub-steps: Scoring Method Preference by Commission Type, Training Environment, and CMF	7
Discussion.....	8
References.....	10
Appendix A. Interview Questions.....	A-1

LIST OF FIGURES

Figure 1. T&EO Usage	3
Figure 2. T&EO Usage by Commission Type, Training Environment, and Other CMFs	4
Figure 3. Scoring Method Preference for T&EO Task Steps	5
Figure 4. Scoring Method Preference for T&EO Task Steps by Commission Type, Training Environment, and Other CMFs.....	6
Figure 5. Scoring Method Preference for T&EO Task Sub-steps	7
Figure 6. Scoring Method Preference for T&EO Task Sub-steps by Commission Type, Training Environment, and Other CMFs	8

TRAINING AND EVALUATION OUTLINES (T&EO): USAGE AND SCORING METHOD PREFERENCE FOR TASK STEPS AND SUB-STEPS

Mission success requires a trained and ready force (Department of the Army, 2016). Thus, reliable and valid measures of unit training readiness are essential. U.S. Army Training and Doctrine Command (TRADOC) publications such as Training Circulars (TC) and Training & Evaluation Outlines (T&EO) provide the Army standards for training and evaluating individual and collective tasks. A T&EO describes the task, conditions, and standards for training, and four sections – Plan, Prepare, Execute, and Assess – that are evaluated. Each section contains a list of task steps and performance measures (sub-steps), with each step and measure rated as GO/NOGO and the overall task assessed as T (Fully Trained), P (Practiced), or U (Untrained). The GO/NOGO checklist and TPU rating of training readiness is used in both individual and collective tasks. Recently, the scale for the task proficiency standards has been expanded to T, T- (Trained), P, P- (Marginally Practiced), or U (Untrained). (Department of the Army, 2016).

Operational units are expected to use Army standards (i.e., T&EO) to evaluate training at home station. However, very little published data are available on the extent to which T&EOs are used. It has been claimed, without supporting evidence, that T&EOs were used extensively for home station training (Fober, 1997). The aim of the present study is to examine how often T&EOs are used, by determining the frequency of use according to commission type (officers versus noncommissioned officers [NCO]), training environment (home station (HS) versus Combat Training Center [CTC]), and different Career Management Fields (CMF).

In the past, a T&EO database was kept at the Joint Readiness Training Center (JRTC), one of the Army's three CTCs, with the goal of enabling trainers and researchers to study progress and trends in unit performance (Fober, 1993, 1997; Fober, Dyer, & Salter, 1994; Nichols, 1991). However, comparisons of TPU ratings on company tasks from five rotational units at JRTC showed that the scale did not discriminate performance among units. For example, performances on most (> 70%) company tasks were rated as U (Fober, 1997), suggesting a need for more sensitive and/or accurate measures of performance. After several years, the database was discontinued for this and other reasons, including i) Observer/Controllers (OC) finding the T&EO format too detailed and cumbersome, ii) OCs finding the tasks of observing and rating too onerous, leading them to often fill in the checklists long after the training session, iii) OCs finding the T&EO evaluations unhelpful for After-Action Reviews (AAR), and iv) the task being evaluated did not always follow the step sequence outlined in the T&EO (Fober, 1997). These findings underscore the fact that OCs at JRTC, who are considered subject matter experts (SME) in unit training and evaluation, found the assessment tool problematic.

One way of improving the T&EO GO/NOGO checklist may be to incorporate a numerical scale. The advantages of having quantitative measures are that they provide enhanced metrics to gauge learning, rank performance, and benchmark different levels of performance. Nevertheless, it must be noted that the standard proficiency evaluation rubric used in the U.S. Army is GO/NOGO. In some cases, both numeric scores and GO/NOGO are used to evaluate

performance, as in the Army Physical Fitness Test (DA Form 705, May 2010). In order to determine the acceptability of using a numeric system over GO/NOGO, the second aim of this effort is to assess the scoring method preference of GO/NOGO and a numeric scale for evaluating T&EO task steps and performance measures (sub-steps).

Method

Participants and Procedure

Participants were recruited from a CTC and a HS via a research support request for platoon leaders, company commanders, platoon sergeants, first sergeants, Abrams master gunners, and Bradley master gunners. There were 106 participants who were given a questionnaire to complete. The questions relevant to this research note were on rank, military occupational specialty (MOS), time in service, T&EO usage, and the relative preferences for GO/NOGO and numerical scoring as the method for evaluating T&EO task steps and sub-steps. The questions asked are in Appendix A. Participants were also asked to identify two collective tasks and the performance indicators that they look for in assessing those collective tasks. The results from these discussions were used to identify two collective tasks for a follow-on project.

Data analyses

Military Occupation Specialties within a CMF were combined for data analyses when appropriate for the category. The average years in service was calculated separately for officers and NCOs. In order to increase statistical power, responses for T&EO usage were consolidated into three categories: ‘≤ Seldom’ (< 40%), ‘Sometimes’ (40-59%), and ‘≥ Often’ (≥ 60%). Likewise, the responses on preference for the two types of scoring methods were combined into three categories: ‘GO/NOGO’ (= Strongly prefer + prefer GO/NOGO over numerical scores), ‘Neither’ (No preference for either method), and ‘Numeric’ (‘Strongly prefer + prefer numerical scores over GO/NOGO). For each of the T&EO questions, an equiprobability model was used to test whether overall responses were equally distributed across the three categories. It must be noted that consolidation of the data into three categories meant that potential differences within the consolidated categories (e.g., ‘Very Seldom’ versus ‘Seldom’, ‘Often’ versus ‘Very Often’) were not tested. This consolidation did not impact the chi-square tests of independence performed to examine differences between groups in usage and preferences. The chi-square test of independence tested for significant differences in pattern of responses across categories according to i) commission type (officers versus NCOs), ii) training environment (CTC versus HS), iii) CMFs (11 (Infantry), 13 (Field Artillery), 19 (Armor) versus other CMFs, and iv) CMFs 11, 13, and 19 at different training environments (CTC versus HS). For all analysis, the alpha level was set at 0.05.

Results

Demographics

Participants were 40 officers (O2-O3) and 66 NCOs (E6-E8). The average years of service was 8.3 for officers and 16.6 for NCOs. There were 69 participants from the CTC and 37

from the HS. Participants from the HS were all in CMFs 11 (Infantry), 13 (Field Artillery), and 19 (Armor) except for one participant whereas those from the CTC were from 12 different CMFs.

T&EO Usage: Overall Frequency

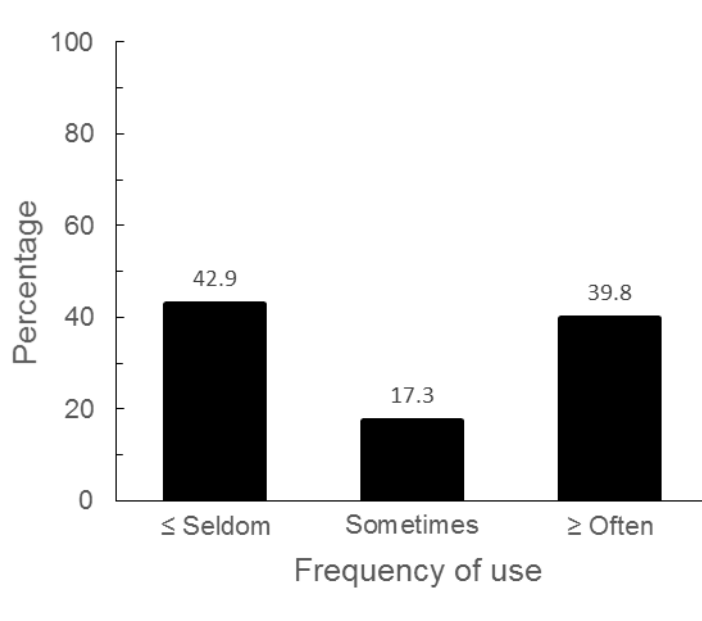


Figure 1. T&EO usage.

The average percentages of T&EO usage in the categories of ‘≤ Seldom’ (< 40%), ‘Sometimes’ (40-59%), and ‘≥ Often’ (≥ 60%) are shown in Figure 1. A chi-square test of equiprobability was significant, $\chi^2 (2, N = 98) = 11.72, p < 0.05$, confirming non-equal distribution across the three categories. As Figure 1 shows, reported usage was bimodal, with the majority (>80%) in either the ‘≤ Seldom’ (< 40%) or ‘≥ Often’ categories.

T&EO Usage: Frequency by Commission type, Training environment, and CMF

Figure 2 shows the average percentages of T&EO usage in the categories of ‘≤ Seldom’, ‘Sometimes’, and ‘≥ Often’ among officers and NCOs (upper left), at CTC and HS (upper right), between CMFs 11, 13, 19 and other CMFs (lower left), and between CMFs 11, 13, 19 at CTC and HS (lower right). A chi-square test of independence between officers and NCOs across the three categories was not significant, $\chi^2 (4, N = 98) = 2.50, p = 0.29$, confirming that both groups showed the same pattern of use. The T&EO usages were bimodal at both CTC and HS, falling mostly in the two opposing categories of ‘≤ Seldom’ and ‘≥ Often.’ A chi-square test of independence was not significant, $\chi^2 (4, N = 99) = 1.06, p = 0.59$, confirming that the frequency of use did not depend on type of training environment (CTC versus HS). There were also no significant differences in pattern of T&EO use between Soldiers in CMFs 11, 13, 19 compared to those in other CMFs, $\chi^2 (4, N = 78) = 1.09, p = 0.54$. For Soldiers in CMFs 11, 13, 19, pattern of

usage was the same at different training environments (CTC and HS), $\chi^2(4, N = 70) = 0.74, p = 0.69$.

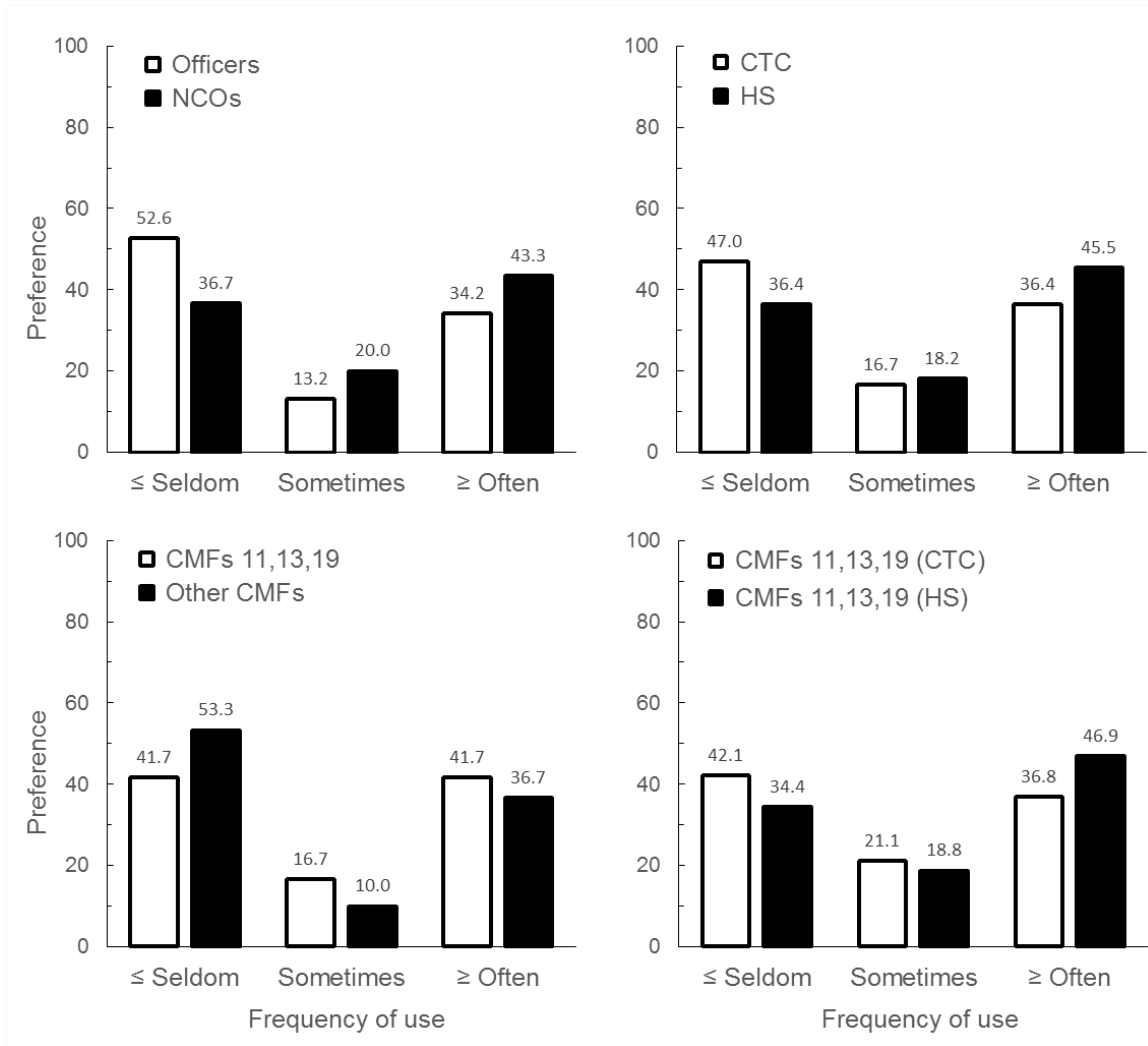


Figure 2. T&EO usage by commission type (upper left: officers and NCOs), training environment (upper right: CTC and HS) and CMF (lower left: CMFs 11, 13, 19 and other CMFs; lower right: CMFs 11, 13, 19 at CTC and HS).

T&EO Task Steps: Overall Scoring Method Preference

Figure 3 shows the average percentages who favored GO/NOGO over Numeric (GO/NOGO), Numeric over GO/NOGO (Numeric), or neither method (No Preference) as the scoring method for evaluating T&EO task steps. A chi-square test of equiprobability was not-significant, $\chi^2(2, N = 98) = 4.03, p = 0.13$, confirming no overarching preference for using either of the two scoring methods for rating performance on T&EO task steps.

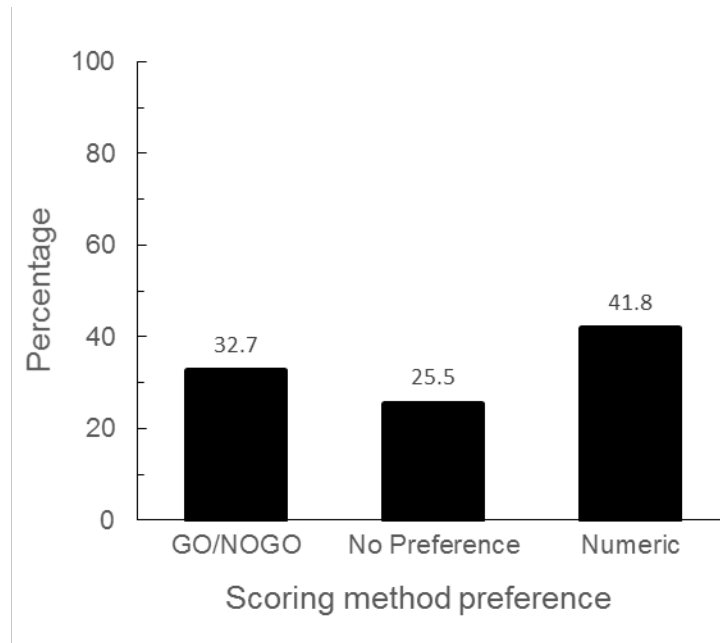


Figure 3. Scoring Method Preference for T&EO task steps.

T&EO Task Steps: Scoring Method Preference by Commission Type, Training Environment, and CMF

Figure 4 shows the scoring method preference by commission type, training environment, and CMF. Officers and NCOs appear to display different patterns of scoring method preferences (GO/NOGO, No preference, Numeric) for evaluating T&EO task steps (Figure 4, upper left). This observation was not confirmed as the chi-square test of independence was not significant, $\chi^2(4, N = 98) = 5.31, p = 0.07$. However, the data suggest that officers are likely to have a stronger preference for numerical scoring whereas NCOs prefer GO/NOGO. At both CTC and HS, the percentages of Soldiers indicating preferences for GO/NOGO, No preference, and Numeric for assessing T&EO task steps did not differ significantly, $\chi^2(4, N = 100) = 2.63, p = 0.27$. Similarly, Soldiers in CMFs 11, 13, 19 and those in other CMFs showed the same pattern of scoring method preference, $\chi^2(4, N = 79) = 2.08, p = 0.35$. For Soldiers in CMFs 11, 13, and 19, the same scoring method preferences were displayed irrespective of training environment (CTC versus HS), $\chi^2(4, N = 70) = 2.97, p = 0.23$.

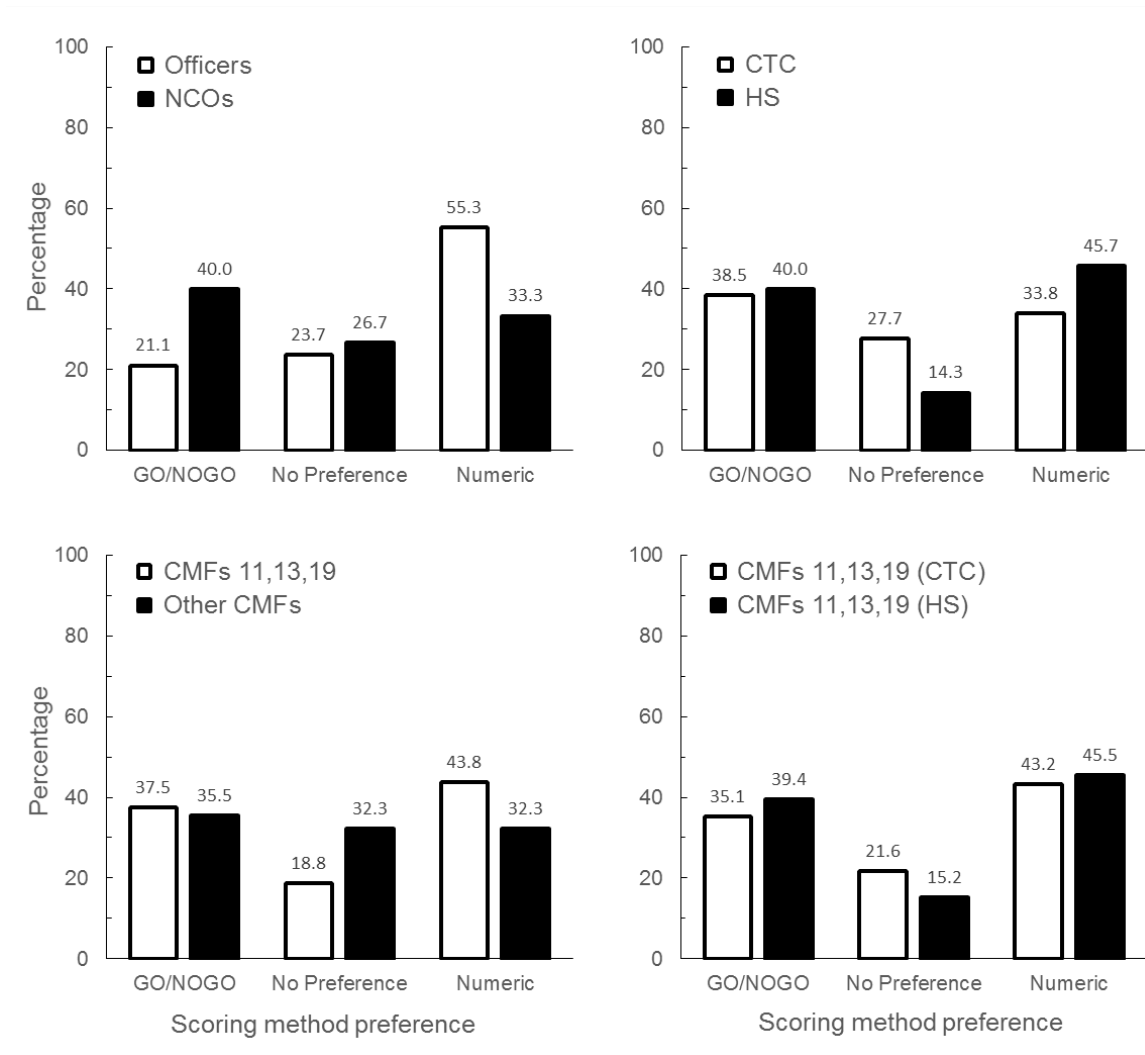


Figure 4. Scoring Method Preference for T&EO task steps by commission type (upper left: officer and NCOs), training environment (upper right: CTC and HS), and CMF (lower left: CMFs 11, 13, 19 and other CMFs; lower right: CMFs 11, 13, 19 at CTC and HS).

Performance Measures (Task Sub-steps): Overall Scale Preferences

The average percentages for scoring method preference for GO/NOGO over Numeric (GO/NOGO), Numeric over GO/NOGO (Numeric), or neither method (No preference) to evaluate T&EO task sub-steps are shown in Figure 5. A chi-square test of equiprobability was statistically significant, $\chi^2(2, N = 96) = 20.34, p < 0.05$, confirming higher percentages in one or more of the categories compared to the others. As can be seen from Figure 5, more respondents favored GO/NOGO than numeric or no preference.

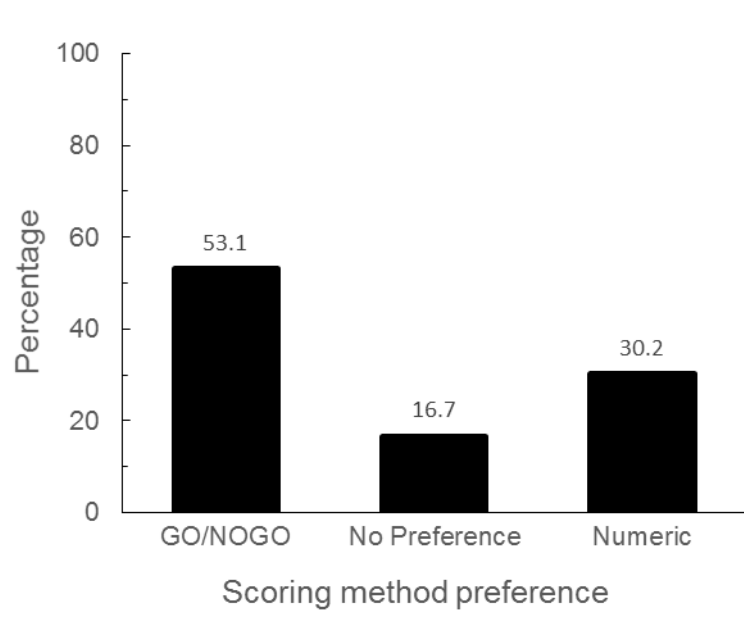


Figure 5. Scoring method preference for T&EO task sub-steps.

Task Sub-steps: Scoring Method Preference by Commission Type, Training Environment, and CMF

Figure 6 shows that percentage of responses in GO/NOGO, No preference, and Numeric categories on T&EO performance measures (sub-steps) according to commission type (officers versus NCOs, upper left), training environment (CTC versus HS, upper right), and different CMF (lower left and lower right). Officers and NCOs showed similar preferences which was confirmed by a lack of significance in the chi-square test of independence, $\chi^2(4, N = 96) = 1.48, p = 0.48$. There was also no significant differences in the way Soldiers at CTC and HS rated their preferences, $\chi^2(4, N = 98) = 2.00, p = 0.37$. Similarly, Soldiers in CMFs 11,13, 19 and those in other CMFs showed the same preferences for the scoring methods because a chi-square test of independence was not significant, $\chi^2(4, N = 76) = 3.24, p = 0.20$. For Soldiers in CMFs 11, 13, 19, preferences for the different scoring methods were the same irrespective of training environment (CTC versus HS), $\chi^2(4, N = 71) = 1.23, p = 0.54$.

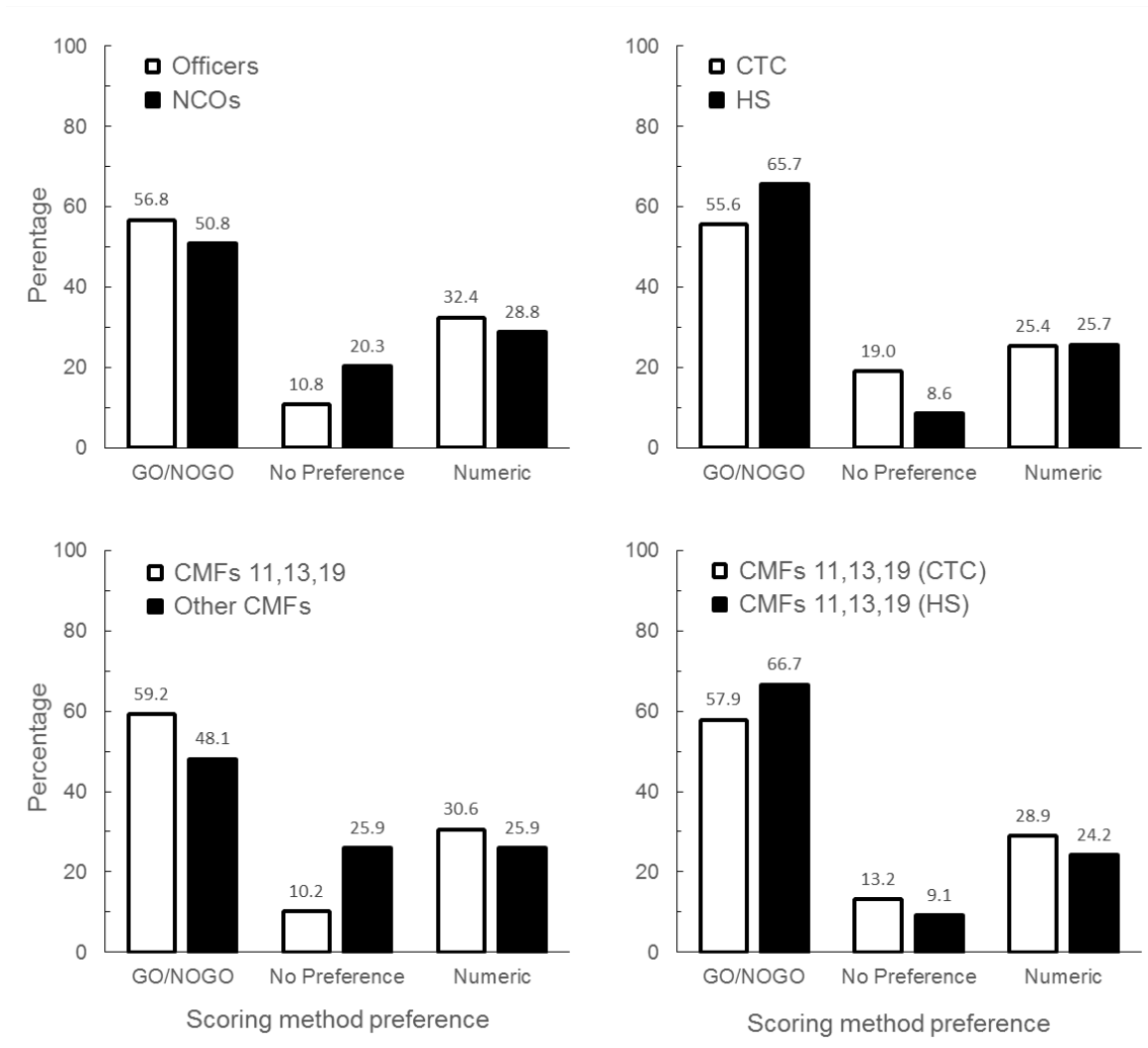


Figure 6. Scoring Method Preference for T&EO task sub-steps by commission type (upper left: officer and NCOs), training environment (upper right: CTC and HS), and CMF (lower left: CMFs 11, 13, 19 and other CMFs; lower right: CMFs 11, 13, 19 at CTC and HS).

Discussion

The present results show that self-reported T&EO use fell mainly in the categories of <40% (Seldom + Very seldom) or $\geq 60\%$ (Often + Very often). This pattern of use was evident among officers and NCOs, at both CTC and HS, across Soldiers in diverse CMFs, and among Soldiers in Infantry, Field Artillery, and Armor. Specifically, only 39.8% of respondents reported using T&EOs $\geq 60\%$ of the time. This finding is in contrast to a previous report of extensive T&EO use for HS training (Foher, 1997). These differences suggest a substantial decline in T&EO usage for evaluating HS training in the last 20 years, although it is unknown when and why the shift occurred. Comments from participants indicated that potential reasons were that some did not know about T&EOs and what T, P, and U meant and others who knew about T&EOs found them unhelpful. Additionally, while use of T&EOs is expected for HS training, the practice is not enforced. This situation will change with the recent Army wide

implementation of Objective-T that requires unit to use T&EOs to evaluate and report training proficiency of mission essential tasks (MET).

The present research found no overall preference for using GO/NOGO over a numeric scoring for evaluating T&EO task steps. This lack of preference was evident regardless of training environment (CTC versus HS) and CMF (CMFs 11, 13, 19 versus other CMFs; Soldiers in CMFs 11, 13, 19 based at CTC or HS). However, officers may favor using a numeric scale to GO/NOGO whereas NCOs appear to prefer GO/NOGO over numeric scoring (Figure 4, upper left). In contrast to the lack of preferences for a particular scoring method for evaluating T&EO task steps, there was a robust overall preference for using GO/NOGO over numeric scores for evaluating the task sub-steps. This preference was shown among officers and NCOs, whether Soldiers were based at a CTC or a HS, and across all CMF surveyed. Taken together, these findings suggest that if a numerical scoring method was introduced into T&EOs, it may be taken more favorably by officers than NCOs and would be better incorporated for the task steps than the task sub-steps. These findings has informed the follow-on research and current effort to develop 5-point Likert based Behaviorally Anchored Rating Scales (BARS) for T&EO task steps of company/troop level collective METs. Incorporation of behaviorally observable and verifiable measures into T&EOs would provide better ways to assess, track, and guide HS training of collective METs by increasing the objectivity, specificity and discriminability of the performance.

The T&EO task steps and sub-steps were constructed using a task-oriented approach. This approach works extremely well for assessing individual tactical skills, such as assembly of a weapon. Use of a highly structured step-by-step task analysis for assessing individual (leader) and collective (unit) performances on collective tasks is not optimal because socio-cognitive factors are critically important for successful teamwork. There is already substantial evidence that leadership experience and styles likely impact unit performance. For example, units with more experienced leaders, and leaders with a transformative leadership style, showed better field performances at JRTC (Bass & Avolio, 2000; Dyer et al., 1992). Other human factors considered critical for military team performance include morale, motivation, communication, shared mental models, and unit cohesion (Bass & Avolio, 2000; Holz, Hiller, & McFann, 1994; Smith, Borgvall, & Lif, 2007). The U.S. Army's view of leadership encompasses the attributes of character, presence, and intellect and the competencies of leading, developing, and achieving (Department of the Army, 2012). Yet, most of these attributes and competencies, and other "soft" skills vital for successful and efficient teamwork, are not formally trained or assessed during field training exercises. In the follow-on research, socio-cognitive factors such as the quantity and quality of communication and shared understanding among team members will be incorporated into a T&EO quick reference guide as key elements of performance.

References

- Bass, B. M., & Avolio, B. J. (2000). Platoon readiness as a function of leadership, platoon, and company cultures. (ARI Technical Report 1104). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Dyer, J., Fober, G. W., Pleban, R. J., Salter, M. S., Valentine, P. J., & Thompson, T. J. (1992). Light infantry performance at the Combat Training Centers: Home station determinants. (ARI Research Note 92-33). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Fober, G. W. (1993). The Joint Readiness Training Center's training and evaluation outline data base. (ARI Research Note 93-11). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Fober, G. W. (1997). The utility of the training and evaluation outline data base as a performance measurement system at the Joint Readiness Training Center. (ARI Research Note 97-05). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Fober, G. W., Dyer, J. L., & Salter, M. S. (1994). Measurement of performance at the Joint Readiness Training Center: Tools of assessment. In R. F. Holz, J. H. Hiller, & H. H. McFann, (Eds.) *Determinants of effective unit performance* Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Holz, R. F., Hiller, J. H., & McFann, H. (1994). Editors in *Determinants of Effective Unit Performance*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Nichols, J. J. (1991). Analysis of the scope, functionality, and usability of the Joint Readiness Training Center (JRTC) data base archives. (ARI Research Note 91-24). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Smith, E., Borgvall, J., & Lif, P. (2007). *Team and collective performance measurement*. Bedfordshire, UK: DSTL, Policy and Capability Studies Bedford Technology Park Thurleigh.
- U.S. Department of the Army. (2012, August). ADRP 6-22. *Army Leadership*. Washington, DC: Author.
- U.S. Department of the Army. (2016, October). FM 7-0. *Train to Win in a Complex World*. Washington, DC: Author.

Appendix A

1. What is your rank, MOS, Branch, and position?
2. How long have you served in the Army?
3. How often do you use Training & Evaluation Outlines (T&EO) to assess collective tasks?

Very seldom (< 20%)

Seldom (20 – 39%)

Sometimes (40 – 59%)

Often (60 – 79%)

Very Often (80 – 100%)

4. Please select your preference for a method to assess collective task steps.

Strongly prefer GO/NO-GO over numerical scores

Prefer GO/NO-GO over numerical scores

No preference for either method

Prefer numerical scores over GO/NO-GO

Strongly prefer numerical scores over GO/NO-GO

5. Please select your preference for a method to assess performance measures within a task step.

Strongly prefer GO/NO-GO over numerical scores

Prefer GO/NO-GO over numerical scores

No preference for either method

Prefer numerical scores over GO/NO-GO

Strongly prefer numerical scores over GO/NO-GO