

AB-47C - Symposium

Assessing Individual Productivity in the Workplace: An Improved Criterion for Training Evaluation

**Brice Stone
Larry Looper
Bronwyn Salathiel**

19990423 047

The objective of this research effort was the analysis of the relationships among task performance time, aptitude, task training time, and proficiency. The research focused on aircraft maintenance personnel and training issues. The Air Force was interested in a complex network of functional relationships or tradeoffs, involving manpower, personnel, and training issues. For example, can the Air Force obtain the same level of performance with fewer people if it selects those with higher aptitude levels and provides more efficient training. The functional relationships of concern involve training time estimates and task performance time estimates as they are effected by aptitude and experience (Lance, Hedge, & Alley, 1987).

In past research, Metrica Inc. developed four equations, one each for the Mechanical, Administrative, General, and Electronics (MAGE) categories of the ASVAB, which were generalizable across all jobs in the MAGE category. The goal was to relate maintenance task performance time to ASVAB Indicator (MAGE) Scores (aptitude), Total Active Federal Military Service (experience), and OMS difficulty measures (average task difficulty). The equations were based on data from the Productive Capacity project (Stone, Turner, Wiggins & Looper, 1996), which involved two maintenance AFSs: Aerospace Ground Equipment and Avionics Communication and Navigation Systems. In contrast, this research effort focused on twenty F-15 and F-16 aircraft maintenance AFSs, from the Mechanical and Electronics ASVAB categories. Two additional career fields were included to ensure that the methodology and analysis were generalizable to all four MAGE category career fields.

The two relationships to be captured in this effort include the relationship between 1) a person's time required to attain full proficiency on a task (training time) and aptitude, on-the-job (OJT) training time, experience, and other factors; and 2) a person's time to perform a task (performance time), and aptitude, formal training time, experience, and other factors. The goal is to produce a model to predict training and performance time. This model will use as predictors a number of variables, one of which will be the person's aptitude. The model is said to be defined when all of the predictors have been identified, their method of measurement defined, and their coefficients in a general linear regression model determined.

Data Collection

The data collection plan for this research effort was based on as many as 200 subjects in each of 19 AFSs associated with F-15 or F-16 weapons system maintenance (see Table 1) and included between 30 and 60 tasks for each AFS. A "percent of proficiency" was defined as a continuous variable (percent of proficiency) ranging from 0 to 100 percent, where 100 percent proficiency was defined as "task performance with a minimum amount of assistance and supervision." This was the definition used by Perrin et al. (1988) which was shown to yield time estimates with moderate reliability. This recommended "percent of proficiency" was expected to be more easily related to measures of task training time and task performance time.

Task-level measures of performance time, training time, proficiency, and experience were collected. At technical training centers, task level data was collected from course instructors/supervisors (as many as were available) with respect to: a) the number of hours required to train the average student to the specified Specialty Training Standard (STS) performance level, b) the number of hours to train to the next lower STS performance level, and c) the "percent of proficiency" corresponding to these hour estimates (the "lower" estimate provides the "second point" needed to specify relationships). The collected data was used to specify the relationships among task formal training time length, aptitude, and proficiency.

At ACC bases, data was collected from both incumbents (5-skill levels) and the incumbent's immediate supervisor. Subjects (incumbents) were drawn from the population of 5-skill levels assigned to CONUS Air Force bases currently supporting F-15 or F-16 aircraft. Within this population, preference was given to recently upgraded subjects, i.e., those subjects with the lowest TAFMS. Sampling was continued until a sample size of 200 was reached or until the available population was exhausted. Special emphasis was given to insuring the

widest possible aptitude range (E or M composite aptitude value) within each AFS.

At ACC bases, incumbents (within each of the 19 AFSs) were asked to provide absolute task performance time estimates (the duration component used by Albert et al., 1994) at each of two time points, i.e., task performance time upon arrival from technical school, and current task performance time. A set of at least 30 tasks (within each AFS) were used to elicit these absolute time duration estimates. Incumbents were also asked to provide task performance experience data in the form of the frequency of task performance within the last 60 days and the number of days since the task was last performed. Additionally, incumbents were asked to provide background information such as time in present job (ITIPJ), level of job satisfaction (ISOA), time with present supervisor (ITWPS), and time in present career field (ITICF).

Supervisors of these incumbents (in conjunction with unit trainers, if necessary) were asked to provide task-level estimates of the incumbent's percent of proficiency upon arrival from technical school and the incumbent's current percent of proficiency. Supervisors/unit trainers were also asked to estimate the number of OJT hours the incumbent required to reach his/her current percent of proficiency level (Bennett Segó, Teachout, & Phalen, 1994). Additionally, supervisors/unit trainers were asked to provide incumbent task performance times (both initial and current) for possible use as an indicator of the validity of time estimates provided by incumbents.

The initial (upon arrival) task-level data provided by incumbents and supervisors/trainers characterized incumbent performance on each separate task (in terms of percent of proficiency and task performance time). This data yielded (across incumbents who will vary with respect to aptitude level) the gain from formal training (in terms of percent of proficiency and task performance time) as a function of aptitude. "Current" percent of proficiency and task performance times yielded the gain from OJT hours expended as a function of aptitude. Finally, use of the "common" percent of proficiency scale allows percent of proficiency to be expressed as a function of both task performance time and task training time.

Data collection efforts were automated, i.e., diskettes containing data collection software were mailed out to base survey control monitors (SCMs). As previously noted, the utility of micro-based survey software has been demonstrated (Albert et al., 1994 and Mitchell et al., 1994). Additionally, survey software (for ACC bases) was easily developed that allowed data collection of both incumbent and supervisor/trainer responses using the same diskette without risk of data contamination, i.e., data entry software was designed to secure supervisors' ratings from incumbents, and secure incumbent responses from supervisors.

Measurement precision (reliability) of proficiency estimates, performance time estimates, and training time estimates were assessed using two relatively different approaches: a) test-retest reliability estimates and b) generalized reliability estimates. The first approach involves resurvey (test-retest of approximately 10 percent of incumbent/supervisor sample) with a small intervening period (approximately five workdays) to estimate the stability of ratings. The period between surveys was short to insure that test-retest differences are a function of error and not systematic changes in incumbent task performance speed or proficiency level. The test-retest approach was feasible only because data collection was automated.

Validation

A portion of the incumbent data was the test/retest data which was collected for approximately 300 of the incumbents. The purpose of the test/retest survey was to validate the time to perform responses. The initial test which was used to determine the validity of the time to perform estimates was a *t*-test of the sample mean for the difference between the test/retest estimates, along with basic mean and standard deviation statistics. If the time to perform estimates are accurately provided by the incumbent, then one would expect for the mean values of the differences between the test and retest values across incumbents by task should center about zero. Thus, the *t*-test for the mean differences would test the hypothesis as to whether the sample mean of the differences is statistically different from zero.

The means test for the differences between the test and retest time to perform values were performed at the task level for each AFS and across tasks for each AFS. The means test by task provided very few statistically significant differences (99% level of confidence) across all 20 AFSs. In most cases, the standard deviations were larger than the mean values. In addition, *t*-test values were calculated and the *t*-test values by AFS indicated that none of the 20 AFSs displayed average differences across tasks which were statistically different from zero.

Another factor which was used to test the credibility of the incumbent time to perform data is the comparison of the incumbent and supervisor estimates. These results were more mixed, especially across AFSs, though in general the majority of the tasks displayed statistically insignificant differences between incumbent and supervisor estimates.

Empirical Results

PF/TTP Equation

The estimated coefficients for the PF/TTP relationship estimated across AFSs is consistent with a quadratic hypothesis. Twenty-eight of the 40 explanatory variables specified are statistically significant at the 99 percent level of confidence. The benchmarked task difficult value (BTDV) (Garcia, Ruck, & Weeks, 1985) was statistically significant and negative which indicates that as the difficulty of learning the task increases the level of proficiency declines, e.g., the more difficult-to-learn tasks generally reflect lower levels of proficiency. AFQT was statistically significant and positive, e.g., higher aptitude incumbents have higher levels of proficiency. The incumbent's experience as reflected by the incumbent's time in the present job (ITIPJ) was statistically significant, as well as the incumbent's job satisfaction (ISOA). The relationship of the supervisor and the incumbent was also a large player in the explanation of the incumbent's proficiency as reflected by STWTA, DSROT, DSRTL, DSRISU, DSROTLP, and DSRWPR. In addition, other incumbent/supervisor factors were statistically significant such as ITIPJ, ITWPS, DSKILL5 (as compared to skill level 3), and DSKILL7.

The explanatory factors which accounted for differences among AFSs, e.g., C2A3X1B, C2A3X1C, were all statistically significant in 18 of the 19 cases. AFS 1C1X1 (Air Traffic Controller) was not statistically different from AFS 2A6X3 (Aircrew Egress - the AFS in the intercept term).

CPF/HT Equation

The estimated coefficients for the change in proficiency/hours of training (CPF/HT) relationship estimated across AFSs is consistent with the expected signs of a quadratic hypothesis. Thirty-four of the 40 explanatory variables specified are statistically significant at the 99 percent level of confidence. The coefficient for HT was positive and statistically significant at the 99 percent level of confidence. In addition, the assumption of a cubic relationship for CPF/HT was also statistically supportable (HT and HT² were statistically significant at the 99% level of confidence, and HT was positive while HT² was negative).

The benchmarked task difficult value (BTDV) was statistically significant and positive which indicates that as the difficulty of learning the task increases the magnitude of change in proficiency increases for a given level of training. AFQT was not statistically significant, though positive, e.g., higher aptitude incumbents have larger changes in the level of proficiency for a given level of training. The incumbent's experience as reflected by the incumbent's time in the present job (ITIPJ) was statistically significant, as well as the incumbent's job satisfaction (ISOA). The relationship of the supervisor and the incumbent was also a large player in the explanation of the incumbent's proficiency as reflected by STWTA, DSROT, DSRTL, DSRISU, DSROTLP, and DSRWPR. In addition, other incumbent/supervisor factors were statistically significant such as ITIPJ, ITWPS, DSKILL5 (as compared to skill level 3), and DSKILL7.

The explanatory factors which accounted for differences among AFSs, e.g., C2A3X1B, C2A3X1C, were all statistically significant in 18 of the 19 cases. AFS 1C1X1 (Air Traffic Controller) was not statistically different from AFS 2A6X3 (Aircrew Egress - the AFS in the intercept term).

Trade-offs Between Aptitude, Hours of Training and Time-to-Perform

Tradeoffs were found to exist between aptitude, hours of training, and time to perform across the AFSs and tasks used in the analysis. Figure 1 provides an example of the tradeoff one might expect between training time (HT) and time to perform (TTP). The empirical analysis performed on the survey data supported the relationships exhibited in Figure 1.

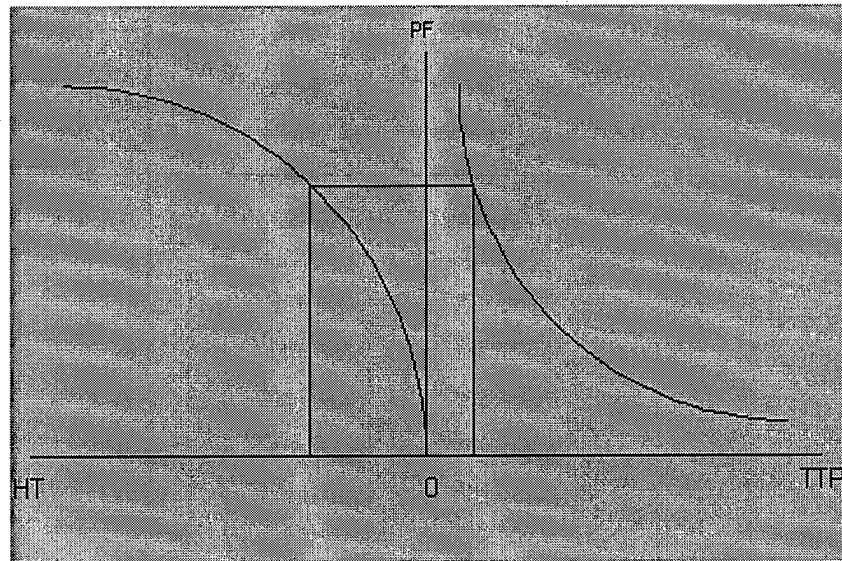


Figure 1. Training/Performance Time Tradeoff

As training time increases, the time-to-perform decreases. The essential mapping between HT and TTP is proficiency (PF). Assuming all other explanatory factors held constant, changes in aptitude will cause changes in TTP and HT.

Conclusions and Recommendations

The data collection and analysis performed for the estimation of the PF/TTP and CPF/HT relationships are relatively unique to the literature. The results afforded by the estimated PF/TTP and CPF/HT equations supported several hypotheses which here-to-fore had not been easy to either gather sufficient information for testing or the data collected did not lend itself readily to the testing of the hypothesis.

Proficiency, based on the GO/NO-GO decision of the immediate supervisor representing a benchmark of 100%, has never successfully gone beyond a concept. The data collected has provided reliable estimates of proficiency and changes in proficiency, sufficient to relate time-to-perform at the task level to training time. The methodology used in this project for collecting the proficiency data and the use of the proficiency data as a tool for modeling training and performance opens new avenues for research and analysis in the training and operational communities. It is a methodology which needs to be further tested and refined, but the present study provides strong evidence of its credibility.

A key point which can not be minimized in the development of the methodology was the intent to develop a scale for proficiency to which operational supervisors and incumbents could easily relate without being required to abstract. Benchmarking was critical to establishing a scale for proficiency which could be understood by the operational community and used by the research community for analyzing manpower, personnel and training issues.

In addition, the technology of computer-assisted surveying greatly enhances the ability to collect information which previously was collected by paper and pencil survey. Computer-assisted surveying provides the opportunity for improved accuracy as this study proves. The design, development, and testing of the survey instrument itself is also critical to the collection of credible data and was quite important for this study.

Task difficulty has, in the past, been expected to be a strong predictor of performance, but very few instances of strong statistical support have been reported (Burch, Lipscomb, & Wissman, 1992). In both sets of estimated relationships, the task difficulty (TDV), both AFS specific and benchmarked values across AFSs, displayed high levels of statistical confidence (99%) and the expected relationship with proficiency (inverse) and the change in proficiency (inverse). Aptitude also displayed signs of statistical significance, though general and administrative composites did not contribute much to the explanation of proficiency to the change in proficiency.

Another key result which was established in the analysis was the importance of other factors (e.g., supervisor position and time in job) in explaining the variation in proficiency. Though not always statistically strong by

AFS, across AFSs these other factors were one of the key reasons for the strong relationships displayed by the key variables (such as time-to-perform, training time, aptitude, and experience). Proper specification or accounting for the predominant factors which affect the variation in task level proficiency was important to improving the likelihood of observing expected key relationships.

References

Albert, W.G., Phalen, W.J., Selander, D.M., Dittmar, M.J., Tucker, D.K., Hand, D.K., Weissmuller, J.J., & Rouse, I.F. (1994). Large-Scale Laboratory Test of Occupational Survey Software and Scaling Procedures. *Proceedings of the 36th Annual Conference of the Military Testing Association, Rotterdam, The Netherlands.*

Bennett, W., Sego, D.J., Teachout, M. S., & Phalen, W. J. (1994). On-The-Job Training Time as a Criterion for Training: A Comparison of Several Task-Based Estimations Approaches. *Proceedings of the 36th Annual Conference of the Military Testing Association, Rotterdam, The Netherlands.*

Burtch, L. D., Lipscomb, M. S., & Wissman, D. J. (1992). *Aptitude Requirements Based On Task Difficulty: Methodology for Evaluation (AFHRL-tr-81-34)*. Brooks AFB, TX: Manpower and Personnel Division Air Force Human Resources Laboratory.

Garcia, S.K., Ruck, H. W., & Weeks, J. L. (1985). *Benchmark learning Difficulty Technology: Feasibility of Operational Implementation (AFHRL-TP-85-33)*. Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Lance, C.E., Hedge, J.W., & Alley, W.E. (1987) *Ability, Experience, and Task Difficulty Predictors of Task Performance (AFHRL-TP-87-14)*. Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Lecznar, W.B. (1971). *Three Methods for Estimating Difficulty of Job Tasks (AFHRL-TR-71-30)*. Lackland AFB, TX: Personnel Division Air Force Human Resources Laboratory.

Leighton, D.L., Kageff, L.L., Mosher, G.P., Gribben, M. A., Faneuff, R.S., Demetriades, E.T., & Skinner, M.J. (1992). *Measurement of Productivity Capacity: A Methodology for Air Force Enlisted Specialties (AL-TP-1992-0029)*. Brooks AFB, TX: Manpower and Personnel Research Division, Human Resources Directorate.

Mitchell, J.L., Yadrick, R.M., & Bennett, W. R. (1993). Estimating Training Requirements From Job and Training Models. *Military Psychology*, 5(1),1-20.

Perrin, B.M., et al, (1988). Training Decisions System: Development of the Task Characteristic Subsystem (AFHRL-TR-88-15). Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Stone, B.M., Turner, K.L., Wiggins, V.L., & Looper, L.T. Measuring Airman Job Performance Using Occupational Survey Data. *Military Psychology*, Vol. 8, Number 3, 1996.

INTERNET DOCUMENT INFORMATION FORM

A . Report Title: Assessing Individual Productivity in the Workplace: An Improved Criterion for Training Evaluation

B. DATE Report Downloaded From the Internet 4/21/99

C. Report's Point of Contact: (Name, Organization, Address, Office Symbol, & Ph #): Navy Education and Training
Professional Development and
Technology Center
Navy Advancement Center Dept
Dr. Grover Diesel, (850) 452-1815
6490 Saufley Field Road
Pensacola, FL 32509-5237

D. Currently Applicable Classification Level: Unclassified

E. Distribution Statement A: Approved for Public Release

F. The foregoing information was compiled and provided by:
DTIC-OCA, Initials: VM_ Preparation Date: 4/22/99__

The foregoing information should exactly correspond to the Title, Report Number, and the Date on the accompanying report document. If there are mismatches, or other questions, contact the above OCA Representative for resolution.