DOCUMENT RESUME

| | | |
|---|---|---|
| ED 064 576 | | AC 012 650 |
| AUTHOR | Prophet, Wallace W. | |
| TITLE | Performance Measurement in Helicopter Training and Operations. | |
| INSTITUTION | Human Resources Research Organization, Alexandria, Va. | |
| SPONS AGENCY | Office of the Chief of Research and Development (Army), Washington, D.C. | |
| REPORT NO | HumRRO-PP-10-72 | |
| PUB DATE | Apr 72 | |
| NOTE | 17p. | |
| EDRS PRICE | MF-$0.65 HC-$3.29 | |
| DESCRIPTORS | Aircraft Pilots; Automation; *Flight Training; *Measurement Techniques; Military Personnel; *Military Training; *Performance Tests; Research; Student Evaluation; Trainees; *Training Techniques | |
| IDENTIFIERS | United States Army | |

ABSTRACT

For almost 15 years, HumRRO Division No. 6 has conducted an active research program on techniques for measuring the flight performance of helicopter trainees and pilots. This program addressed both the elemental aspects of flying (i.e., maneuvers) and the mission- or goal-oriented aspects. A variety of approaches has been investigated, with the stress on nonautomated techniques feasible for operational use. This paper describes the work and illustrates its application to implications for training management, quality control, manpower resources management, and operational capability. Automated human performance monitoring in flight simulators and its implications for automated training is also described. (Author)

Professional Paper 10-72
umRRO-PP-10-72



HumRRO

DoD
AC



# Performance Measurement in Helicopter Training and Operations

Wallace W. Prophet

HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street • Alexandria, Virginia 22314

Presentation at
American Psychological Association
79th Annual Meeting
Washington, D.C. September 1971

April 1972

Prepared for
Office of the Chief of Research and Development
Department of the Army
Washington, D.C. 20310

The Human Resources Research Organization (HumRRO) is a nonprofit corporation established in 1969 to conduct research in the field of training and education. It is a continuation of The George Washington University Human Resources Research Office. HumRRO's general purpose is to improve human performance, particularly in organizational settings, through behavioral and social science research, development, and consultation. HumRRO's mission in work performed under contract with the Department of the Army is to conduct research in the fields of training, motivation and leadership.



Published
April 1972
by
HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street
Alexandria, Virginia 22314

| BIBLIOGRAPHIC DATA SHEET | 1. Report No. HumRRO-PP-10-72 | 2. | 3. Recipient's Accession No. |
|---|---|---|---|
| 4. Title and Subtitle<br>PERFORMANCE MEASUREMENT IN HELICOPTER TRAINING AND OPERATIONS | | | 5. Report Date<br>Apr 72 |
| | | | 6. |
| 7. Author(s)<br>Wallace W. Prophet | | | 8. Performing Organization Rept. No.<br>PP 10-72 |
| 9. Performing Organization Name and Address<br>Human Resources Research Organization (HumRRO)<br>300 North Washington Street<br>Alexandria, Virginia 22314 | | | 10. Project/Task/Work Unit No. |
| | | | 11. Contract/Grant No.<br>DAHC 19-70-C-0012 |
| 12. Sponsoring Organization Name and Address<br>Office, Chief of Research and Development<br>Department of the Army<br>Washington, D.C. 20310 | | | 13. Type of Report & Period Covered<br>Professional Paper |
| | | | 14. |

15. Supplementary Notes HumRRO Division No. 6 (Aviation), Fort Rucker, Ala.; paper presented at American Psychological Association Annual Convention, Washington, D.C., Sept 1971

16. Abstracts

For almost 15 years, HumRRO Division No. 6 has conducted an active research program on techniques for measuring the flight performance of helicopter trainees and pilots. This program addressed both the elemental aspects of flying (i.e., maneuvers) and the mission- or goal-oriented aspects. A variety of approaches has been investigated, with the stress on nonautomated techniques feasible for operational use. This paper describes the work and illustrates its application to and implications for training management, quality control, manpower resources management, and operational capability. Automated human performance monitoring in flight simulators and its implications for automated training is also described.

17. Key Words and Document Analysis. 17a. Descriptors

*Flight tests
*Flight simulation
*Performance evaluation
*Quality control
*Pilot training

17b. Identifiers/Open-Ended Terms

Performance measurement
Proficiency measurement
Flight proficiency
Helicopter training

17c. COSATI Field/Group 05 09 Behavioral and social sciences/Personnel selection, training and evaluation

| 18. Availability Statement | 19. Security Class (This Report) | 21. No. of Pages |
|---|---|---|
| Approved for public release; distribution unlimited. | UNCLASSIFIED | 15 |
| | 20. Security Class (This Page) UNCLASSIFIED | 22. Price |

FORM NTIS-35 (10-70) USCOMM-DC 40329-P71

## Prefatory Note

The research described in this paper was performed for the Department of the Army by the Human Resources Research Organization, Division No. 6 (Aviation), at Fort Rucker, Alabama. The material was derived primarily from Work Unit LIFT, The Development of a Comprehensive Review of Psychological and Sociological Literature on Organized Leadership, but also stemmed from Work Unit ECHO, Synthetic Flight Training Programs and Devices, from INTACT, Integrated Contact/Instrument Training, and from SYNTRAIN, Modernization of Synthetic Training in Army Aviation. Dr. Wallace W. Prophet is the Director of Division No. 6.

# PERFORMANCE MEASUREMENT IN HELICOPTER TRAINING AND OPERATIONS

Wallace W. Prophet

The past two decades have seen a tremendous change in the role of aviation in the U.S. Army. The use of the helicopter has given a dimension and degree of mobility heretofore impossible for ground forces. An idea of the extent of growth in the Army aviation field, can be gai ›d from the following figures. In 1950, the Army had only 715 aviators and 1242 aircraft; by 1960, these totals had risen to 5984 and 5477, respectively; in 1970, there was a total of 22,250 aviators and 11,446 aircraft.

This is truly an amazing growth and reflects the great utility of the helicopter in performing a wide variety of airlift roles. The versatility of these sometimes awkward looking and noisy machines undoubtedly will result in a continuing increase in their application to both civil and military needs. For example, the helicopter is already being used for such diverse activities as transportation of persons; oil exploration; patrolling of forests, game preserves, power lines, and pipelines; traffic control and other aspects of law enforcement; medical evacuation; and heavy construction.

This great increase in the use of aircraft, particularly helicopters, in the Army has brought with it a tremendous expansion in the Army's flight training program. For example, for fiscal years 1966 through 1970, the Army graduated the following numbers of initial entry pilots: 1966, 1869; 1967, 4257; 1968, 5295; 1969, 7699; and 1970, 7525.

The progress of training for the initial entry student has been approximately as follows: The helicopter, or rotary wing, student began his primary training (110 flight hours over 16 calendar weeks) at the U.S. Army Primary Helicopter School, Fort Wolters, Texas. He then moved on to either Hunter Army Air Field, near Savannah, Georgia, or to Fort Rucker, Alabama, to complete the remainder of his training (100 hours over 16 weeks). Upon graduation from this sequence of training, the new aviator received his wings and was assigned to an operational unit, usually in Vietnam. The fixed wing trainee received a similar amount of instruction, except that his primary instruction (110 hours over 16 weeks) was given at Fort Stewart, Georgia, and he then moved to Fort Rucker to complete his training (100 hours over 16 weeks).

As in any educational or training system, measurement plays an extremely critical role in flight training. The requirements for achieving psychometrically sound performance measures are well known. However, these problems are vastly compounded when the performance measured is as complex as that required in flying. Furthermore, the functional uses to which the measurements may be put are multifarious. Overlying these considerations are the facts that aviation training is very expensive and adequacy of performance may have life or death consequences for the pilot and perhaps others.

HumRRO has just completed its 20th year as a research and development organization concerned primarily with human functioning and performance in the world of work. For the last 15 of those 20 years, HumRRO has been working actively on problems related to aviation training and flight performance. This paper will describe some of the work relating to performance measurement.

An examination of some of the uses to which flight performance measures may be put will provide a background for the discussion of HumRRO's flight performance research program. Our research is marked by emphasis on pragmatic, utilitarian aspects. First, the most obvious application of performance measurement to the individual trainee is the determination of who passes and who fails. Here, performance measures refer to measures of achievement in the flight training program, such as the periodic checkrides given at various points during training. Failure to perform satisfactorily results in elimination from the training program. However, the extent to which these measures are predictive of, or relevant to, future pilot performance is also of concern. This predictive function is of special importance to the Army, because, unlike his fellow pilots in the other services, the Army pilot assumes duties in an operational unit (usually in combat) immediately after the completion of his undergraduate pilot training (UPT). In contrast, the Air Force UPT graduate goes on for further training and assessment at a Combat Crew Training School, while the graduate of Navy UPT goes to a Replacement Air Group Squadron for such work.

Other uses of flight performance measures focus on the individual pilot trainee, principally by the individual flight instructor. He may use daily flight grades or performance records as a basis for counseling the trainee and for modifying the training presentation. He may also use grades in an attempt to motivate the student through selective reinforcement.

The great majority of past research on flight performance measurement has tended to concentrate on the use of such measures with the individual student. Hence, we have seen much research dealing with reliability, as in the stability of trainee performance (or measures of his performance) from one occasion to another, based upon the classical test-retest paradigm. Studies of checkpilot flight standards (i.e., interobserver reliability) also fall in this area.

Although this concern with the individual in training is paramount, over the past decade flight performance measures have also been used as a management tool. In these functional areas fall recruitment, selection, and general manpower management. Also, we have been quite concerned with applications of the concept of quality control in aviation training systems. Here, the focus is on feedback to the training system from those external operational systems with which it interfaces, as well as with feedback loops entirely within the training sytem. One particularly critical feedback loop between the training system and the criterion world of flight operations involves aviation safety. The goal of such efforts is the development of a continuing, dynamic means for adjusting and regulating aviation training sytems. All applications assume sound indices of performance.

One of the first areas of aviation psychology investigation undertaken by HumRRO was helicopter flight performance evaluation methods. In a series of studies by Greer, Smith, and Hatfield (1), attempts were made to develop more objective and reliable means of evaluating student performance in helicopters. Initial investigation indicated that the traditional subjective grading system then in use had quite low reliability, a finding consonant with those reported in earlier summaries of research on reliability of subjective checkrides (e.g., see Erickson 2, and Ben-Avi, 3). Correlations of daily training grades and flight check grades during rotary wing primary training were typically less than .30, while those between checkrides given at various points in training were as low or lower.

Building on previous Air Force work by Smith, Flexman, and Houston (4), Greer *et al*, (1) developed a series of relatively objective flight performance checklists called Pilot Performance Description Records (PPDR). In constructing the PPDR, each maneuver was analyzed in detail, and as many items or scales as possible describing specific pilot and aircraft behaviors during the maneuver were developed. Where feasible, objective indices were used, such as airspeed, altitude, and RPM indications. In each case, the item and conditions for its observation were carefully defined. Figure 1 illustrates a page from a rotary wing PPDR. Similar instruments have been developed by Prophet and Jolley (5) for use in fixed wing flight measurement.

## Sample Page From the Pilot Performance Description Record (PPDR)

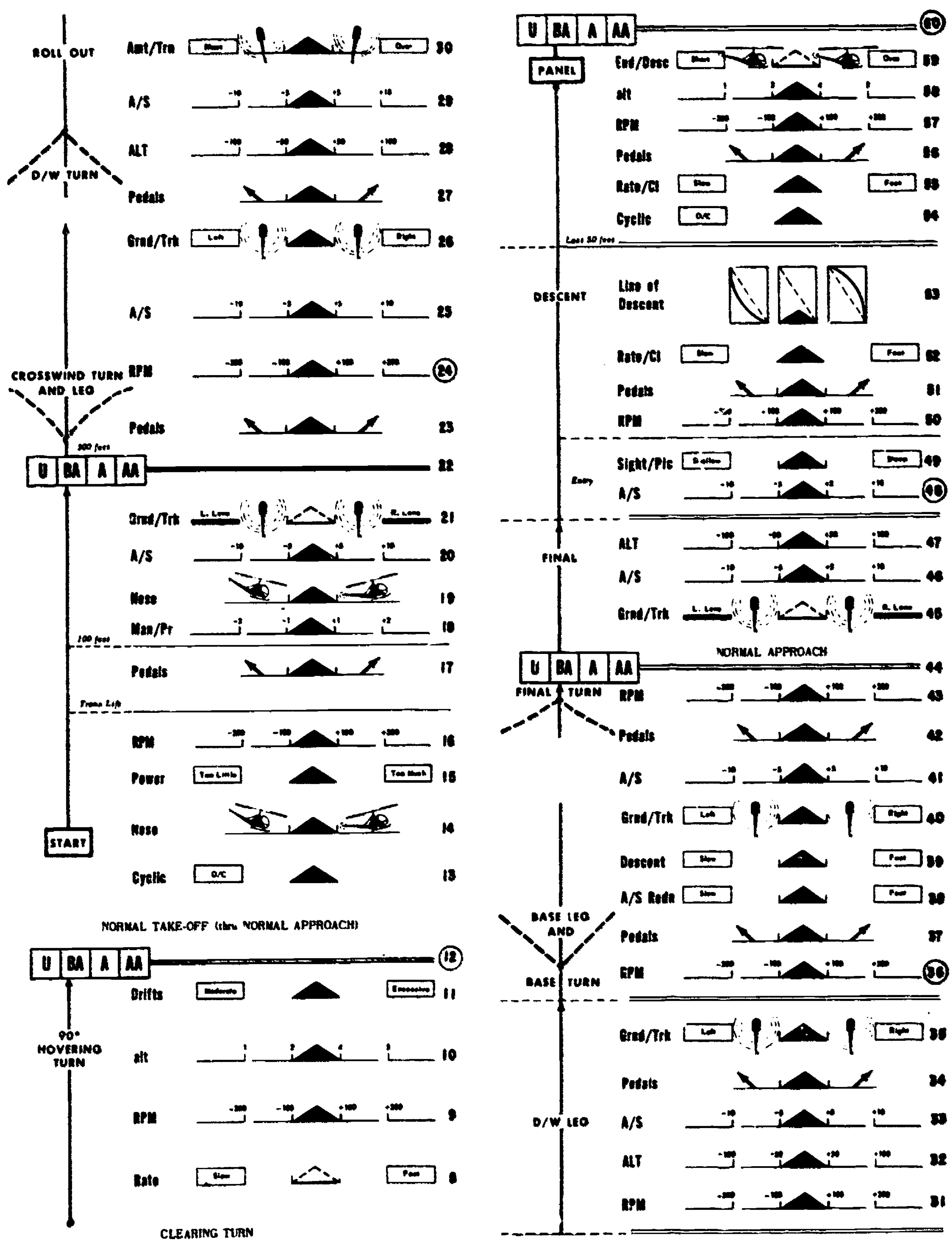


Figure 1



7

After much research effort, the PPDR was installed as an integral part of the flight evaluation program at the Primary Helicopter School. Change from an existing subjective flight evaluation system to a new, albeit more objective, system such as the PPDR does not come easily. Without the support of top management such a change could not have been made. Note should be taken of the utter necessity for systematic and thorough training of the checkpilots in the use of the new instrument *before* they try it operationally with real students.

Before-and-after results of the use of the PPDR are shown in Table 1. These data show correlations between mean daily flight grade for a phase of training with the grade on the checkride administered at the end of that phase. The subjective system did not yield statistically significant correlations, while grades derived from the PPDR correlated significantly with training grades.

Table 1

**Correlation of Mean Daily Grade and Checkride Grade by Stage of Primary Training**

| System | Stage of Training | |
|---|---|---|
| | Intermediate | Advanced |
| Subjective | .08 | .09 |
| Objective (PPDR) | .42* | .51* |

*$p < .05$.

The PPDR has been used to provide a greater degree of standardization and objectivity to the flight evaluation process. Because of the considerable detail in which it describes the desired or proper performance of a maneuver, the PPDR is quite useful as a pedagogical tool. First, it conveys to the student rather precisely the performance objectives he seeks to achieve and the items on which he will be evaluated. Typically, for a checkride of an hour's duration, approximately 250 separate behavioral observations are recorded on the PPDR by the checkpilot. The PPDR also standardizes and defines for the student and the checkpilot, the sequence of events on the checkride. A second major use of the PPDR is for detailed postflight feedback to the student on his performance. This feature of the PPDR has been found to be very useful.

These uses of the PPDR are examples of research applications to problems of teaching and evaluating the individual student. However, we have extended this approach to problems that are more systemic in nature, through use of the PPDR as part of a training quality control system. The individual performance items or scales are used as input to automatic data processing by which the performances of large groups of students, for example, a flight class, can be summarized. These performances may then be compared with a school standard (Figure 2). Probability tables were developed to allow evaluation of the statistical significance of a given maneuver's variation from its school standard, based upon the number of cases involved, and the intramaneuver performance variability.

The analysis is carried a step further in Figure 3 in which individual critical maneuvers are examined. In this way, causes of deviant performance can be examined at a more detailed level, and specific remedies can be developed.

Data summaries of this sort provide management with a powerful tool to use in evaluating and adjusting the training system. The advantage of such data is that they are

## Evaluation of Class Performance (Errors vs. Maneuvers)

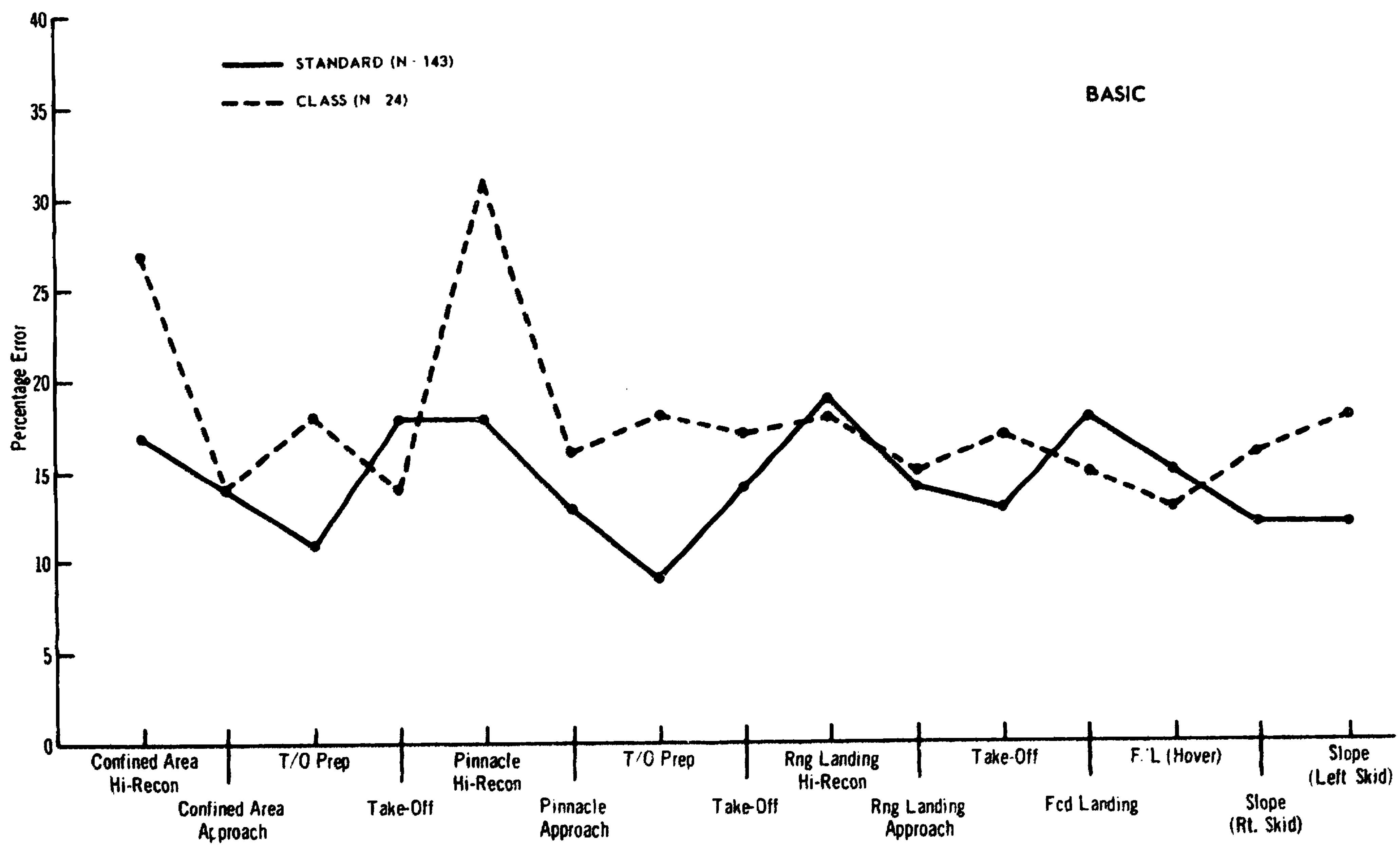


Figure 2

performance *observations* rather than performance *evaluations*. The functioning of instructors and checkpilots, both by groups and individually, can also be examined in this way, and corrective actions taken if the data indicate shortcomings of either the instruction or evaluation systems. Details of this flight training quality control system are presented in the work of Duffy and Colgan (6). Implications of this approach for other training systems have been discussed in other HumRRO publications by Smith (7, 8).

**Class Performance on Critical Maneuvers (Error vs. Items)**

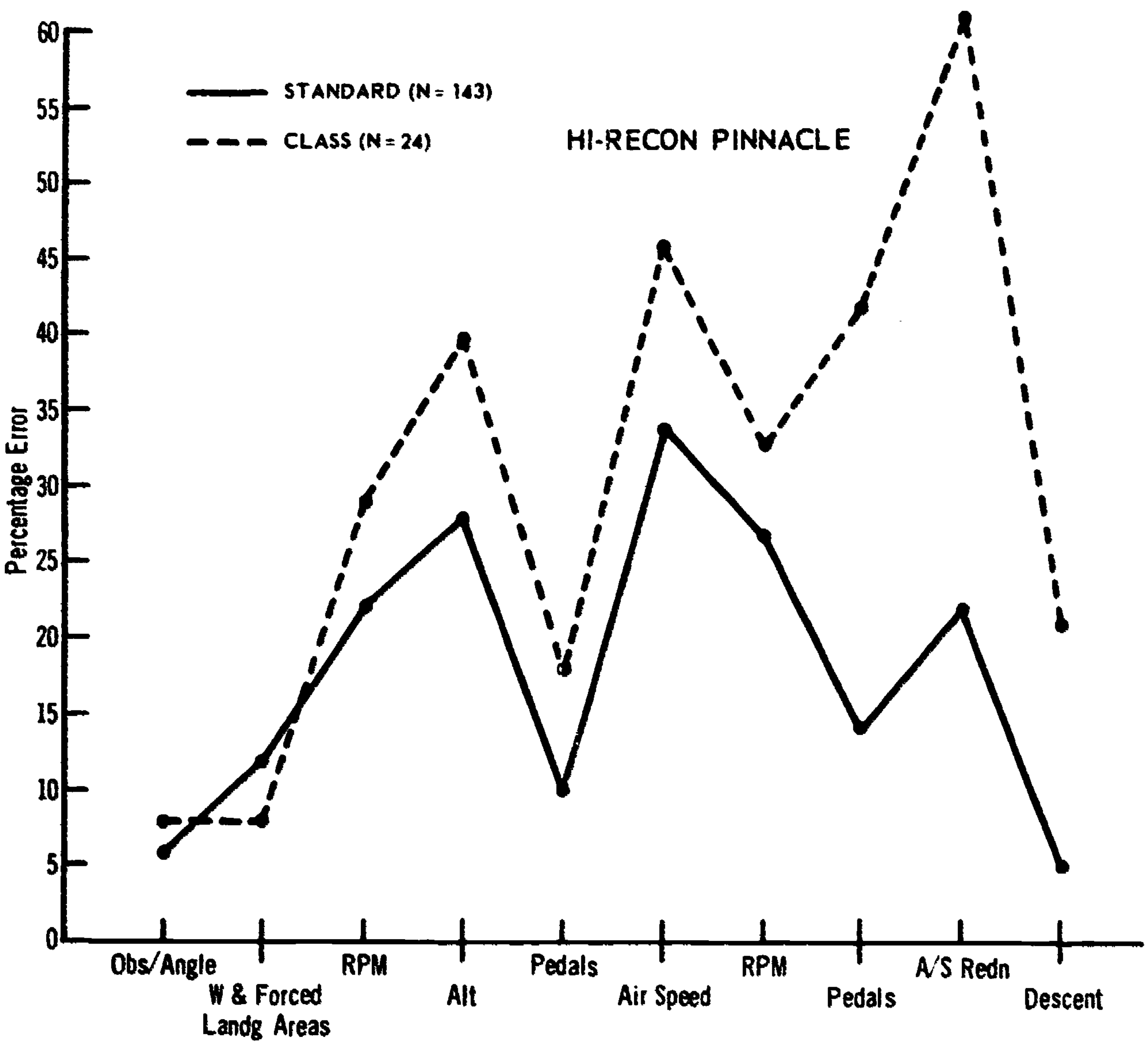


Figure 3

This application of the quality control concept might also be described under the accountability concept that has come into prominence recently in education circles. The instructor is held accountable, so to speak, for the quality of his output—that is, the performance of his students. The detailed PPDR data summary for a number of students of a given instructor provides specific means of counseling with that instructor and for adjusting and standardizing his instruction as necessary.

This means of evaluating the flight instructor assumes a random assignment of students to instructors so that student aptitude differences do not account for instructor output differences. In more recent work, we have been seeking to control nonrandom, interstudent difference effects through use of multiple regression predictions of student performance in flight training. The discrepancies between the actual performances of all students of a given instructor and those performances predicted from multiple regression equations will provide a more precise measure of the instructor's effect on his students than discrepancies based on the random assignment model.

In another series of studies, Caro (9) has investigated the effects of prior knowledge on checkride evaluations. In one portion of advanced helicopter training, checkrides are administered by instructor pilots from within the same instructional flight. The instructors simply trade students at checkride time, giving the checkpilot (instructor) an opportunity to learn something about the prior performance of the student before he administers the checkride. At the least, he knows who the student's instructor is and probably has ideas on what kind of student that instructor usually turns out.

In several classes we brought in qualified checkpilots from *outside* the instructing flight to administer checkrides to a portion of the class. The remainder of the class was administered checkrides in the usual manner by instructors from *within* the flight. The correlations between instructor evaluation and checkride grade for these two conditions are contrasted in Table 2. Those checkrides involving no prior information (i.e., the "Special" group) showed negligible correlation, whereas those done from within the instructing flight (the "Regular" group) showed substantial correlation. From these data, Caro concluded that prior knowledge of the student, rather than similarity of evaluation standards, may have accounted for the higher correlations of the "Regular" group.

The focus of research so far described is the standardization process. Aviation training managers recognize the need for standardization and devote much effort to its achievement. In spite of this extensive effort, it is difficult, and perhaps impossible, to achieve a substantial degree of standardization using highly subjective measures. These are matters of considerable concern in military flight training programs. For example, Figure 4 shows mean checkride grade and ±1 standard deviation range for checkride grades given by 17 checkpilots. The interrater differences are considerable. Analysis of variance shows these differences to be statistically significant ($p < .001$). Such variation is obviously not desirable, but it seems to be an inevitable part of subjective evaluation.

Checkpilot variation among seven checkpilots at Fort Wolters is shown in Figure 5. Four of these were then given training in the use of the PPDR. The effect of this training on their relative standardization is shown in Figure 6.

Table 2

**Correlations Between Instructor and Checkpilot Evaluations[a]**

| Checkride | Stage of Training | | | | | |
|---|---|---|---|---|---|---|
| | Pre-Solo | | Advanced | | Instrument Cross-Country | |
| | N | Correlation | N | Correlation | N | Correlation |
| Special | 36 | .28 | 40 | .20 | 44 | .18 |
| Regular | 24 | .55[b] | 20 | .73[b] | 18 | .64[b] |

[a]Classes 63-1W and 63-3.
[b]Coefficiencies significantly greater than zero ($p < .01$).

Mean and ± 1 SD Range for Checkride Grades
Assigned by 17 Checkpilots

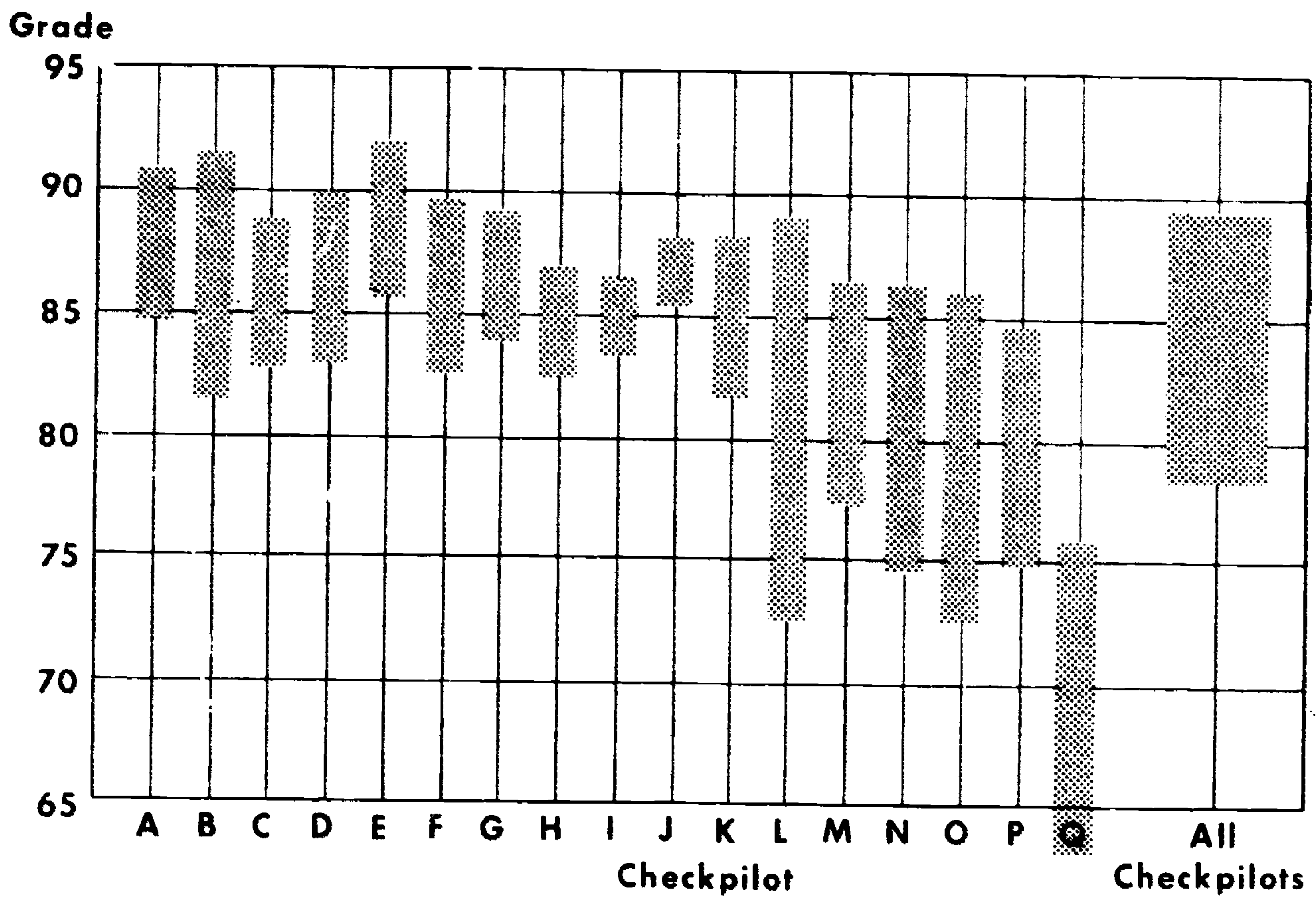


Figure 4

Greer *et al* (1) report that inter-checkpilot flight-check agreement increases as a function of their degree of similarity in classroom evaluations of already marked PPDRs. For example, checkpilots whose evaluations during classroom training correlated between .95 and .99 showed flight interobserver, or test-retest, correlations of .70 for the intermediate level check, whereas those unselected on the basis of their classroom agreement showed flight correlation of only .42. Similar data for the advanced checkride showed correlations of .61 and .52 for the classroom--similar and unselected groups. Thus, use of the PPDR or similar techniques offers an indirect means of increasing checkpilot standardization.

More recently, our work has been concerned with multiple regression approaches to predicting student performance. In this effort, described by Boyles and Wahlberg (10), a computerized data bank was developed for predicting a variety of aviator performances. Our system, which owes much to the conceptions so ably developed by Miss Ambler and her colleagues at Pensacola (e.g., see Schoenberger, Wherry, and Berkshire, 11), presently contains over 100 predictor variables. Included are variables such as aptitude and ability measures, demographic data, education, academic grades, and daily and checkride flight grades.

Data are in the computer for over 12,000 students now, with several thousand more records in partial stages of completion. We are building toward not only the prediction of training performance, but prediction of a variety of operational flight performances such as combat, flight safety, and instructor effectiveness. We view the data bank as a

Mean and ± 1 SD Range for Checkride Grades Assigned by Seven Checkpilots (Primary Checkride - Before PPDR)

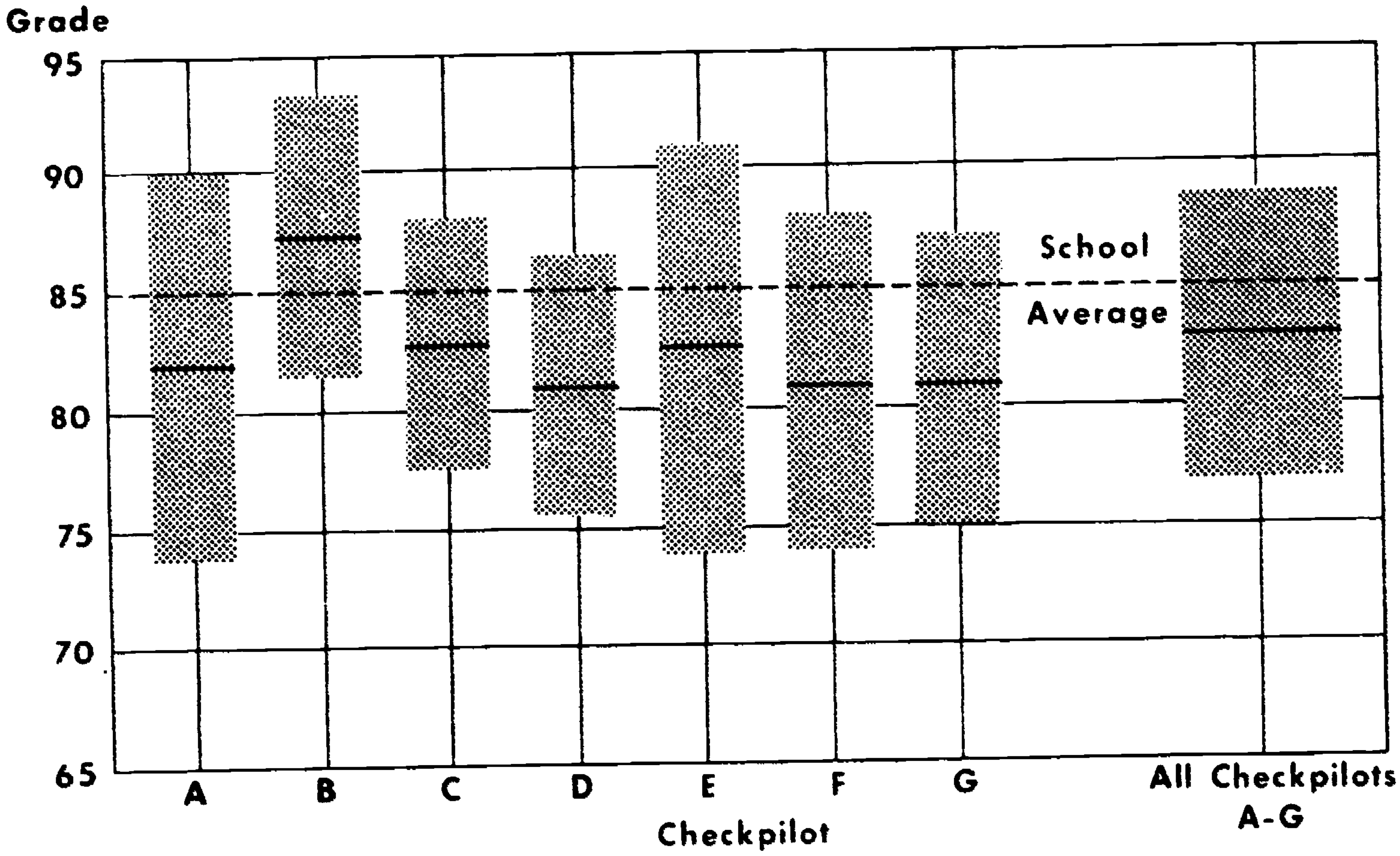


Figure 5

Mean and ± 1 SD Range for Checkride Grades Assigned by Four Checkpilots (Primary Checkride - After PPDR)

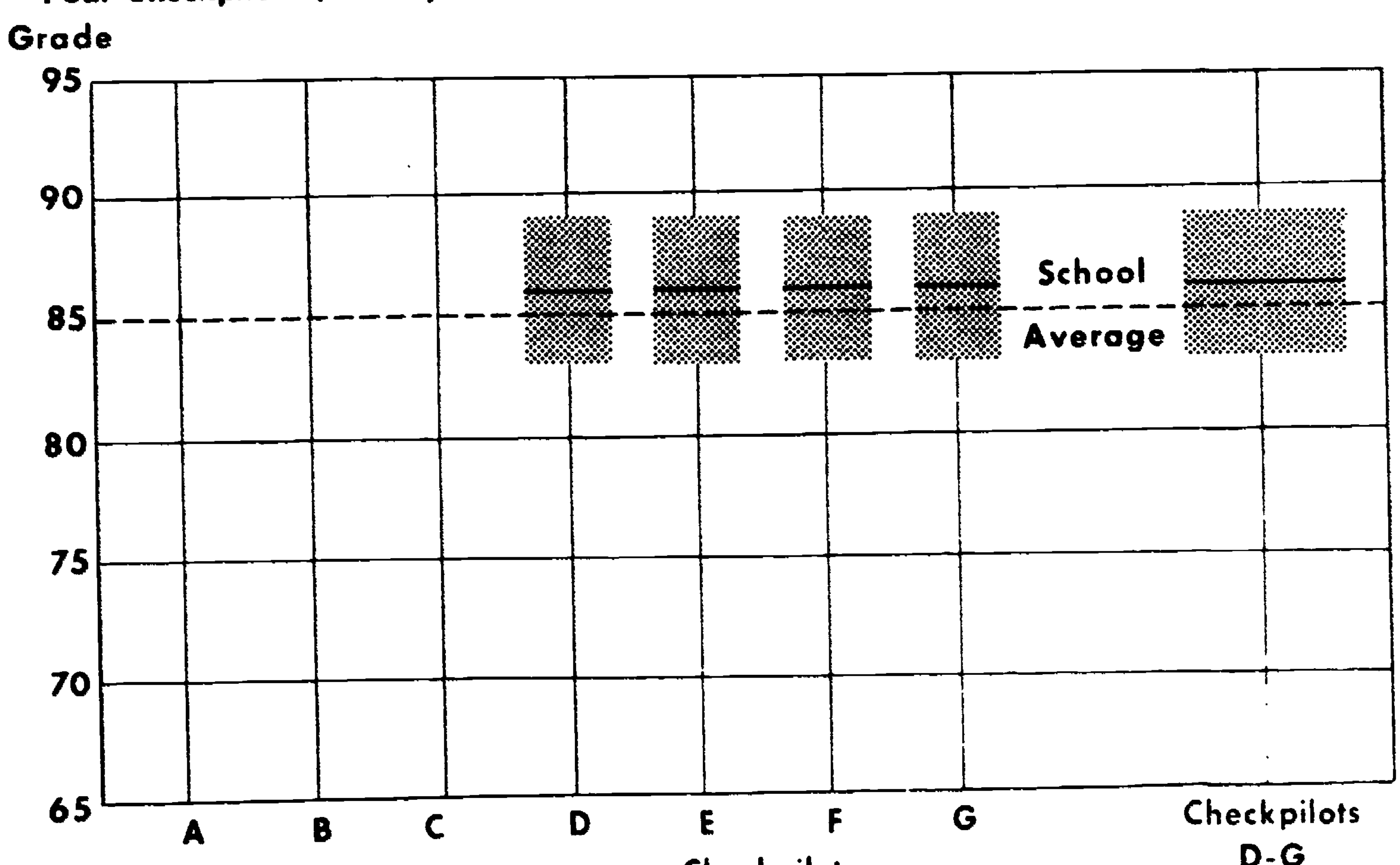


Figure 6

longitudinal one. Our biggest need now is for predictor variables to account for aspects of operational performance variance independent of training performance, that is, motivational factors.

There are several questions relating to the quality or kind of data from which multiple predictions are made that are of interest here. In a study of the use of a captive helicopter as a training device, Caro, Isley, and Jolley (12) report data concerning the predictability of subsequent flight performance from performance on the device.

They gathered 50 separate objective measures of performance on the captive helicopter device during a preflight device training program either 3 1/4 or 7 1/4 hours in duration. These measures were then correlated with mean daily flight grade, time to checkride, and checkride grade for the pre-solo, intermediate, and advanced stages of primary training.

Maximum correlations with the three pre-solo stage criterion measures were shown by certain device measures reflecting cumulative time to achieve basic hovering control of the device; these correlations ranged from .52 to .60. At the intermediate stage (i.e., the first 50 hours), these same measures, plus a measure of lateral right tracking error and one of turn rate during right turns, showed the maximum correlation with flight performance; correlations ranged from .38 to .46. At the advanced stage (i.e., the 100-hour level), several measures of precision hovering, which involved maintaining a probe attached to the front of the device inside either a 10-inch or 14-inch hoop without touching it, showed the highest correlations; values ranged from .44 to .52.

Considering the time lapse between the device training and the advanced checkride (over four months) and the previous comments on inter-checkride correlation, these latter relationships are quite high.

The precision hover task involving the hoop and probe is particularly interesting. Students were able to master this task relatively easily on the device and to perform it quite proficiently. However, expert helicopter pilots had great difficulty with this particular task, even though they could hover the device well. Their difficulty stemmed from their inability to use the visual cue sources for the hoop and probe so close to their eyes (about six feet). Experienced pilots gather their hovering information from more distant sources. It is interesting that this artificial task put in for training purposes only, one that lacked face validity in the eyes of experienced pilots, was one of the more effective for predicting performance at all stages of training and was the most effective for predicting advanced performance.

These data would suggest that early flight performance—for the student's performance on the device can be considered early flight performance—should be predictive of subsequent flight performance. However, data from our multiple prediction study show that the first five graded helicopter flights correlate only .32 with subsequent pass-fail, a correlation similar to that reported by Schoenberger *et al* (11) for presolo grades at Pensacola.

In contrast, an earlier study of fixed wing training by Prophet and Jolley (5) showed substantial correlation between early flight performance and subsequent success in the program. A score based on the sum of errors made on seven selected flight maneuvers on the first three days of flight showed product-moment correlation of .50 with checkride performance at the 35-hour level. Inclusion of the first five days raised the correlation to .63. Biserial correlations of these errors with pass-fail were .62 and .76 for the three- and five-day periods, respectively.

These three sets of data present contrasting results on the predictability of subsequent flight performance from measures of early or preflight performance. Those early or preflight measures which showed substantial correlation with later flight performance were based upon objective or relatively objective indices. The predictor measures of Caro *et al* (12) were based upon time to criterion, frequency counts, time measures, and similar data. Those of Prophet and Jolley (5) were based upon a PPDR-like daily flight record on which specific performances were noted for such indices as altitude and airspeed. In contrast, the data reported in our multiple regression study and by the Navy are based on subjective grades of daily flight performance, (above average, average, below average, and unsatisfactory). Thus, flight performance may be reliably predicted only if the proper kinds of data are gathered. No only do these observations have implications for the kinds of data that should be provided by checkpilots and instructors, but they suggest that the area of psychomotor selection testing is ground that needs replowing.

These points are illustrated in Table 3. These are the same fixed wing data (Prophet and Jolley, 5) that produced correlations of .62 and .76 with pass-fail for three and five days, respectively. It can be seen that most of the successful students were different from the washouts from the very first day of training. Also, note that the washouts show practically no improvement in performance over the entire five days. This suggests the possibility that the initial selection screen let through a number of students who should not have entered training. Perhaps a good psychomotor test might have picked them up.

These same data may also indicate that our training is grossly inappropriate for substantial segments of our input population. The results of Caro *et al.* (12) would support this hypothesis, for they found a significant reduction in flight deficiency attrition as a result of the preflight device training. Perhaps, we should ask whether our training systems possess sufficient flexibility to individualize instruction to meet the needs of these students who have difficulty, seemingly, from the beginning of the program. The captive helicopter provided the students of Caro *et al.*, a relatively low stress environment in which to learn certain skills and to develop confidence in themselves. It was also unique in that the students received full, immediate, and often emphatic feedback concerning the results of their control actions. While this is somewhat peripheral to the main subject, there appears to be a crying need for definitive research on what flight students do or don't learn and why.

While the PPDR approach has done much to improve the quality of flight performance measurement for portions of the Army's flight training system, there is need for examination of other new approaches. The PPDR requires a thoroughly trained checkpilot, and this training may not always be feasible. For example, we have often used time-lapse photographic techniques to gather flight data (Isley and Caro 13). However,

Table 3

**Percent Error for Seven Selected Manuevers by Day of Training and 35-Hour Check Grade**

| Group Based on 35-Hour Check | N | Percent Error: Training Day 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Pre-35-hour washouts | 13 | 63 | 62 | 60 | 57 | 57 |
| 35-hour washouts | 3 | 60 | 51 | 58 | 45 | 55 |
| 35-hour grade=70-74 | 6 | 66 | 54 | 48 | 34 | 35 |
| 35-hour grade=75-79 | 10 | 42 | 40 | 36 | 30 | 40 |
| 35-hour grade=80-84 | 8 | 50 | 35 | 34 | 32 | 23 |
| 35-hour grade=85-89 | 7 | 49 | 40 | 29 | 28 | 22 |
| 35-hour grade=90-94 | 9 | 45 | 35 | 29 | 22 | 24 |

data reduction is time consuming. Airborne videotape techniques seem to offer promise, as do other airborne data recorders. We are following the work of the Air Force in this area with considerable interest.

There seems little doubt that future major gains in the effectiveness and efficiency of flight-proficiency measurement techniques will involve forms of automated measurement. This is particularly relevant for the basic perceptual-motor control skill areas. However, we must not lose sight of the fact that operational flying involves complex decision making and cognitive factors overlaid on these control skills. This is the real challenge for measurement research in aviation. I am not convinced that these complex, mission-oriented factors can be sensed, transduced, and then recorded adequately by hardware.

The only real application of automated performance-measurement techniques in helicopter training is in the Army's Synthetic Flight Training System (SFTS) currently undergoing test at Fort Rucker. This system has the capability of automatically administering training, recording and evaluating trainee performance, adapting problem difficulty level to manifest performance, and sequencing the trainee to the next step in the training program. The lack of hard data on how trainees actually perform in maintaining various flight parameters within tolerance envelopes and the manner in which these envelopes change over time makes automatic measurement difficult. However, shortly we should be able to develop a much more complete and valid picture of training performance as we work with this device. We also intend to explore quality control applications with the SFTS equipment, both in the school training situation and in operational helicopter units.

In summary, the Army has made progress in its flight-measurement programs over the past 15 years. The PPDR system is the most objective and detailed flight performance measurement system in operational use in a military flight training program. The applications of quality control techniques in the Army represent substantial advancement, both at the line instructional level and at the training system management level. However, we have a long way to go before the students no longer perceive that their fate is largely in the sometimes capricious hands of the "Santa Clauses" and "Hardnoses" who check them.

## LITERATURE CITED

1. Greer, George D., Jr., Smith, Wayne D., and Hatfield, Jimmy L. *Improving Flight Proficiency Evaluation in Army Helicopter Pilot Training*, HumRRO Technical Report 77, May 1962.

2. Erickson, Stanford C. *A Review of the Literature on Methods of Measuring Pilot Proficiency*, Air Training Command Human Resources Research Center, Research Bulletin 52-25, Lackland Air Force Base, San Antonio, Texas, August 1952.

3. Ben-Avi, Avrum H. "Studies of Subjective Measures of Flying Proficiency," in Miller, Neal E. (ed.), *Psychological Research on Pilot Training*, Aviation Psychology Program Research Report No. 8, U.S. Army Air Forces, 1947, pp. 83-109.

4. Smith, J.F., Flexman, R.E., and Houston, R.C. *Development of an Objective Method of Recording Flight Performance*, Air Training Command Human Resources Research Center, Technical Report 52-15, Lackland Air Force Base, San Antonio, Texas, December 1952.

5. Prophet, Wallace W., and Jolley, Oran B. *Evaluation of the Integrated Contact-Instrument Concept for Army Fixed Wing Flight Instruction*, HumRRO Technical Report 69-26, December 1969.

6. Duffy, John O., and Colgan, Carroll M. *A System of Flight Training Quality Control and Its Application to Helicopter Training*, HumRRO Consulting Report, June 1963.

7. Smith, Robert G., Jr. *Controlling the Quality of Training*, HumRRO Technical Report 65-6, June 1965.

8. Smith, Robert G., Jr. *An Annotated Bibliography on Proficiency Measurement for Training Quality Control*, HumRRO Research Memorandum, June 1964.

9. Caro, P.W., Jr. *Flight Evaluation Procedures and Quality Control of Training*, HumRRO Technical Report 68-3, March 1968.

10. Boyles, Wiley R., and Wahlberg, James L. *Prediction of Army Aviator Performance: Description of a Developing System*, HumRRO Professional Paper 5-71, April 1971.

11. Shoenberger, R.W., Wherry, R.J., Jr., and Berkshire J.R. *Predicting Success in Aviation Training*, Research Report No. 7, U.S. Naval School of Aviation Medicine, Naval Air Station, Pensacola, Fla., September 1963.

12. Caro, Paul W., Jr., Isley, Robert N., and Jolley, Oran B. *The Captive Helicopter as a Training Device: Experimental Evaluation of a Concept*, HumRRO Technical Report 68-9, June 1968.

13. Isley, Robert N., and Caro, Paul W., Jr. "Use of Time-Lapse Photography in Flight Performance Evaluation," *Journal of Applied Psychology*, vol. 54, no. 1, February 1970; issued as HumRRO Professional Paper 10-70, April 1970.