

A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS

**Bart B. Cobb, M.S.
John J. Mathews, B.A.
FAA Civil Aeromedical Institute
P.O. Box 25082
Oklahoma City, Oklahoma 73125**



May 1972

Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia 22151, for sale to the public.

Prepared for
**DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
Office of Aviation Medicine
Washington, D.C. 25091**

Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute, Oklahoma City, Oklahoma. A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS, by Bart B. Cobb and John J. Mathews, May 1972, 26 pp. Report No. FAA-AM-72-18.

I. Cobb, B. B.
II. Mathews,
J. J.

Descriptors
Air Traffic
Controllers
Aptitudes
Personnel
Selection

This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations

of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.

Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute, Oklahoma City, Oklahoma. A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS, by Bart B. Cobb and John J. Mathews, May 1972, 26 pp. Report No. FAA-AM-72-18.

I. Cobb, B. B.
II. Mathews,
J. J.

Descriptors
Air Traffic
Controllers
Aptitudes
Personnel
Selection

This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations

of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.

Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute, Oklahoma City, Oklahoma. A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS, by Bart B. Cobb and John J. Mathews, May 1972, 26 pp. Report No. FAA-AM-72-18.

I. Cobb, B. B.
II. Mathews,
J. J.

This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations

Descriptors
Air Traffic
Controllers
Aptitudes
Personnel
Selection

Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute, Oklahoma City, Oklahoma. A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS, by Bart B. Cobb and John J. Mathews, May 1972, 26 pp. Report No. FAA-AM-72-18.

I. Cobb, B. B.
II. Mathews,
J. J.

This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations

Descriptors
Air Traffic
Controllers
Aptitudes
Personnel
Selection

of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.

of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.

1. Report No. FAA-AM-72-18	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS		5. Report Date May 1972	6. Performing Organization Code
7. Author(s) Bart B. Cobb, M.S. John J. Mathews, B.A.		8. Performing Organization Report No.	
9. Performing Organization Name and Address FAA Civil Aeromedical Institute P. O. Box 25082 Oklahoma City, Oklahoma 73125		10. Work Unit No.	11. Contract or Grant No.
12. Sponsoring Agency Name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, D. C. 20591		13. Type of Report and Period Covered OAM Report	
15. Supplementary Notes This research was conducted under Tasks No. AM-A-70-PSY-23, AM-B-71-PSY-23, and AM-B-72-PSY-23.		14. Sponsoring Agency Code	
16. Abstract <p>This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols; and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.</p>			
17. Key Words Air Traffic Controllers Aptitudes Personnel Selection		18. Distribution Statement Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia 22151, for sale to the public.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 26	22. Price \$3.00

ACKNOWLEDGMENTS

The assistance of Peter Nelson, Carolyn Lay, Dana Hauserman, and Barbara Mallett in the collection, processing, and/or analysis of the data is gratefully acknowledged.

A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS

I. Introduction.

This report pertains to the development and experimental validation of a new, highly-speeded, and rather novel paper-and-pencil type aptitude test and discusses its potential as an additional screening device for improving the validity of the Federal Aviation Administration's (FAA) selection of personnel for training as Air Traffic Control Specialists (ATCSs). The test, referred to as the "Directional Headings Test" (DHT), was developed by the Aviation Psychology Laboratory of the FAA's Civil Aeromedical Institute (CAMI). In seeking to establish whether the findings justify operational implementation of the DHT, it is important first to consider the nature and effectiveness of tests currently being used in the selection process.

The Current ATC Aptitude Test Battery

A. *Validation of CSC Test Battery.* Prior to 1964, the FAA's screening and selection of personnel for ATC training generally involved no formal assessment of an applicant's mental abilities or aptitudes. Beginning in July 1962, and for 18 months thereafter, aptitude measures were considered in the selection of a limited number of trainees who possessed little or no previous job-related experience. From January 1964 through October 1968, a qualifying aptitude index, reflecting performance on a battery of six Civil Service Commission (CSC) tests, constituted a major eligibility requirement of *all* applicants—regardless of their experience and other qualifications. The aptitude screening index was an outgrowth of extensive research conducted by CAMI during 1961 through 1963.^{4 18 19}

Specifically, the six CSC tests were identified from among 27 experimentally administered instruments as yielding the best composite measure (of aptitudes) for prediction of training performance. One of the tests, designated as

CSC Booklet 24, provided a measure of numerical ability. Others were: CSC 51, which pertained to spatial relationships; CSC 135, "Following Oral Directions"; "Abstract Reasoning" and "Letter Sequence," both of which were parts of CSC Booklet 157; and a special, omnibus test called "Air Traffic Problems" which the American Institute of Research developed in 1950 for the CAA. The latter was ultimately designated as "CSC 540." In that research, the CSC Battery was administered to a total of 893 men as they entered Academy ATC training. The results of subsequent analyses included a product-moment correlation of .51 between the CSC composite scores and a variable which reflected an average of each individual's examination grades and laboratory performance ratings, and a point-biserial correlation of .40 between the CSC measure and Pass-Fail status. More importantly, an examination of the data to determine a "cutting score" for screening purposes revealed that 182 (or 67.1 per cent) of the 271 training-course attritions failed to attain CSC scores of 190 or higher as compared with 222 (or only 37.5 per cent) of the 622 who successfully completed the course. However, in developing the new standards for implementation in January 1964, officials chose to establish 210, rather than 190, as the minimum qualifying score. This relatively high screening hurdle was instituted because records indicated that the number of applicants had, for several months, vastly exceeded the number of ATC trainee positions, and a further reduction in the recruiting of ATC personnel was anticipated. As shown in the upper portion of Appendix A, 576 (or 64.5 per cent) of the 893 trainees involved in the validation study scored less than 210 on the CSC Battery. Although 357 (or 62 per cent) of the 576 passed Academy training, 219 did not, and the 219 represented 81 per cent of the 271 training failures within the entire sample. Selection

officials were particularly impressed by the latter finding and, inasmuch as further curtailments in recruiting were anticipated, they subsequently decided that candidacy for appointment to ATC training should require a CSC score of at least 210.

B. Current Aptitude Screening Standards. Under current procedures, which were adopted in November 1968, applicants having exceptional amounts or types of previous job-related experience (and particularly former military controllers with specialized ATC experience) may be exempted from the aptitude screening requirement and, in many instances, may be appointed to training at a pay grade (or General Schedule level) of GS-9 or higher—rather than the GS-7 level at which the majority of the less experienced are hired. In developing the exemption standards, officials presumed that men having substantial amounts of specialized experience would, in general, also possess aptitudes permitting rapid advancement in FAA training. During 1968-1970, less than one-fourth of the entrants into ATC training have established their eligibility on the basis of the specialized experience standard, and a recently completed study,⁵ involving a sample of over 300 of the highly experienced trainees, has shown that many of them are indeed of very high mental calibre. However, the same study has also shown that about 45 per cent would not have been selected had they been screened with the CSC ATC-Aptitude Test Battery.

With minor exceptions, the selection standards for other applicants have remained much the same as those under which all trainees were recruited during 1964 through 1968. They must take the group of CSC tests and about half are eliminated from candidacy because they fail to attain passing scores of 210 or higher. The scores of those who do pass are subsequently combined with evaluations of experience and other variables to establish overall eligibility ratings and selections are then made on a competitive basis. Since 1968, the FAA has selected over three-fourths of its ATC trainees from among such aptitude-screened candidates and, as alluded to above, previous research has indicated that about 55 per cent of the remaining one-fourth, who entered under the specialized-experience standard, would probably have been

able to qualify on the test battery had they not been granted a waiver of the requirement.

The probable implications of such recruiting practices upon research concerning the validity of the operational battery with current inputs are rather obvious—findings based on highly-screened personnel are apt to be grossly attenuated. Moreover, several unpublished CAMI studies involving Academy trainees have yielded correlations of .80 to .85 between scores on the battery of CSC tests and summary measures of performance on other groups and types of aptitude-assessment instruments. One such investigation, which preceded adoption of the battery for operational purposes, revealed a correlation of .81 between the summary CSC aptitude scores of several hundred trainees and their so-called "I.Q.'s" ("Intelligence Quotients") derived from performance measures on the California Test of Mental Maturity (CTMM). The aptitude scores for the group averaged 193 and the mean "I.Q." was 110.5. Although the study indicated that the CSC scores were more useful than the latter for prediction of training outcomes, the relationship between the two variables provided the basis for development of a regression equation which has subsequently permitted estimation of the general intelligence level of operationally-examined recruits relative to that of a general population. For example, it has been determined that a minimally qualifying score (of 210) on the CSC Battery approximates a CTMM I.Q. of 120 which, according to the CTMM Manual,¹⁴ corresponds to the 88th percentile for a "normal" (or general) population and about the 70th percentile for college freshmen.

While such findings indicate that it would probably be unwise to consider any upward revision of the CSC ATC Aptitude Test eligibility standard, they do not necessarily imply that any further improvement in the screening of personnel for ATC-related aptitudes is unattainable. As stated or implied by CAMI researchers in published reports^{3, 4} and in communications with FAA officials, it is doubtful that *all* mental abilities and skills bearing upon controller performance have yet been identified and it is questionable as to whether some of those which have been determined as pertinent are being optimally assessed.

C. Purported Loss of Screening Effectiveness. Based on the findings obtained in the validation

study with 893 Academy entrants during 1962 and 1963, it had been anticipated that adoption of the CSC Battery for screening purposes would result in a reduction of the attrition rate at the Academy from approximately 30 to about 16 per cent. A reduction did occur but not to the extent expected. For example, 22 per cent of the 2,822 men who presumably met or exceeded the aptitude screening requirement and who entered the Academy during November 1968 through March 1970 failed to successfully complete their training course. Also, the Air Traffic Controller Career Committee, in its 1970 report,¹ stated that a longitudinal study of all ATC personnel hired during 1 January 1967 through 30 November 1969 indicated "losses as high as 23 per cent" for some of the FAA regions. Much of the attrition, they pointed out, occurred at the Academy. The committee deemed the elimination rate rather high and concluded that deficiencies in the selection procedures, including aptitude screening, were responsible. Several analyses, all of which involved relatively small groups of trainees recruited since 1968, have produced (predictably) relatively low correlations between the global scores on the CSC Battery and summary measures of training performance. Specifically, the validity coefficients have ranged as low as .04 and no higher than .26 for prediction of the pass-fail training criterion and from .03 to .39 for training grade averages. These and other similar findings, however, should be expected on the basis of restriction of range and other factors noted below.

As mentioned earlier, a total of 27 different tests were involved in the study wherein the six CSC instruments were identified; a series of multiple correlational analyses¹⁰ selected these tests as the most valid group for prediction of training performance. A point which warrants emphasis, however, is that the 27 tests covered a variety of aptitude-factor areas. Twelve of the 27 had been selected for tryout on the basis of findings and recommendations reported by Brokaw² (for a study in which 198 ATC trainees were experimentally examined with a heterogeneous battery of 19 tests). In seeking to determine other types of tests most likely to validate, CAMI scientists familiarized themselves with the Academy's training courses and materials, engaged in lengthy consultations with manage-

ment and training officials, and also undertook a number of field trips to make on-the-job observations. Such efforts produced a reservoir of knowledge and information which proved helpful in, but never constituted the sole basis for, selection of tests for the experimental battery. In fact, the majority of the 27 instruments were chosen for study because it was felt they would provide reliable measures of specific mental abilities and aptitudes deemed relevant to those activities and functions identified in contractual research^{6 7 8 9 12 13} as most important in control work at ARTCC and high-density Tower facilities.

An aptitude screening battery predicated on such a vast and diverse amount of research would be expected to maintain considerable effectiveness as an initial screening device. Inasmuch as the Civil Service Commission adheres to rigid control-and-accountability procedures, there is little likelihood that the aptitude measuring devices have been subjected to compromise. Changes within the air traffic management system have admittedly occurred during the interim, but it is doubtful that the job requirements and task demands of control personnel are now markedly different from those of seven or eight years ago.

It was ascertained in the validation study with the 893 pre-1964 ATC trainees that Academy-basic-training attrition rates were: (1) 30.3 per cent for the total group, (2) only 16.4 per cent for the 317 who attained composite raw scores of 210 and higher on the CSC test battery, and (3) 38 per cent for the group of 576 having scores of 209 and lower. As mentioned earlier, 22 per cent of the 2,822 aptitude-screened men who entered the Academy during November 1968 through March 1970 failed to satisfactorily complete their training course. An attrition rate of 22 per cent is *significantly lower* than the 30.3 obtained for the total validation sample of 893 trainees but somewhat above the failure rate found for the 317 who had aptitude scores greater than 209.

If granted assurance that ATC training performance evaluation standards have remained unchanged since 1963, a failure rate of somewhat less than 22 per cent would be anticipated. While little or no evidence is available to conclusively prove or disprove equivalency of the performance criteria, the issue is one frequently con-

fronted in research aimed at progressive improvement of personnel selection procedures. It is commonly referred to as "the problem of creeping criteria" and, inasmuch as officials responsible for training traditionally strive to upgrade their programs, it is a problem that is normally and hopefully anticipated.

The authors of the present report hypothesize that the FAA has also followed much the same pattern: that ATC training programs and performance evaluation standards have been upgraded; that the changes (the exact nature and extent of which would be difficult to determine reliably) are partially attributable to training requirements associated with the increasing complexity of the air traffic management system and in part to the inherent tendency of training personnel to "customize" instructional material, instruction, and/or subjective performance evaluation standards relative to the general level and range of mental abilities characterizing the incoming classes. Moreover, the vast majority of trainees recruited since 1963 have been of exceptional mental calibre and, thereby, comprise groups which would be characterized by substantial restriction of range of aptitudes.

D. Rationale for Improving Aptitude Screening. There are a number of reasons why it may be assumed that some improvement in the current aptitude screening process might be achieved. First, the air traffic management system has undergone change in recent years, and it is possible that the job tasks now require certain mental skills supplementary to those currently being assessed.

Moreover, the research in which the six CSC tests were initially validated also yielded appreciable validities for a number of other CSC tests and several commercially-published copyrighted instruments. The latter group included several tests which apparently involved a speeded perceptual-discrimination factor and some which pertained to either coding skills, comprehension, rapid integration or processing of information, or memory. Although none of these tests appeared in the group which the multiple regression analyses indicated most useful for selection purposes, the results led us to suspect that the instruments were measuring important ATC-related aptitudes albeit inadequately. Such findings, supplemented by those obtained for a number of different spatial tests,

were interpreted as illustrating the need for development of tests "tailored" to fit the purpose.

While none of the various spatial tests included in CAMI's numerous experimental batteries has failed to validate at statistically significant levels, the validities of each have generally been much more variable from sample to sample than for non-spatial tests. This also applies to CSC Booklet 51 of the operational screening battery. We concluded several years ago that a "visual imagery" type of speeded spatial ability was a major determinant of ATC performance, but we were unable to locate an instrument fully appropriate in this respect. The development of a special test, the complexity of which would have required a major research effort, was not undertaken due to other research commitments having higher priorities.

II. Procedure.

The Directional Headings Test should be considered as a fortuitous outgrowth of developmental efforts which were undertaken on a somewhat intuitive basis rather than a consequence of extensive background research, job activity analyses, and the like. It was developed "in house," within a period of only four weeks by personnel of CAMI's Psychometrics Unit who interpreted the research findings discussed immediately above as indicating that the aptitude screening process might be further improved through more adequate assessment of the applicants' speeded perceptual-discrimination and coding (or decoding) skills. In conceiving the type of task ultimately embodied in Part I of the DHT, the initial exploratory efforts focused upon the use of diagrams simulating a navigational compass. However, the pictorial approach was abandoned on the hypothesis that if the data pertinent to the solution of each problem were presented in a different format, such as described below, the test would be more difficult and therefore provide better opportunity for individual differentiation.

The DHT is a 3-part test in which the subject is allowed a total of only 90 seconds for the solution of 60 items, or problems, of each part. In each problem, the subject is presented one, two, or three "bits" of information relating to the cardinal points on a mariner's compass. For example, the letter "N," the symbol "∧," and

the notation "360°" each denote "North." Similarly, the letter "E," the symbol ">" and "090°," either separately or in combination, denote "East." Other letters, symbols, and degrees correspond to "South" and "West." In Part I of the test, the examinee must rapidly interpret and collate the bits of presented information in order to ascertain whether the data are of a conflicting nature and, if not conflicting, then determine the directional heading to which they correspond; in Part II, he must determine the exact opposite of the heading; and, in Part III, he must ascertain the exact opposite of the directional heading while being subjected to aural distraction. In each item of Form A, the bits of information are followed by one of five questions: "North?", "East?", "South?", "West?", or "Conflict?". The subject is offered two response categories, "Yes" or "No." Approximately one-fourth of the items present "conflicting" data. For example, if the letter "S" were presented in an item of Part I with the symbol ">" and/or "270°" and if the question were "Conflict?", a correct answer would be indicated by marking the space under "Yes." On the other hand, the same conflicting data presented with "West?" (or "North?", "East?", or "South?") would warrant a "No" answer.

The first tentative version of the test consisted of only one part, with the task being essentially the same for all 60 items. Small groups of Academy ATC trainees were examined with the instrument and, although no validation analyses were possible at that time, considerable range and variance were found in the distribution of scores. Moreover, during test sessions, some individuals in every group appeared to become rather confused, or frustrated, when confronted with those items in which "conflicting" bits of information were presented and also hesitant in marking either "Yes" or "No" to many of those items where the questioned direction (e.g., "North?") failed to conform to the heading indicated by the data provided. When the CSC Battery was being validated some eight years earlier, reactions of much the same type had been noted for many of the trainees during their assessment with CSC 135 Following Oral Directions. Inasmuch as CSC 135 proved to be one of the most valid components of the screening battery, we hypothesized that many of the trainees who had performed poorly on the in-

strument might have done so because they had a "low frustration tolerance level." Although the latter phrase and the hypothesis to which it alludes are rather loosely formulated, they grossly reflect the reasoning which led to the development of Parts II and III of the DHT.

As mentioned earlier, Part II is similar to Part I except that in each item the subject is asked whether a specified direction represents the direct opposite of the heading. Part III is like Part II but the examinee is subjected to possible distraction by the aural presentation of the randomly arranged words, "West," "South," "East," "North," and "Conflict." The resulting 3-part instrument, in which the examinees were offered no response categories other than "Yes" or "No," was designated as "CAMI Directional Headings Test—Form A."

An alternate version of the test, Form B, was eventually developed; it offered five response categories. In each item of the latter, the "bits" of information are presented in the same manner as in Form A but the "question element" does not appear and the examinee merely indicates the answer he deems correct by marking a response space under one of the five columns labeled, "East," "South," "West," "North," and "Conflict."

Form A of the DHT was administered on an experimental basis to a total of 586 students as they enrolled at the FAA Academy during 20 January 1970 through 13 March 1970 in either the basic ARTCC or TATC training course. Form B was administered on a similar basis to 199 trainees, all of whom entered the Air Route course on 27 March 1970. The vast majority of the 785 trainees examined with either Form A or Form B were new hires and none possessed an FAA ATCS certificate. For purposes of analysis, however, the cases were divided to establish the samples and subsamples described below.

A. Sample 1. Form A was administered to 388 ARTCC, or En Route, students who were designated as Sample 1. Of the 388, one entered training at the GS-5 level and 361 at the GS-7 level; the former case was added to the 361 and the group was thereafter referred to as subgroup "1A—En Route—GS-7." The remaining 26 En Route examinees entered at the GS-9 level. They constituted the subgroup "1B—En Route—GS-9."

B. *Sample 2.* The 198 students assessed with Form A of the DHT as they entered the TATC basic training course represented Sample 2. One was a GS-5, 141 were GS-7's, 55 were GS-9's, and one was GS-11. The GS-5 was merged with the GS-7's to establish the subgroup "2A-Terminal-GS-7." The remaining cases comprised the subgroup "2B-Terminal-GS-9."

C. *Sample 3.* Only 199 trainees were examined with Form B of the DHT. They were designated as Sample 3. The subgroups were "3A-En Route-GS-7" and "3B-En Route-GS-9," with N's of 187 and 12, respectively.

The question had arisen during the development of the DHT as to whether past experience in control work would serve as a major determinant of test performance. Unfortunately, official records of the experiential backgrounds were unavailable for many of the trainees who took the test. However, there were reasons to believe that, by maintaining separateness of the test data for the GS-7's and GS-9's, certain phases of the analyses might yield some information bearing upon the issue. It was presumed that most subjects who entered with a pay grade above GS-7 received their appointment to training on the basis of exceptional pre-FAA experience (usually in military air traffic control work) and that, on the same basis, they were automatically exempted from the aptitude screening requirement. Knowledge of FAA recruiting practices suggested that the vast majority of the GS-7 DHT examinees were probably men with moderate-to-appreciable amounts of pre-FAA ATC-related experience who, in establishing their eligibility for selection, were required to pass the CSC ATC Aptitude Screening Test Battery. It was recognized that the experience backgrounds of the GS-7's and GS-9's could not be considered representative of applicant groups. However, it was reasoned that *if* the analysis revealed *no* significant differences between the mean DHT scores of the GS-9's and GS-7's, the issue regarding possible differential effects of lesser amounts of experience (than possessed by most GS-7's) would warrant little concern because the FAA traditionally recruits a very small proportion of its trainees from among those applicants having relatively little or no previous ATC-related experience. On the other hand, if the GS-9's performed at a significantly

higher level than the GS-7's, it would then be unwise to recommend that the DHT be used as an operational screening device.

D. *DHT Scoring Procedures.* A formula of "right responses minus wrong" was employed in the scoring of Form A of the DHT because only two response alternatives (i.e., "Yes" and "No") were offered for each item. Form B, with a 5-choice response format, was scored in accordance with the formula "rights minus one-fourth the wrongs." Scores involving fractions were always rounded to the nearest whole number.

Inasmuch as differences, as well as similarities, characterized the tasks presented in Parts I, II, and III, it was postulated that the parts might prove unequal in terms of difficulty and it was also felt that the parts might yield contrastingly different validities. Consequently, the three parts were scored separately. An *arithmetic mean* of each subject's part scores, rather than a sum of the part scores, was obtained to serve as the global measure. The mean score is referred to in the present report as either "The Total Average Score" or simply as "The DHT Score." The latter, as well as the part scores, were analyzed to determine their respective relationships with the criteria. In addition, a number of similar analyses were undertaken regarding the differences between the summary DHT measure (i.e., "Average Score") and each of the part scores of the subjects and differences between the part scores themselves.

E. *Criterion Variables.* Academy training officials provided CAMI with an "Evaluation of Performance Record" for each of the trainees who participated in the experimental testing program. The evaluation form for the ARTCC trainees was not identical to that employed in evaluating TATC training-course performance. Nevertheless, each provided grades which permitted derivation of a similar, corresponding measure of performance for each subject, regardless of training option. The summary measure, referred to as the "Combined Academic and Laboratory Grade Average" or "A+L," represented an arithmetical mean of two separate averages; one based on all examination grades relating to academic materials, instruction, and the like, and the other based on performance grades for simulated air-traffic-control work in the laboratory.

The second criterion variable was "Pass-Fail Status" or "P-F." Students whose records bore the notation "Withdrawn in Failing Status" were designated as "Fails"; those who were withdrawn for other reasons were deleted from the study, and all others were considered as "Passes."

In accomplishing some of the analyses, every subject who passed training was categorized either as a "Marginal Pass" or "Non-Marginal Pass." The "Marginal Pass" category was established separately for each sample, irrespective of GS level. First, the "Fail" cases were extracted; next, the A+L Grades of all non-fail subjects in a given sample were arrayed from low to high, and those comprising approximately the lower one-fourth of the grade distribution were designated as "Marginal Passes."

F. CSC Test Scores. It was postulated that if the DHT scores of the pre-screened groups of trainees should validate at statistically significant levels for prediction of the training performance measures just described, there would be little doubt that the instrument would validate at higher levels with applicant groups. Nevertheless, additional information bearing upon the interpretation of the attenuated validities could be obtained if the CSC Test Scores and/or evaluations of pre-FAA ATC-related experience were included in the scheduled analyses. Such data were available for a portion of the trainees comprising each validation sample but, as will be discussed, there were reasons to suspect bias in the data.

The development and validation of the DHT was undertaken in conjunction with, and as part of, a comprehensive project which had been initiated in response to a request by the FAA Director of Personnel (PN-1) for updated research on factors associated with successful performance in ATC training. (Other aspects and findings of the latter will be discussed in subsequent reports.) In connection with the requested research, PN-1 wrote to all regions specifying that CAMI be provided a report on each trainee at time of entry into the Academy; the report was to reflect the trainee's overall eligibility rating and other types of information. Most importantly, if the trainee had been assessed with the operational CSC ATC Aptitude Test Battery, the report was to include all subscores

(i.e., scores on each CSC test booklet) as well as the complete aptitude measure.

Records were ultimately forwarded to CAMI for about nine-tenths of the 3,579 trainees involved in the larger project but many of them were incomplete. Researchers received CSC Test Scores for less than 52 per cent of the entire group although it was estimated that 70 to 75 per cent had probably been operationally examined with the battery. Unfortunately, insofar as research purposes were concerned, scores on the aptitude screening battery were forwarded for a much smaller proportion of the training failures than for Academy graduates.

In the present study, the 504 GS-7 trainees of Samples 1 and 2 who were administered Form A of the DHT were also administered a questionnaire which included the following two items: (1) "Have you ever been examined with the CSC ATC Aptitude Screening Test Battery?" (2) "If so, please indicate the number of times you took it, and for each time, also indicate the approximate date and the city in which examined." Exactly 95 per cent of the total group of 504 GS-7's indicated that they had been tested one or more times. (Slightly over 97 per cent of the 102 training-course failures within the group claimed they had taken the battery on an operational basis.) However, performance scores on the battery were forwarded to CAMI for only 281 of the 504. Moreover, the 281 represented 62.2 per cent of the 402 pass cases but only 30.4 per cent of the 102 failures. Inasmuch as the CSC Test Scores were available for disproportionately fewer training-course failures than for pass cases, it was presumed that the analyses to be undertaken in this study would yield grossly attenuated validity coefficients for the operationally-derived aptitude measures and that the correlations involving the latter and age and other variables would also be affected.

G. Pre-FAA ATC-Related Experience. The questionnaire which CAMI administered to the DHT examinees also contained a section which pertained to ATC-related experience. Specifically, the instructions requested that each trainee indicate whether he had ever held a license, certificate, or rating as a "pilot" and/or in "air traffic control work" and/or in the field of "communications." Although 170 trainees

replied "No" to each of the three areas, the remainder checked one or more. These responses provided the basis for assigning each case to one of eight mutually exclusive categories, with one of the categories reflecting certificated experience in all three fields. An analysis was then scheduled to assess possible effects of experience on DHT scores.

H. *Limited Research on Form B of the DHT.* As pointed out earlier, Form A of the DHT was administered to a total of 586 trainees (Samples 1 and 2) whereas only 199 trainees (Sample 3) were examined with Form B. Academy training schedules were rearranged soon after Form B was developed, thereby precluding further experimental test sessions. Several different analyses accomplished on the data of Form A examinees were not undertaken with Sample 3. Inasmuch as it was determined very early in the study that the two forms were highly comparable in terms of both validity and reliability, it was assumed that if other corresponding analyses of data for the Form A and Form B groups were undertaken, they also would yield comparable findings. Research efforts were therefore focused primarily upon Form A.

III. Results and Discussion.

After scoring the DHT answer sheets, each examinee's scores on Parts I, II, and III were averaged to obtain the overall measure of performance (i.e., "The DHT Score"). The first analysis, which pertained to the Form A examinees only, involved a comparison of the frequency distributions of the DHT Scores of the graduates and non-graduates of the ARTCC and TATC training courses. The distributions are shown in Table I. As may be noted, the graduates of each training course tended to attain higher scores than the non-graduates. The 329 ARTCC graduates averaged 38.5 on the three parts of the test whereas their 59 attrited classmates averaged 31.0. Similarly, the mean score for the 142 TATC graduates was 39.0 while the 56 who failed averaged 31.0. Both differences were statistically significant at the .01 level. Moreover, three of the En Route graduates attained perfect scores of 60 whereas the highest score of any attrited ARTCC trainee was 55; one subject within each subgroup scored only 01. Three of the TATC graduates scored 56 on the

test and none scored lower than 13, whereas the scores for the Terminal failures ranged from 50 down to 05. Other differences between the score distributions of the passes and fails, and their bearing upon a recommended screening standard, will be discussed in one of the later sections of this report.

When the frequencies of the Form-A DHT Scores of the passes and fails were merged, the resulting distributions (which are not shown) proved to be much alike. Although that of Sample 1 reflected somewhat greater range and variability of performance, both were slightly skewed in a negative direction. Moreover, the difference between the means of 37.4 and 36.7 (for Samples 1 and 2, respectively) was not statistically significant. Inasmuch as 60 represented the maximum possible (average) score on the three parts, the range and distribution of performance measures for each group were considered quite satisfactory in that they provided no basis for questioning the difficulty level and/or timing of the test.

A. *Intercorrelations and Empirical Validities of DHT and CSC Test Scores.* A point which warrants emphasis is that the three parts of the DHT present different tasks or task situations. Considering these differences, the authors expected that the intercorrelations of the part scores would be of relatively moderate magnitude. The Pearson product-moment correlation coefficients,¹¹ reflecting the interrelationships, are shown, along with other data, in Table II. The correlations (or "r's") ranged from .46 to .77 for Part I versus Part II, .29 to .69 for I versus III, and from .67 to .88 between the scores on Parts II and III. In each instance, however, the lowest of the coefficients pertained to Subsample 2A. Otherwise, the coefficients ranged from .64 to .77 for I versus II, .58 to .69 for I versus III, and from .81 to .88 for II versus III. The narrow range of the latter would imply that the test possessed a fair degree of reliability. The authors re-examined all data relating to Subsample 2A but found no computational errors or reasons why the intercorrelations were lower than obtained for any other group.

The hypothesis that examinees would generally experience greater difficulty with Part II than Part I proved rather tenuous. For every sample and subsample, the mean score on Part

TABLE I.—Distribution of Form-A Directional Headings Test scores for 388 ARTCC and 198 TATC Trainees who were examined upon entry into FAA Academy training.

Form-A D.H. Test Score	Sample 1 Enroute or ARTCC Trainees						Sample 2 Terminal or TATC Trainees					
	Graduates			Non-Graduates			Graduates			Non-Graduates		
	f	Cum. f	Cum. %	f	Cum. f	Cum. %	f	Cum. f	Cum. %	f	Cum. f	Cum. %
56 & >	13	329	100.0		59	100.0	3	142	100.0		56	100.0
54 - 55	1	316	96.0	3	59	100.0	4	139	97.9		56	100.0
52 - 53	10	315	95.7	3	56	94.9	5	135	95.1		56	100.0
50 - 51	17	305	92.7		53	89.8	7	130	91.5	1	56	100.0
48 - 49	17	288	87.5	3	53	89.8	6	123	86.6		55	98.2
46 - 47	19	271	82.4	2	50	84.7	12	117	82.4	2	55	98.2
44 - 45	16	252	76.6	1	48	81.3	5	105	73.9	1	53	94.6
42 - 43	33	236	71.7	4	47	79.7	18	100	70.4	4	52	92.9
40 - 41	32	203	61.8	1	43	72.9	11	82	57.7	6	48	85.7
38 - 39	38	171	52.0	3	42	71.2	12	71	50.0	3	42	75.0
36 - 37	27	133	40.4	2	39	66.1	12	59	41.5	4	39	69.6
34 - 35	27	106	32.2	4	37	62.7	12	47	33.1	2	35	62.5
32 - 33	16	79	24.0	4	33	55.9	6	35	24.7	6	33	58.9
30 - 31	14	63	19.1	3	29	49.2	9	29	20.4	2	27	48.2
28 - 29	10	49	14.9	3	26	44.1	2	20	14.1	4	25	44.6
26 - 27	9	39	11.9		23	39.0	8	18	12.7	3	21	37.5
24 - 25	4	30	9.1	5	23	39.0	3	10	7.0	7	18	32.1
22 - 23	7	26	7.9	5	18	30.5	2	7	4.9	3	11	19.6
20 - 21	2	19	5.8	2	13	22.0		5	3.5	1	8	14.3
18 - 19	2	17	5.2	3	11	18.6	1	5	3.5	3	7	12.5
16 - 17	7	15	4.6	1	8	13.6	2	4	2.8	1	4	7.1
14 - 15		8	2.4		7	11.9	1	2	1.4	1	3	5.4
13 & <	8	8	2.4	7	7	11.9	1	1	.7	2	2	3.6
Mean D.H. Score*												
Graduates:		38.5						39.0				
Non-Grads:					31.0						31.0	
Std. Deviation												
Graduates:		10.2						9.2				
Non-Grads:					14.3						9.9	

* Each mean and S.D. is based on ungrouped test performance data

TABLE II.--Intercorrelations and validities of Directional Headings Test measures and CSC-ATC Test scores.

Sample and Type	Variable	N	Mean	S.D.	Part II	Part III	Total Avg. Score	CSC*	Training Criteria	
								Batt. Score	A+L	r pbi** P-F
1A, Enroute GS-7's	DH-A, Part I	362	37.4	9.5	.64	.58	.78	.24	.28	.14
	DH-A, Part II	362	36.6	12.4		.82	.93	.23	.31	.21
	DH-A, Part III	362	39.1	13.9			.93	.30	.36	.25
	DH-A, Avg. Part Score	362	37.5	11.2				.29	.38	.24
	CSC-ATC Test Battery	212	248.2	20.7					.13	.04
	Acad.+Lab Trng. Grade	362	88.1	5.3						.63
1B, Enroute GS-9's	DH-A, Part I	26	35.8	10.7	.77	.59	.80		.64	
	DH-A, Part II	26	34.0	12.4		.84	.95		.46	
	DH-A, Part III	26	37.1	15.8			.93		.35	
	DH-A, Avg. Part Score	26	35.2	12.4					.50	
	CSC-ATC Test Battery*	9	245.8	34.7						
	Acad.+Lab Trng. Grade	26	90.8	4.4						
2A, Terminal GS-7's	DH-A, Part I	142	37.1	8.5	.46	.29	.64	.12	.32	.28
	DH-A, Part II	142	33.9	13.4		.67	.91	.16	.45	.31
	DH-A, Part III	142	37.3	14.6			.85	.11	.45	.28
	DH-A, Avg. Part Score	142	36.2	10.0				.15	.50	.35
	CSC-ATC Test Battery	69	238.1	21.4					.14	-.04
	Acad.+Lab Trng. Grade	142	79.1	8.2						.73
2B, Terminal GS-9's	DH-A, Part I	56	38.0	9.0	.68	.63	.81		.33	.36
	DH-A, Part II	56	36.1	11.6		.88	.96		.33	.28
	DH-A, Part III	56	40.9	11.9			.93		.39	.32
	DH-A, Avg. Part Score	56	38.2	9.9					.38	.34
	CSC-ATC Test Battery	4	231.8							
	Acad.+Lab Trng. Grade	56	80.4	8.9						.86

*Correlation coefficients were not obtained because CSC Test scores were available for only 9 cases of Sample 1B and 4 cases of Sample 2B

**Point-biserial coefficients were not computed for Sample 2 because only 2 of the 26 cases were classified as training-course failures.

II was lower than that on Part I, but only one such difference—involving Subsample 2A—was statistically significant. (The performance means pertaining to Form A appear in Table II and those relating to Form B appear in Figure 1.) The authors also hypothesized that a sizable proportion of the trainees would be adversely affected by the “conflicting noise” imposed during administration of Part III. Contrary to expectation, the Part III mean score exceeded that of Part II for every sample and subsample and all mean differences except those relating to Sample 3 and Subsample 1B were statistically significant. Moreover, a review and comparison of each individual’s part scores revealed relatively few instances of performance decrements from Part II to Part III. However, no definite conclusions can be made regarding either of the above hypotheses because the results are presumably confounded by learning effects.

Scores on the separate and combined parts of each version of the DHT validated at substantially higher levels for prediction of the A+L Grades than for pass-fail status (see Table II and the tabular data presented with Figure 1). For the four subgroups of Form A examinees, the correlations between the composite DHT scores and the A+L Grades ranged from .38 to .50 while those relating to pass-fail status ranged from .24 to .35. The overall scores of the 187 Form B examinees correlated .44 with the training-course grade averages and only .15 with the alternate criterion.

All such validity coefficients pertaining to the overall scores on either form of the DHT are statistically significant. Each is also significantly higher than the corresponding coefficient obtained for the CSC ATC-Aptitude-Screening Test score. However, the validity indices presented for both the operational screening battery and the two versions of the DHT should be regarded as grossly attenuated because they are based on data pertaining to individuals who had been selected from among those who had fully met the FAA’s existing ATC-personnel-screening standards. Had formulae been applied to estimate the correction for restriction-of-range effects, the resulting coefficients would have been larger, and in some instances of much greater magnitude, than those actually obtained. For example, if the correlation of .13 (see Table II) between the CSC ATC Test Score and A+L

Grade were adjusted (relative to the ratio of variances in the performance measures of the subjects of the current study and the 893 ATCS trainees who participated in the original validation of the CSC Battery), a coefficient of .25 would result. Such estimated values are not shown in the present report. Instead, the authors have followed a more conservative approach and have presented only the un-corrected validity coefficients.

B. Correlations Between Scores on Odd- and Even-numbered Items. Correlations of scores based on odd-numbered items versus scores based on even-numbered items were computed for Form A for the 362 GS-7 En Route trainee subjects only. They were: .92 for Part I; .94 for Part II; .93 for Part III, and .93 for the scores representing an average of the three parts. Similar procedures were employed with the response data of the 187 GS-7 En Route Form B examinees. The resulting correlations were .93, .96, .93, and .95 for Parts I, II, III, and the combined parts, respectively. Such coefficients, however, should be regarded as inflated estimates of reliability because the DHT is admittedly a highly-speeded test and it is known that most speeded tests are apt to yield relatively high correlations between measures reflecting performance on odd- and even-numbered items, particularly when the items, like those of the DHT, do not represent a wide range of difficulty.

C. Effectiveness of DHT Scores in Predicting “Fail,” “Marginal Pass,” and “Non-Marginal Pass” Cases. The effectiveness with which the scores on Form A of the DHT might be used to forecast three categories of performance in the ARTCC training course is illustrated in Figure 2. Prior to undertaking the series of analyses which produced the results depicted in Figures 2, 3, and 4 (which all follow the same format), we inspected the frequency distributions presented in Figure 1 and decided that a DHT Score (i.e., an average of the part scores) of 29 represented the best choice for a “cutting score” or theoretical screening cut. Analyses were then accomplished using score intervals of “29 <,” “30-39,” and “40 >.”

Looking first at the upper set of bar graphs in Figure 2, it may be noted that 69 of the 362 En Route GS-7 trainees scored 29 or less on the DHT and that 25 (36.2 per cent) of the 69

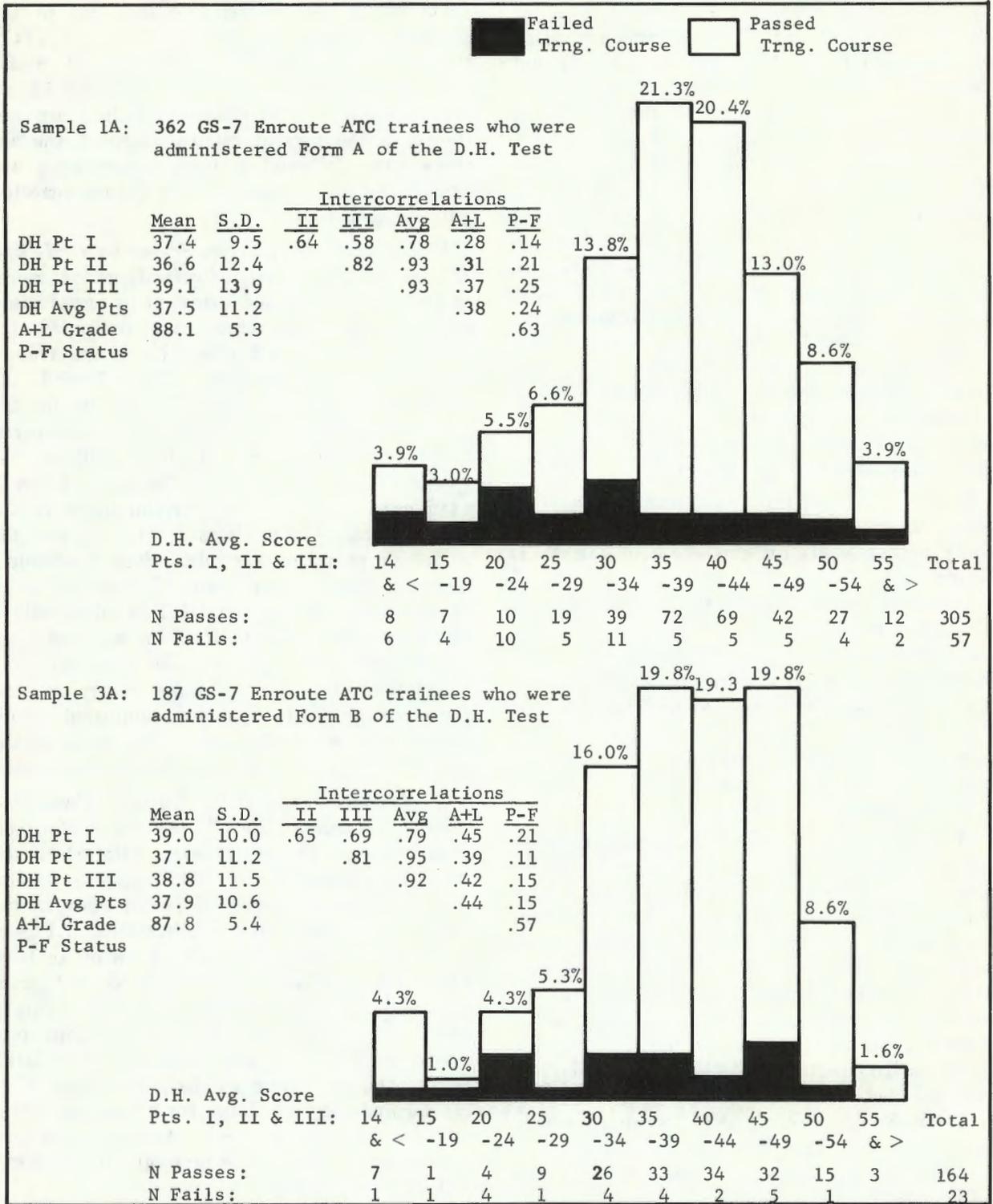


FIGURE 1. Comparison of performance data for groups of En Route GS-7 level ATCS trainees who were administered the Directional Headings Test Form A or B.

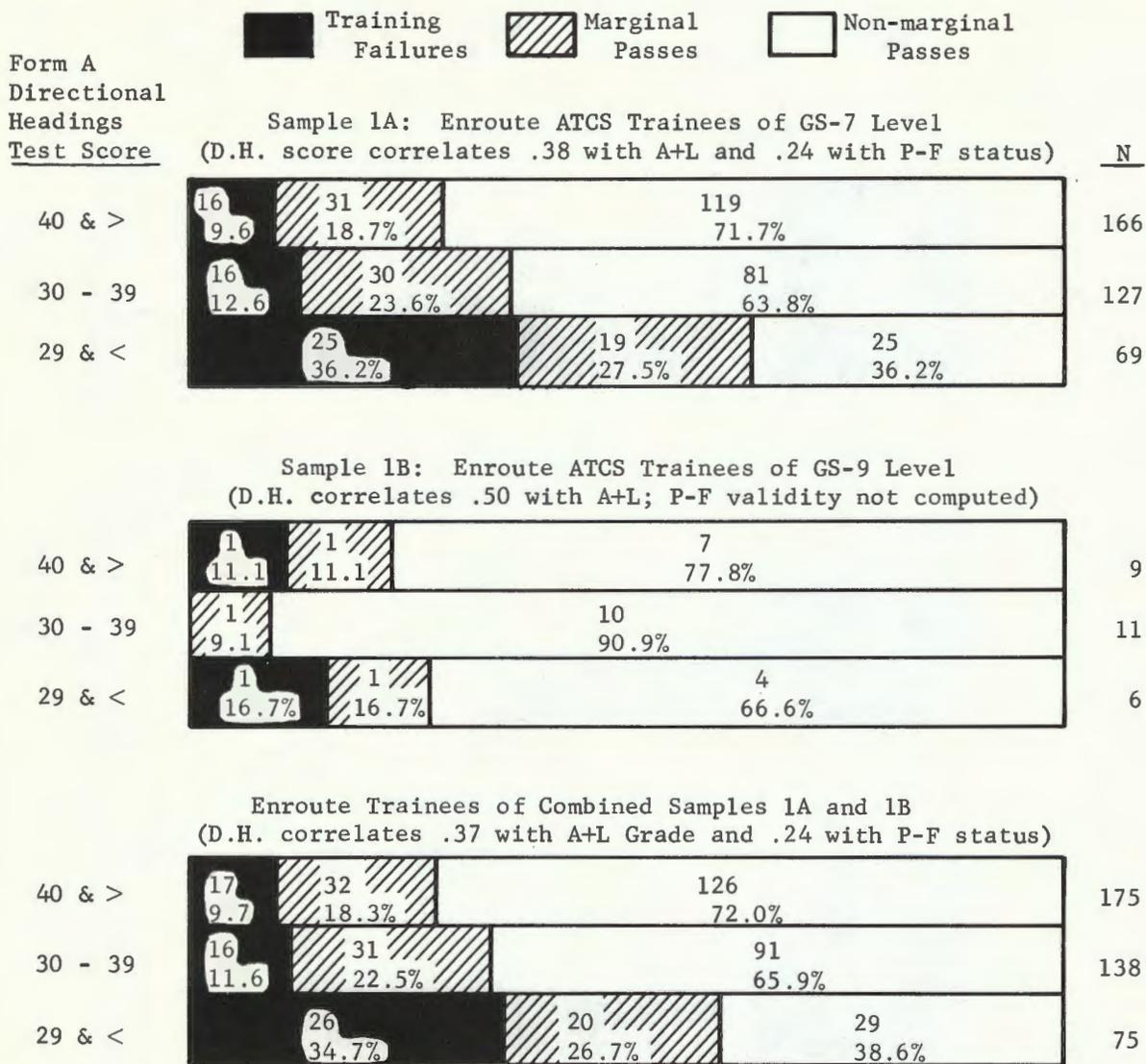


FIGURE 2. Proportion of En Route ATCS trainees within each of three Directional-Headings-Test-Score groupings who failed the Academy's basic ARTCC training course, "marginally passed," or passed with grades comprising the approximate upper three-fourths of the A+L Grade distribution for the passes only.

failed to complete training; 19 (27.5 per cent) marginally passed with A+L grades represented in the approximate lower one-fourth of the grade distribution of the passes only, and that only 25 (36.2 per cent) completed their course as "non-marginal passes." In contrast, 63.8 per cent of the 127 with DHT scores of 30 to 39 were "non-marginal passes." The results for the 166 GS-7 ARTCC trainees who scored 40 and higher on the experimental aptitude test are impressive but only slightly more so than those pertaining to trainees who scored 30 to 39. No conclusions can be derived on the basis of the results ob-

tained for the 26 ARTCC trainees of GS-9 level. Only two of the 26 failed, three were "marginals," and the other 21 were "non-marginal passes." It is possible that the pre-FAA ATC experience of these trainees enabled them to master more readily the training materials than their less-experienced and lower GS-rated classmates, but such a hypothesis should be viewed as rather tenuous because the sample of En Route GS-9's was so small.

The results of corresponding analyses on the data of the GS-7 and GS-9 Terminal students were much more impressive than obtained for

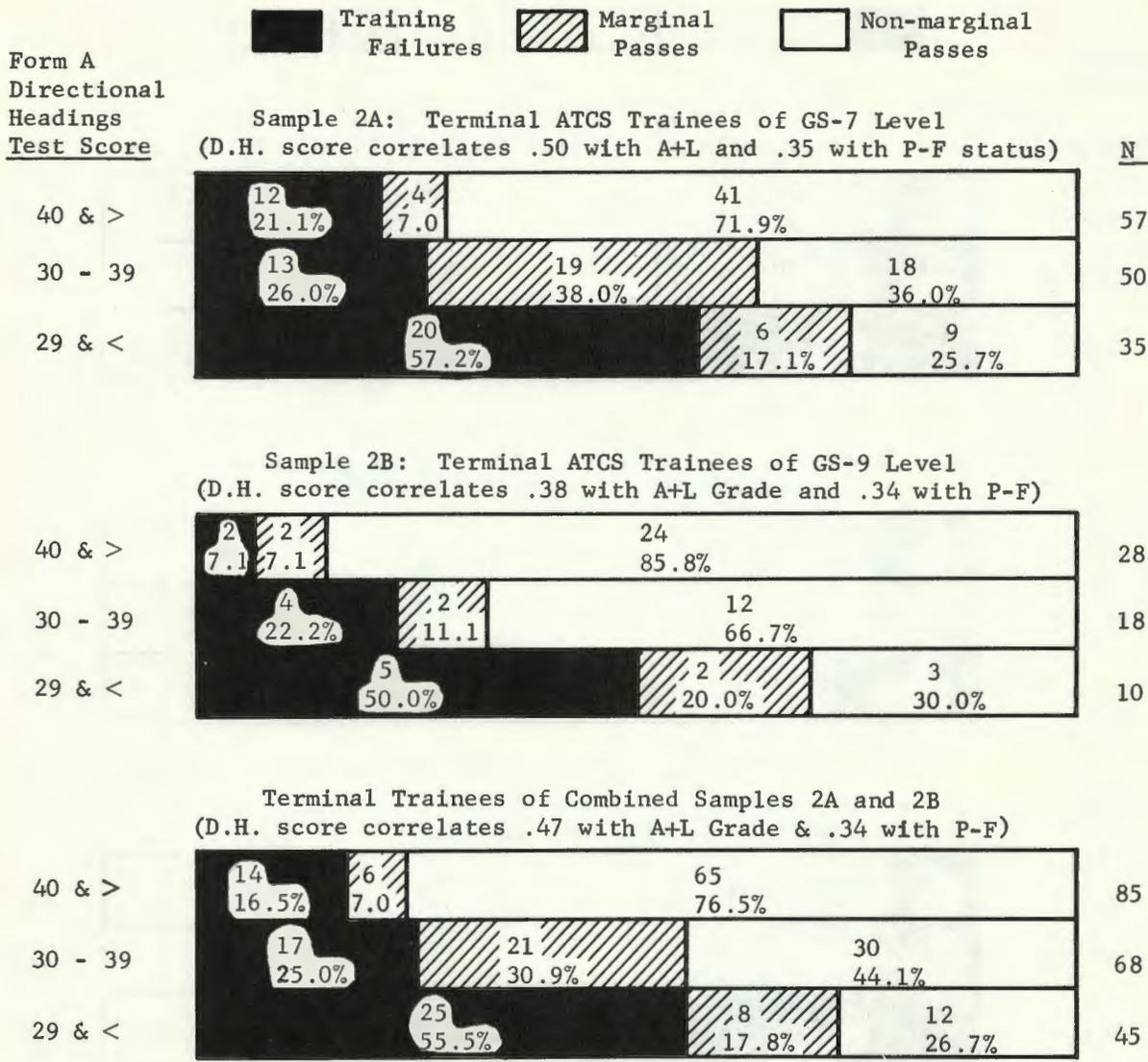


FIGURE 3. Proportion of Terminal ATCS trainees within each of three Directional-Headings-Test-Score groupings who failed the Academy's basic Terminal training course, "marginally passed," or passed with grades comprising the approximate upper three-fourths of the A + L Grade distribution of the passes only.

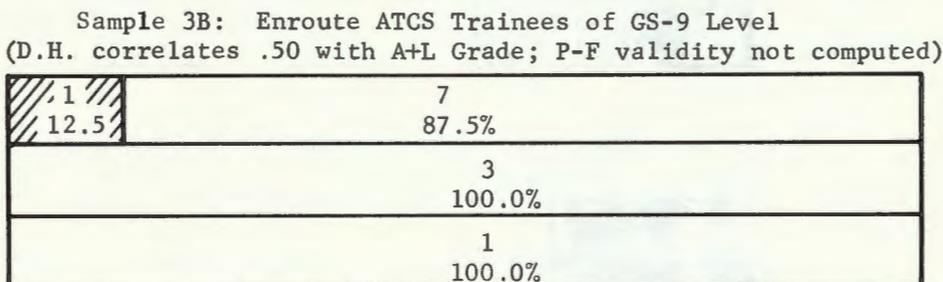
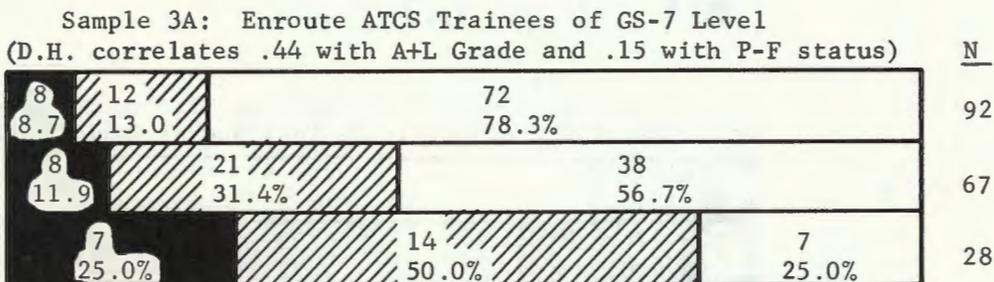
the groups of En Route trainees. In the upper set of bar graphs shown in Figure 3 for the 142 GS-7 TATC trainees, it should be noted that 57.2 per cent (N=20) of the 35 men having DHT Form-A scores of 29 and less failed training. Of these 35, an additional 17.1 per cent "marginally passed" and only 9 (or 25.7 per cent) passed as "non-marginals." Among the 50 having test scores of 30 to 39, a total of 31 (or 64 per cent) either failed or "marginally passed." In contrast, only 28.1 per cent of the 57 with scores of 40 and above failed or "marginally passed." The findings depicted for Sample

1B (Figure 3) illustrate that the DHT Form-A scores also possess potential for discriminating between the training-performance criteria of the Terminal GS-9's. Seventeen of the 56 TATC GS-9's either failed or passed with "marginal" grades and seven of the 17 scored 29 or less on the DHT, six scored between 30 and 39, and only four attained scores of 40 or above.

Relatively few of the ARTCC trainees who were examined with DHT Form B failed training and consequently the results shown in Figure 4 are not nearly as dramatic as those previously discussed for Form A. Yet, among the 187 GS-7

Training Failures
 Marginal Passes
 Non-marginal Passes

Form B
Directional
Headings
Test Score



Combined Samples 3A and 3B: Total of D.H. Test Form Examinees
(D.H. score correlates .45 with A+L Grade and .15 with P-F)

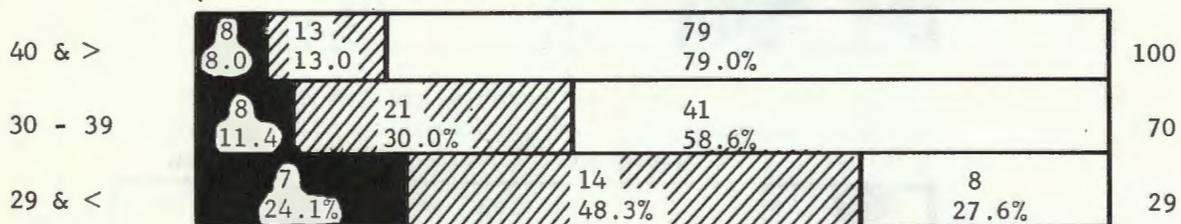


FIGURE 4. Proportion of Directional-Headings-Test Form B examinees (En Route only) within each of three D.H. performance score groupings who failed Academy basic training, "marginally passed," or passed with grades represented in the approximate upper three-fourths of the A+L Grade distribution of passing Ss only.

ARTCC trainees, the 28 with DHT scores of 29 and less included only 7 "non-marginal passes" and 21 who either failed or performed "marginally." For each of the higher DHT score intervals, a progressively higher proportion of the GS-7 subjects were "non-marginals."

D. *Screening Potential of the DHT Versus the CSC Test Battery.* As mentioned earlier, performance measures on the operational CSC ATC Aptitude Test Battery were received for 281, or only 55.8 per cent, of the 504 GS-7 ATCS trainees who were experimentally assessed with

DHT Form A. Moreover, the 281 cases included only 31 training-course failures. Inasmuch as the 31 attritions represented only 11 per cent of the sample of 281, whereas 102, or over 20 per cent, of the 504 were known to have failed, the sample should be regarded as biased. The reasons why bias occurred in the collection of the data are unknown. (Over 97 per cent of the 102 non-graduates, like 95 per cent of the 402 graduates, checked "Yes" to an item of the CAMI Questionnaire indicating that they had taken the CSC Battery in the process of establishing eligibility for appointment to training.)

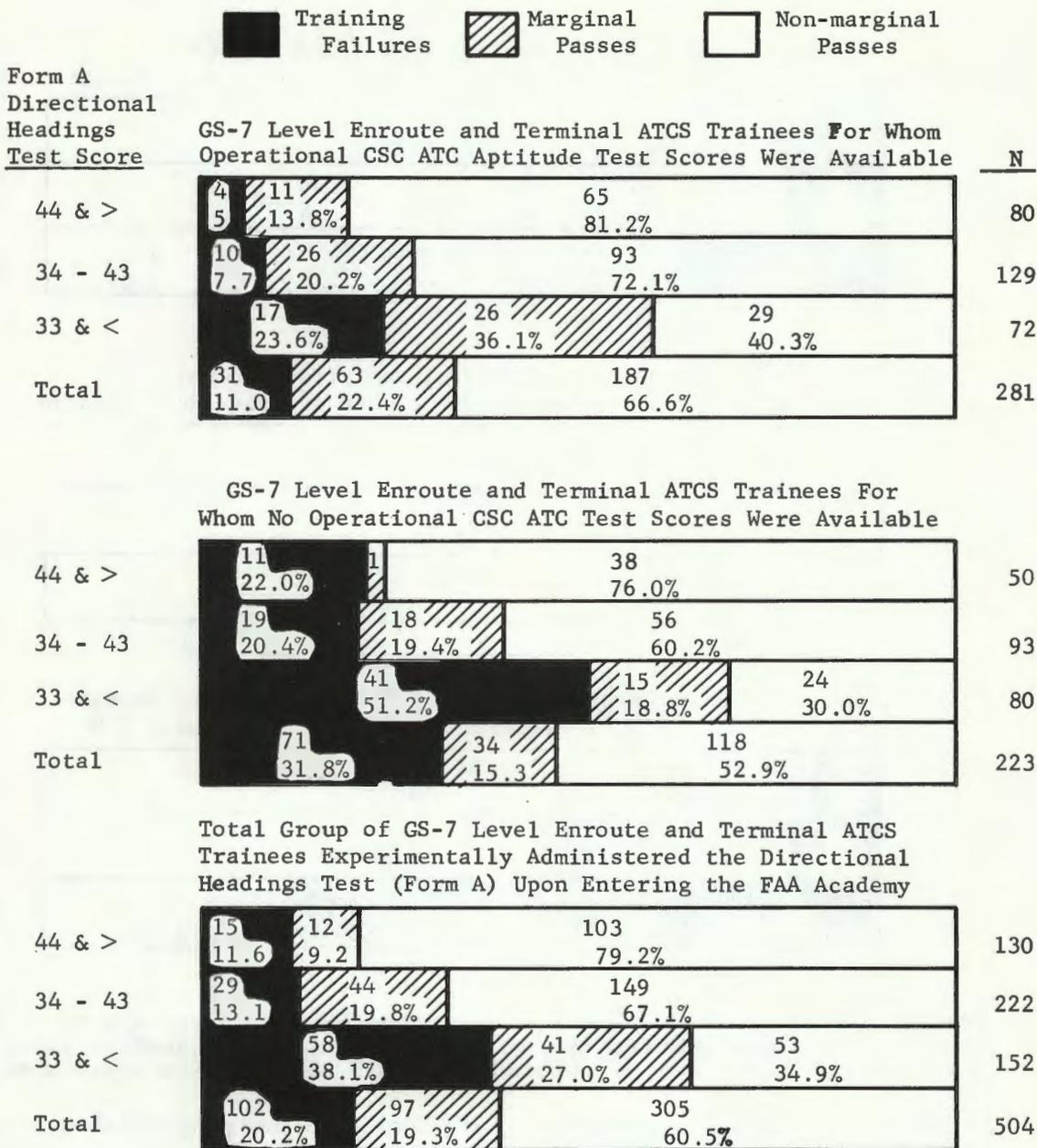


FIGURE 5. Distribution of training-course failures, "marginal" and "non-marginal" passes within each of three Directional-Headings-Test-Score groupings for 281 GS-7 level ATCS trainees (of Samples 1A and 2A) known to have been operationally examined with the CSC-ATC-Screening Battery and for a similar group of 223 trainees for whom no CSC-Test scores were forwarded.

Notwithstanding the shortcomings of the sample, the authors had no recourse but to deal with the 281 cases in analyses aimed at comparing the screening potential of the DHT with that of the CSC Battery.

The DHT Form A Scores of the 281 GS-7 ATCS trainees for whom CSC Test data were available were arrayed from low to high and the resulting distribution was consequently divided into approximate fourths. The lower fourth

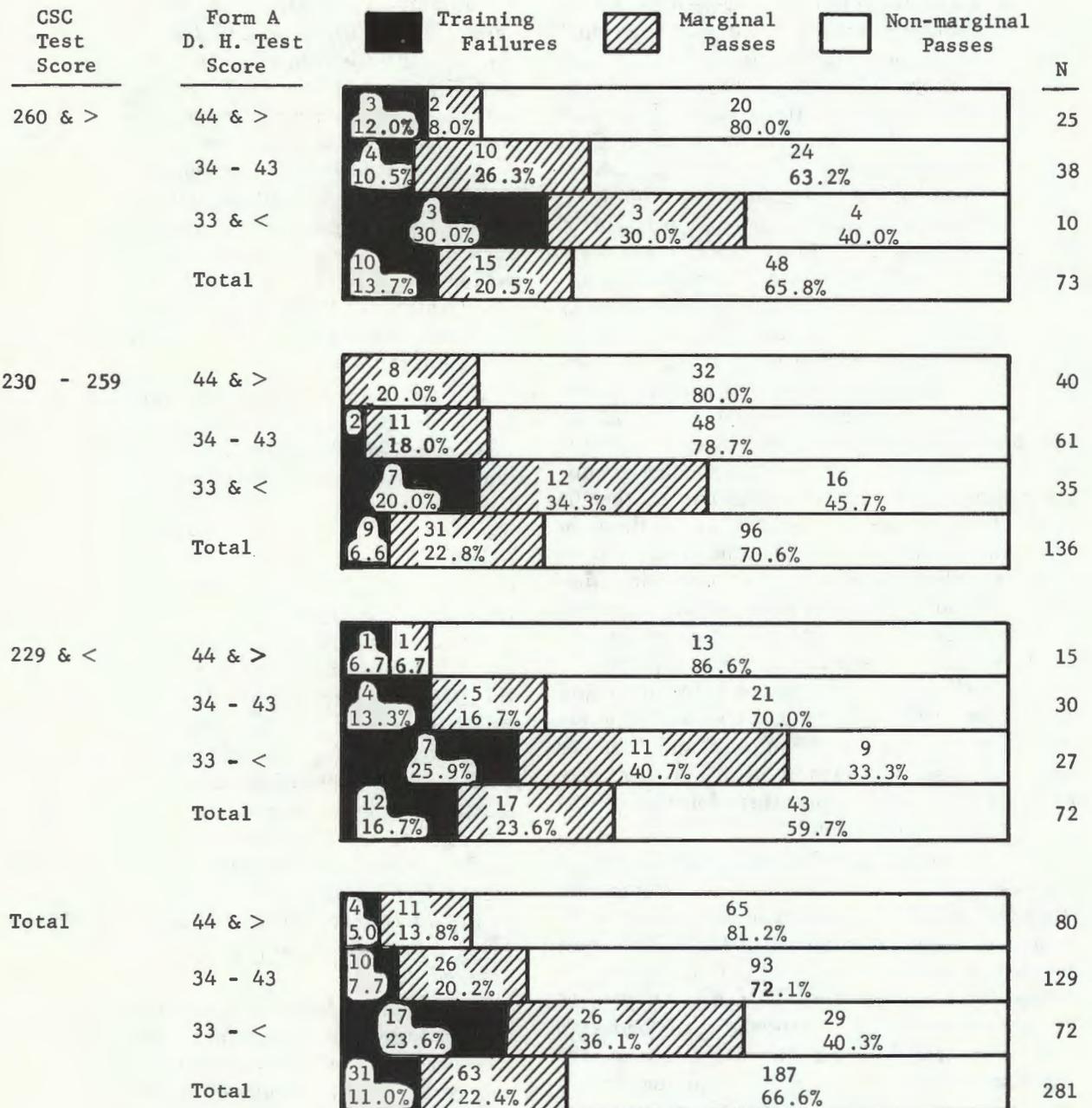


FIGURE 6. Percentages of ATC trainees by CSC-Test-Score and Directional-Headings-Test-Score categories who failed Academy basic training, passed with grades comprising the approximate lower one-fourth of the distribution of "A+L" training grades of the passes only (i.e., "marginals"), or passed with grades in the upper three-fourths (i.e., "non-marginals"). The data pertain to 212 ARTCC trainees of Sample 1A and 69 TATC trainees of Sample 2A, all of GS-7 level and for whom operational CSC-Test-Scores were forwarded to CAMI.

comprised 72 cases, all with scores of 33 or less; 80 cases with scores of 44 and above represented the upper fourth; the remaining 129 cases, with scores of 34 to 43, were designated as the intermediate half. Each of the three subgroups were then reviewed to determine the proportion of cases representing "training failures," "marginal passes," and "non-marginal passes." Retaining the DHT Score intervals established with the group of 281, a similar analysis was accomplished with the group of 223 subjects for whom no CSC Test data were available. In comparing the results obtained for the two groups (see Figure 5), it should first be noted that only 11 per cent of the 281 subjects for whom CSC scores were available failed to successfully complete Academy ATC training, whereas an attrition rate of 31.8 per cent occurred for the 223 for whom no operational aptitude screening measures were available. The latter group also performed less well than the former on the DHT. Although not shown, the mean DHT score was 38.2 for the group of 281 and 35.8 for the smaller group; the mean difference was significant at the .05 level. Over one-half of the training failures in each group scored 33 or less on the DHT. Moreover, over three-fourths of the trainees in either group who attained DHT scores of 44 or greater passed Academy training with "non-marginal" A+L Grades. Also, the majority of subjects within both groups who scored between 34 and 43 on the DHT were "non-marginals" whereas relatively few (i.e., 40.3 per cent and 30.0 per cent) of those with scores of 33 or less were represented in the upper three-fourths of the A+L Grade distribution.

An additional analysis was accomplished on the test data of the 281 GS-7 En Route and Terminal ATCS trainees in order to more fully assess the potential with which the DHT might be used in conjunction with the CSC Battery to further enhance the aptitude screening process. In accordance with a procedure analogous to that just described for the DHT scores, the distribution of CSC scores was examined to determine those cases representing the approximate upper and lower fourths of the array as well as the intermediate half. It was determined that 73 of the 281 subjects scored 260 or higher on the CSC Battery, that 136 scored between 230 and 259, and that the lower fourth consisted of 72 subjects with scores of 229 or less. The cases

within each of the three CSC score intervals were then subdivided in terms of the three previously established levels of DHT performance (i.e., score intervals of "33 &<," "34-43," and "44 &>") and also in terms of the three categories of training performance. In reviewing the results (Figure 6), it should first be noted that there is little evidence of any significant relationship between the CSC aptitude scores and training performance; however, previous studies have illustrated that the *power of the CSC Battery is in its successful screening of applicant groups*, whereas the differential scores of those *within the qualifying range* offer little potential for discriminating between levels of training performance. (This is an expected result: See *The American Psychologist*, 27:236-239, 1972, for discussions of this topic in a different selection context.)

The results depicted by the upper set of bar graphs in Figure 6 for the 73 trainees having CSC scores of 260 and above constitute convincing evidence that the DHT can be used effectively to improve the aptitude screening process. Ten of the 73 subjects represented in the upper fourth of the CSC score distribution scored 33 or less on the DHT and three of the ten failed training, three passed the course with "marginal" A+L Grades and only four were "non-marginals." Disproportionately fewer trainees within the DHT score ranges of "34-43" and "44 &>" failed or marginally passed. Similar results were obtained for the 136 students who had CSC scores of 230 to 259. Thirty-five of the 136 scored 33 or lower on the new test and only 45.7 per cent of the 35 passed with "non-marginal" grades, whereas over three-fourths of those with higher DHT scores did so. Results pertaining to the 72 trainees with CSC scores of 229 and lower follow the same pattern, but in more pronounced degree.

Such findings imply that the DHT is providing measures of certain aptitudes other than those encompassed by the CSC Battery. This was verified by the results of a series of multiple regression analyses. Preparatory to the latter analyses, intercorrelations and validities of the DHT and the various tests of the CSC Battery were first computed for the group of 281 GS-7's for whom CSC scores were available. Although not presented in any table, the correlations between the DHT and the different tests of the

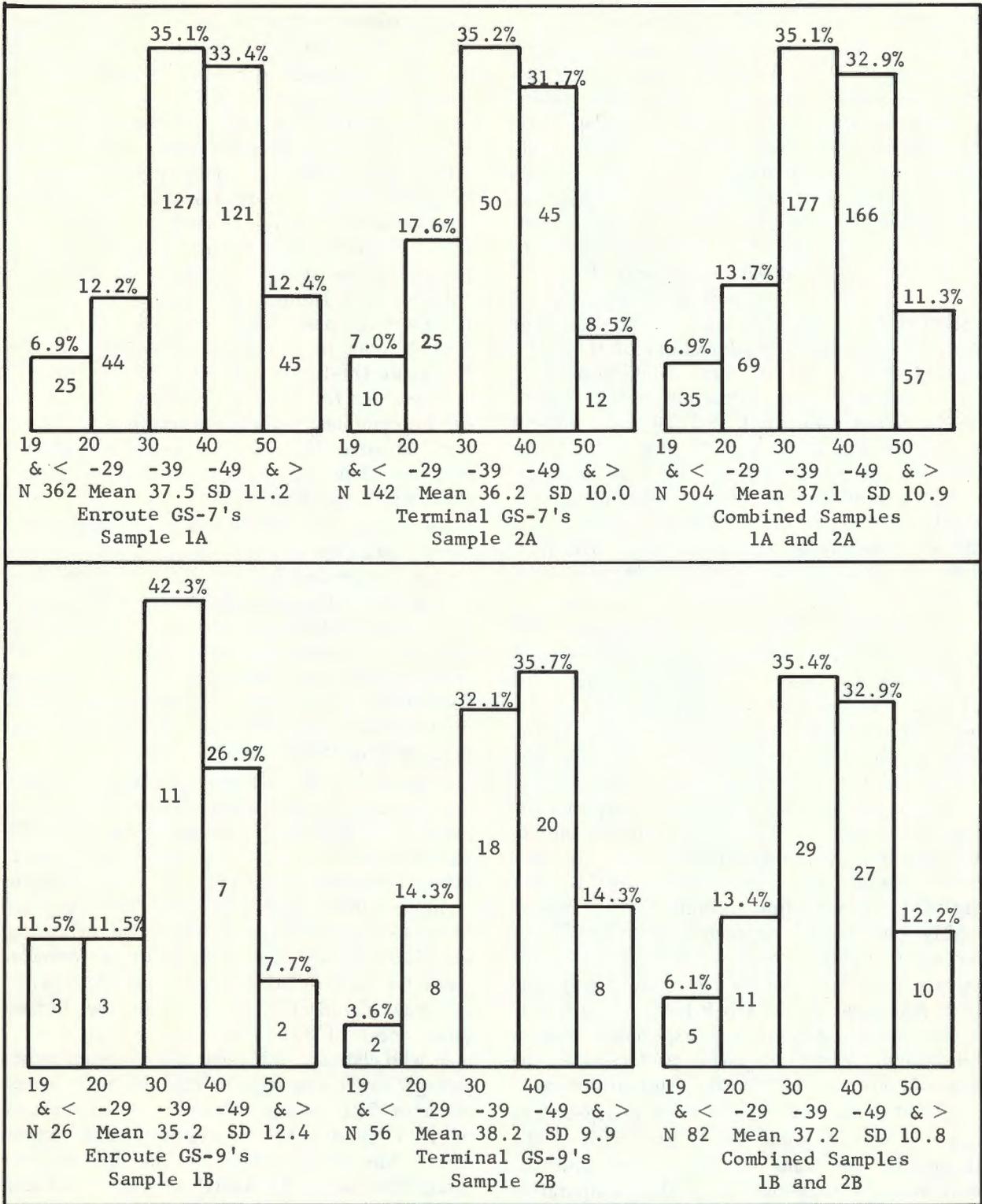


FIGURE 7. Comparison of ATC trainees of GS-7 and GS-9 levels in terms of performance scores on Directional Headings Test Form A.

CSC Battery ranged from .05 to .21; an r of .25 was obtained between the DHT score and the Composite Score for the entire group of CSC instruments. With A+L Grade and Pass-Fail Status serving as the criteria, the DHT yielded validity coefficients of .37 and .20, respectively. In contrast, the Composite CSC Scores correlated .13 with the A+L Grades and only .05 with the Pass-Fail criterion. An analysis was then accomplished, wherein the variances common to the DHT and each of the CSC tests were statistically partialled out (or theoretically nullified) to permit determination of the residual validities of the DHT alone. For prediction of A+L Grade, the residual validity of the DHT was .34, which was only three points lower than the r obtained before extraction of the associative variance. For P-F Status, the resulting validity of the DHT was .19—one point lower than the initial correlation. Such findings, although stemming from an analysis of data for a highly selected group of GS-7's, suggest that the aptitudes measured by the DHT are quite different from those assessed with the CSC Battery.

E. Pre-FAA ATC-Related Experience and DHT Performance. Two different analyses were undertaken to assess the possible influence of previous ATC-related experience upon DHT performance. Both were accomplished upon the data of the Form A examinees only. For the first of the two analyses, separate frequency distributions of the DHT scores were prepared for the GS-7's and GS-9's of each training option, and also for the combined En Route and Terminal groups. The GS-9's were known to have qualified for entry into training on the basis of highly specialized experience whereas the GS-7's possessed insufficient experience to warrant exemption of the aptitude screening requirement or appointment to the GS-9 level. It was reasoned that if ATC-related experience were a determining factor of DHT performance, the distributions would differ for the two groups.

Initially, the test scores were grouped into intervals of two points each. Inasmuch as the frequencies for some of the 2-point intervals were too low to permit a reliable comparative study, distributions were also prepared using both 5-point and 10-point intervals. Regardless of the grouping procedures employed, however, the differences between the distributions were

generally either very small or inconsistent insofar as they related to the GS level of the subjects of the En Route, Terminal, and combined samples. Although the distributions shown in Figure 7 are based on DHT score intervals of ten points each, they illustrate the inconsistency of the differences found between the subgroups. Test scores of 40 and higher were attained by proportionately fewer En Route trainees of GS-9 level than GS-7 level, whereas the opposite was true with regard to Terminal trainees. A comparison of the score distributions for the differentially rated trainees of the combined samples reveals only minor differences. The mean scores were 37.5 for the En Route GS-7's, 35.2 for the En Route GS-9's, 36.2 for the Terminal GS-7's, and 38.2 for the Terminal GS-9's. The GS-7's of the combined samples averaged 37.1 on the test compared to 37.2 for the GS-9's of both samples. None of the mean differences proved statistically significant.

The second analysis concerning experience was accomplished by dividing the 586 cases into subgroups on the basis of responses to those items of the CAMI Questionnaire wherein the subjects were asked to indicate whether they held, or ever had held, a certificate, license, or rating in air traffic control work, as a pilot, or in the general field of communications. The procedure resulted in the establishment of eight mutually exclusive subgroups or "experience" categories.

The number of trainees represented in each of the categories, and their DHT performance means by category, are presented in Table III. In the lower portion of the table for all examinees, irrespective of their GS level and/or training option, it will be noted that only 170 of the 586 indicated they possessed no certificated experience of any type. The DHT performance mean for these relatively non-experienced trainees was significantly lower than the highest group mean of 39.8 which pertained to the 147 men who claimed both pilot and communications ratings, yet it was significantly above the lowest mean of 31.1, which reflected the performance of the 10 pilots who also possessed ATC experience. The DHT scores of the 166 subjects categorized as "ATC Only" averaged 37.4 and those of the 36 pilots having no other experience averaged 37.5, but neither differed significantly from the mean of 36.8 obtained for the 19 trainees with ratings in "Communications Only." A

TABLE III--Directional Headings Test (Form A) performance means by pre-FAA experience groupings of students who entered basic ATC training at the Academy. (The "experience" groups are mutually exclusive, with subjects having a current or previous license, certificate or rating in ATC work and/or as a pilot or in communications work being assigned to only one of the categories.)

Samples		Type of Experience								Total Group
		ATC Only	Pilot Only	Commun. Only	ATC & Pilot	ATC & Comm.	Pilot & Commun.	ATC, Pil. & Comm.	No ATC P. or C.	
1A Enroute GS-7's	N:	44	27	16	1	1	121	6	146	362
	M:	36.5	37.4	38.7	41.0	40.0	40.2	35.7	35.5	37.5
	SD:	9.5	12.6	7.4			10.1	12.6	12.2	11.2
1B Enroute GS-9's	N:	14			6	3		2	1	26
	M:	38.4			27.7	36.0		37.0	30.0	35.2
	SD:	12.3			14.7					12.4
2A Terminal GS-7's	N:	67	9	3	2	8	25	7	21	142
	M:	36.7	37.9	26.7	31.0	39.6	38.1	33.3	32.9	36.2
	SD:	9.8	6.8			6.1	11.7	9.8	10.7	10.0
2B Terminal GS-9's	N:	41			1	5	1	6	2	56
	M:	39.1			42.0	29.2	42.0	41.2	30.0	38.2
	SD:	9.8						6.4		9.9
1A & 2A Enr. & Terml. GS-7's	N:	111	36	19	3	9	146	13	167	504
	M:	36.6	37.5	36.8	34.3	39.7	39.8	34.4	35.2	37.1
	SD:	9.7	11.4	8.7		5.7	10.4	10.7	12.0	10.9
1B & 2B Enr. & Terml. GS-9's	N:	55			7	8	1	8	3	82
	M:	38.9			29.7	31.8	42.0	40.1	30.0	37.2
	SD:	10.3			14.4	11.7		6.4		10.8
1A, 1B, 2A&2B Enr. & Terml. GS-7's & 9's	N:	166	36	19	10	17	147	21	170	586
	M:	37.4	37.5	36.8	31.1	35.9	39.8	36.6	35.1	37.1
	SD:	9.9	11.4	8.7	13.5	9.6	10.3	9.6	12.0	10.9

NOTE: The standard deviation was not computed for groups involving fewer than six subjects

number of other statistical analyses, including several analyses of variance (not presented in this report), were accomplished on the data of the GS-7's and GS-9's of the respective and combined samples. In each instance, however, the results failed to indicate any definite relationship between DHT performance and the various types of pre-employment experience surveyed in the study.

F. Age Versus DHT Performance. The results obtained in a number of previous studies^{3 5 15 16 17 18 19} have been remarkably consistent in demonstrating that older men, particularly those over 35 years of age, tend to perform less well than their younger colleagues in various stages of training and as journeymen controllers. Moreover, the older subjects also tend to score significantly lower than the younger men on a great variety of aptitude tests, including most of those comprising the CSC Battery.^{6, 8, 5 15 19} Negative correlations, ranging from $-.12$ to $-.38$, have been obtained between the ages and Composite CSC Aptitude scores of ATC trainees.^{6, 8, 5 19} Although no survey of *applicants* has yet been conducted, it is commonly presumed that a far greater proportion of those within the older age brackets are unable to meet the current aptitude screening requirement than those who are relatively young. FAA officials have considered this somewhat advantageous; they have long sought, and are currently seeking, congressional approval to utilize age as a factor in the selection of ATC personnel. Should such approval not be granted, however, the FAA would continue to rely upon aptitude screening instruments in precluding the establishment of eligibility of at least a portion of the older applicants.

On the basis of previous findings^{3 5 15 16 18 19} with a great variety of aptitude tests, we hypothesized that chronological age would correlate negatively with DHT performance. The hypothesis held but the correlations (which are not presented in any table) were all rather low. An r of $-.04$ was obtained between age and DHT performance for the 388 ARTCC trainees of Sample 1; however, only 40, or 10.3 per cent, of the 388 were over 35 years old. In contrast, a significant correlation of $-.22$ was obtained between the two variables for the 198 TATC trainees of Sample 2; slightly over 15 per cent of the TATC trainees were age 36 or older.

The two variables correlated $-.10$ ($p < .05$) for the 586 subjects of the combined samples.

Considering that the majority of the trainees comprising each sample were GS-7's who had qualified on the CSC ATC Test Battery, and inasmuch as studies preceding operational implementation of the battery illustrated significant inverse relationships between the aptitude measures and age, it may be presumed that the groups examined with the DHT included disproportionately fewer men in the older age brackets than would be found among applicant groups. Inasmuch as the correlation coefficient of $-.22$ (between the ages and DHT scores of the subjects in Sample 2) is substantially larger than that obtained for the sample having the smaller proportion of older trainees, it would be expected that a correlation of even greater magnitude (than $-.22$) would be found between the two variables for subjects representative of an unscreened applicant population.

Further evidence of a definite relationship between age and DHT performance was obtained by comparing the distribution of scores of the 70 trainees of age 36 and over (Samples 1 and 2 combined) with that of the remaining 516 younger Form-A examinees (see Table IV). Slightly over 31 per cent, or 22 subjects, of the older group scored less than 30 on the DHT whereas only 19 per cent ($N=98$) of the younger trainees did so. The difference between the latter proportions was statistically significant at the .05 level. However, other findings implied that the relationship was nonlinear. For example, exceptionally high scores of 50 and better were achieved by 10 per cent ($N=7$) of the 70 older trainees and by 11.5 per cent ($N=60$) of the younger men.

The two age groups also differed markedly with respect to training-course failure rates. About 44 per cent ($N=31$) of those over 35 years old were attrited, compared to a significantly lower failure rate of 16.3 per cent for the younger group. Moreover, DHT scores of 29 and less were obtained by 13, or 41.9 per cent, of the 31 older attritions and by only 9, or 23 per cent, of the 39 subjects over age 35 who passed. The difference between these proportions was not significant due to the small number of cases involved. Scores of less than 30 were also obtained by 38, or 45.2 per cent, of the 84 training-course failures who were less than

35 years old and by 60, or 13.9 per cent, of the 432 graduates within the same age group. The difference between the latter proportions was significant at the .01 level. Such findings suggest that a revision of current ATC selection procedures to include consideration of both age and DHT performance would result in an appreciable reduction of the basic-training attrition rate. However, the results illustrate that, although age and DHT performance are substantially related, the DHT could be used with somewhat greater effectiveness than age in the screenout of applicants most likely to fail training.

G. *Addendum: Use of the DHT with Journeyman Controllers.* Prior to the completion of this paper, the DHT was made available to

Education and Public Affairs (EPA), a private company which was performing an ATCS selection study for the FAA. In their final report based on samples of journeymen controllers with 3-10 years of experience ("Selection of Air Traffic Controllers for Federal Aviation Administration" by Ann M. Milne and Joseph G. Colmen, Final Report on Contract DOT-FA 70WA-2371, January, 1972), EPA researchers noted the effectiveness of the DHT in predicting job success by citing it (page VI-6 of their report) as one of the best of large numbers of aptitude tests which they evaluated. Thus the promising findings obtained with ATCS trainees in the present paper has been independently confirmed with a different group of journeyman controllers.

IV. Summary and Conclusions.

This study concerned the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been se-

lected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.

Table IV.--Form-A score distributions by dichotomized age groups for subjects who passed or failed Academy basic training.

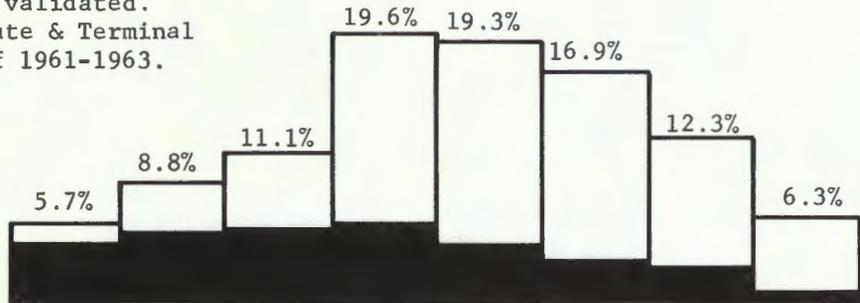
Form A DHT Score	Age 35 & <			Age 36 & >			All Ages		
	Pass	Fail	Total	Pass	Fail	Total	Pass	Fail	Total
50 & >	56	4	60	4	3	7	60	7	67
40-49	159	17	176	10	7	17	169	24	193
30-39	157	25	182	16	8	24	173	33	206
29 & <	60	38	98	9	13	22	69	51	120

Total:									
Passes	432			39			471		
Fails		84			31			115	
P + F			516			70			586

APPENDIX A. Distributions of CSC ATC Aptitude Test Battery performance scores for the 893 experimentally-examined entrants into Academy ATC training during 1961-1963 and for 281 operationally-examined trainees who entered during 1970.

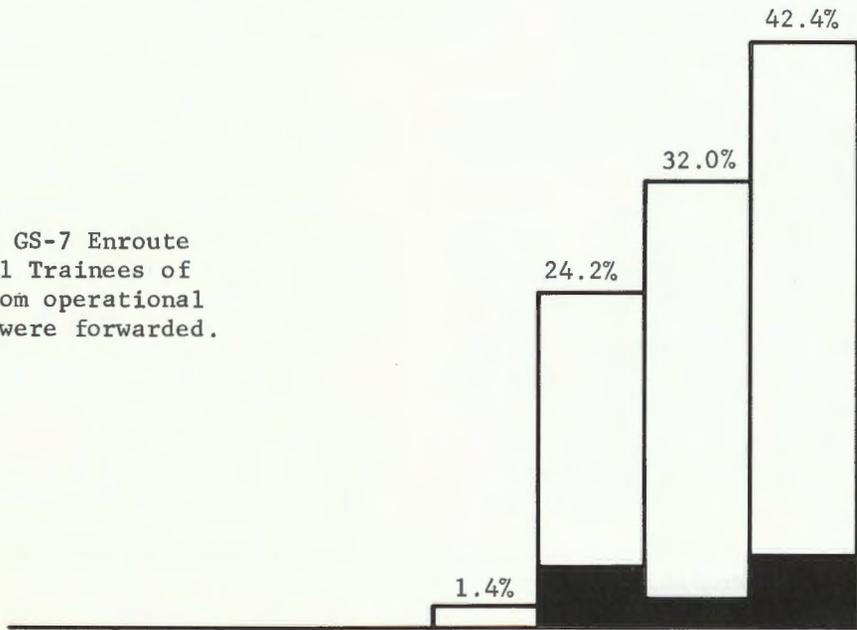
Sample on which the CSC ATC Test Battery was originally validated. N=893 Enroute & Terminal Trainees of 1961-1963.

Academy Trng. course Failures
 Academy Trng. course Passes



CSC Score:	129 & <	130 -149	150 -169	170 -189	190 -209	210 -229	230 -249	250 & >	N Total
N Pass:	13	33	52	124	135	126	87	52	622
N Fail:	38	46	47	51	37	25	23	4	271
N Total:	51	79	99	175	172	151	110	56	893

Sample of 281 GS-7 Enroute and Terminal Trainees of 1970 for whom operational CSC scores were forwarded.



CSC Score:	129 & <	130 -149	150 -169	170 -189	190 -209	210 -229	230 -249	250 & >	N Total
N Pass:					4	56	85	105	250
N Fail:						12	5	14	31
N Total:					4	68	90	119	281

REFERENCES

1. Air Traffic Controller Career Committee Report: The Career of the Air Traffic Controller—A Course of Action. Executive Summary of the Report of the Air Traffic Controller Career Committee. FAA, Washington, D.C., January 1970.
2. Brokaw, L. D.: School and Job Validation of Selection Measures for Air Traffic Control Training. WADC-TN-59-39, Personnel Laboratory, WADD, USAF, Lackland AFB, Texas, 1959.
3. Cobb, B. B.: Problems in Air Traffic Management: II. Prediction of Success in Air Traffic Controller School. FAA Civil Aeromedical Research Institute Report No. 62-2, 1962.
4. Cobb, B. B.: Problems in Air Traffic Management: V. Identification and Potential of Aptitude Test Measures for Selection of Tower Air Traffic Controller Trainees. FAA Office of Aviation Medicine Report No. AM-65-19, 1965.
5. Cobb, B. B., C. D. Lay, and N. M. Bourdet: The Relationship Between Chronological Age and Aptitude Test Measures of Advanced-Level Air Traffic Control Trainees. FAA Office of Aviation Medicine Report No. AM-71-36, 1971.
6. Davis, C. G., R. Kerle, A. W. Silvestro, and W. H. Wallace: Identification of Training Requirements in Air Traffic Control. FAA/BRD-40, Project O, Report 36, 1960.
7. Davis, C. G., R. Kerle, A. W. Silvestro, and W. H. Wallace: An Activity Analysis of the Positions in a High Activity Air Traffic Control Tower. FAA/BRD-40, Project O, Report 39, 1960.
8. Davis, C. G., W. H. Wallace, R. H. Kerle, and A. W. Silvestro: Analysis of Controller Functions at the High Altitude Sectors of the New York ARTC Center. FAA/BRD-40, Project O, Report 35, 1960.
9. Davis, C. G., W. H. Wallace, R. H. Kerle, and A. W. Silvestro: Activity Analysis of Positions in a High Activity ARTC Center. FAA/BRD-40, Project O, Report 37, 1960.
10. DuBois, P. H.: *Multivariate Correlational Analysis*, New York, Harper & Brothers, 1957, 202 pp.
11. Guilford, J. P.: *Fundamental Statistics in Psychology and Education*, York, Pennsylvania, Maple Press Company, 1956.
12. Nagay, J. A.: The Development of a Procedure for Evaluating the Proficiency of Air Route Traffic Controllers. Civil Aeronautics Administration, Report 83, 1949.
13. Nagay, J. A.: Field Tryout of a Procedure for Evaluating the Proficiency of Air Traffic Controllers. Civil Aeronautics Administration, Report 91, 1950.
14. Sullivan, E. T., W. W. Clark, and E. W. Tiegs: *Manual for California Test of Mental Maturity, Advanced Form*, Los Angeles, California Test Bureau, 1957, 39 pp.
15. Trites, D. K.: Problems in Air Traffic Management: I. Longitudinal Prediction of Effectiveness of Air Traffic Controllers. FAA Civil Aeromedical Research Institute Report No. 61-1, 1961.
16. Trites, D. K.: Problems in Air Traffic Management: VI. Interaction of Training-Entry Age With Intellectual and Personality Characteristics of Air Traffic Control Specialists. FAA Office of Aviation Medicine Report No. AM-65-21, 1965.
17. Trites, D. K., and B. B. Cobb: Problems in Air Traffic Management: III. Implications of Age for Training and Job Performance of Air Traffic Controllers. FAA Civil Aeromedical Research Institute Report No. 62-3, 1962.
18. Trites, D. K., and B. B. Cobb: Problems in Air Traffic Management: IV. Comparisons of Pre-Employment, Job-Related Experience With Aptitude Tests as Predictors of Training and Job Performance of Air Traffic Control Specialists. FAA Civil Aeromedical Research Institute Report No. 63-31, 1963.
19. Trites, D. K., and B. B. Cobb: CARI Research on Air Traffic Control Specialists: Age, Aptitude, and Experience as Predictors of Performance. FAA Civil Aeromedical Research Institute, unnumbered report, 1964.

Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute, Oklahoma City, Oklahoma. A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS, by Bart B. Cobb and John J. Mathews, May 1972, 26 pp. Report No. FAA-AM-72-18.

I. Cobb, B. B.
II. Mathews,
J. J.

Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute, Oklahoma City, Oklahoma. A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS, by Bart B. Cobb and John J. Mathews, May 1972, 26 pp. Report No. FAA-AM-72-18.

I. Cobb, B. B.
II. Mathews,
J. J.

This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations

Descriptors
Air Traffic
Controllers
Aptitudes
Personnel
Selection

This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations

Descriptors
Air Traffic
Controllers
Aptitudes
Personnel
Selection

of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.

of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.

Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute, Oklahoma City, Oklahoma. A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS, by Bart B. Cobb and John J. Mathews, May 1972, 26 pp. Report No. FAA-AM-72-18.

I. Cobb, B. B.
II. Mathews,
J. J.

Descriptors
Air Traffic
Controllers
Aptitudes
Personnel
Selection

This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations

of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.

Federal Aviation Administration, Office of Aviation Medicine, Civil Aeromedical Institute, Oklahoma City, Oklahoma. A PROPOSED NEW TEST FOR APTITUDE SCREENING OF AIR TRAFFIC CONTROLLER APPLICANTS, by Bart B. Cobb and John J. Mathews, May 1972, 26 pp. Report No. FAA-AM-72-18.

I. Cobb, B. B.
II. Mathews,
J. J.

Descriptors
Air Traffic
Controllers
Aptitudes
Personnel
Selection

This study concerns the development and experimental validation of a novel aptitude test, referred to as "Directional Headings" (or DHT), for the selection of Air Traffic Control Specialist (ATCS) trainees. The test requires the subject to rapidly interpret letters, symbols, and degrees in order to determine: directional headings (Part I), the exact opposites of headings (Part II), and opposites of headings under conditions of aural distraction (Part III). The DHT was administered on an experimental basis to several hundred men as they entered basic ATCS training at the FAA Academy. The vast majority of the subjects had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations

of pre-FAA experience. Despite these pre-screening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 per cent of the 115 examinees who failed the training course scored no higher than 29 on the DHT whereas over 85 per cent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the scores based on even-numbered items versus scores based on odd-numbered items, was .93.