AD P001383

Shortening of Defense Language Aptitude Battery

John W. Thain
Education Specialist
Defense Language Institute
Foreign Language Center

The Defense Language Institute, Foreign Language Center (DLIFLC) is the proponent for the Defense Language Aptitude Battery (DLAB), used in screening recruits for aptitude to learn foreign languages. DLIFLC is attempting to shorten DLAB without decreasing its reliability or validity. Item analysis data was used to plan eleven possible shortening strategies with different number of items deleted from original test. One hundred sixty-four answer sheets were rescored and recorrelated with a foreign language course grade after items had been deleted according to each strategy. Tests from another sample will also be rescored and recorrelated with the criterion using each of the test shortening strategies in order to cross-validate the original results.

SHORTENING OF THE DEFENSE LANGUAGE APTITUDE BATTERY

John W. Thain
Education Specialist
Defense Language Institute
Foreign Language Center

Background.

The Defense Language Aptitude Battery (DLAB) is administered to determine if enlisted personnel should be programmed for language training. Previous studies by the Defense Language Institute Foreign Language Center (DLIFLC) have shown the correlation between DLAB scores of students selected for language training and their course grades in DLIFLC language courses to be consistently around .40; if this correlation is corrected for restriction of range, it rises to around .60. In recent years, only about 20% of the recruits taking the test have achieved a passing score. The test takes 90 minutes to administer and consists of 119 multiple-choice test items.

The DLAB is administered at several locations; the most important of which are the local Armed Forces Entrance and Examination Station (AFEES) throughout the country, and Lackland Air Force Base in San Antonio, Texas (for Air Force enlisted men). The agencies responsible for administering DLAB at those locations are also responsible for screening recruits for most of the other occupational specialities in the Services. Largely on grounds of administrative convenience, these agencies favor any proposal that would shorten the process of programming enlistees to their subsequent assignments. In particular the AFEES have expressed interest in reducing the current administration time of 90 minutes required to administer DLAB. There are two specific reasons for this interest in shortening DLAB:

- A shorter test would contribute to the shortening of the overall screening process and thus help avert the possibility that the AFEES (Armed Forces Entrance and Examination Stations) would have to pay the overnight expenses of recruits if the processing of the recruits had to be extended an additional day.

- A minority of students taking DLAB pass the test. From the point of view of test examinees, every minute spent by failing examinees on DLAB is wasted.

Of course, shortening a test tends to decrease test reliability and validity. Using a less valid and reliable version of a test will result in less effective screening of potential students and lead to lower student performance and increased student attrition at DLIFLC.

Therefore, DLI designed a research study to determine whether DLAB could be shortened without substantially reducing test validity and reliability.

Procedure.

Two methods of shortening the test were considered - condensing redundant instructions or by deleting poorly functioning test items. At the very start of the project we needed to analyze in detail the factors contributing to length of test administration.

The following table breaks down the time required to administer DLAB. The total time required to administer each part of the test and the test as a whole is given in minutes and seconds. The average time required to administer each item in each part is computed by dividing the time required to administer each part by the number of items in that part. The total administration time and the average per item administration time is broken into two parts, that required by the test items themselves and that taken up by instructions. Although items are numbered 1 through 126, only 119 items are scored; seven practice items are not scored.

TABLE I

TIME REQUIRED TO ADMINISTER DLAB: TIME TAKEN
ITEM TYPE AND CORRESPONDING SET OF INSTRUCTIONS

PART OF TEST	TEST ITEMS	NO. ITEMS	ITEMS TOTAL MIN	ITEMS TOTAL SEC	ITEMS PER ITEM SEC	TIME TAKEN BY INSTRUCTIONS TOTAL MIN	TIME TAKEN BY INSTRUCTIONS TOTAL SEC	TIME TAKEN BY INSTRUCTIONS PER ITEM SEC	BOTH TOTAL MIN	BOTH TOTAL SEC	BOTH PER ITE SEC
I	1-7	7	1	33	13.22	0	36	5.14	2	09	18.36
II	9-26	18	6	25	21.38	1	13	4.06	7	38	28.76
III-1	28-40	13	5	31	25.46	2	46	12.76	8	17	38.22
III-2	42-55	14	5	35	23.96	2	30	10.71	8	05	34.67
III-3	57-73	17	8	46	30.94	3	37	12.76	12	24	43.70
III-4	74.93	20	12	43	38.15	3	58	11.90	16	41	50.05
IV	97-126	30	(22	00)*	44.00	(3	00)*	6.00	25	00	50.00
TOTAL	1-126	119	62	29	31.50	17	44	8.94	90	13	40.44

*ESTIMATED AVERAGE. THIS PART NOT MACHINE PACED.

The Defense Language Aptitude Battery is a multiple-choice test with four parts. The first three parts of the test are paced by an audio tape. Part I is a self-report biographical inventory. In Part II and Part III the examinee learns an artificial language. In Part II the examinee discriminates stress patterns. In Part III the examinee selects the correct translation in the artificial language on the basis of grammatical rules provided in the instructions. Part III has four subparts. In Part IV, which the examinee is to complete in 25 minutes while working at his own pace, the examinee matches pictures to phrases in an artificial language according to rules given in the instructions.

When the test was reviewed, it was found that no meaningful savings would result if all four parts of the test were retained in their present form. Of course, if a whole part or subpart of the test were deleted, the instructions for that part could also be deleted.

Table I shows that every item deleted from Part III and IV will save more administration time than an item deleted from Part I or II; the table also shows that deleting all of Part IV or any subpart of Part III will save more administration time than deleting all of Part I or Part II.

An item analysis was conducted on DLAB in order to rank-order the items in terms of their overall contribution to test validity. 164 answer sheets were included in the sample. Final course grade at DLIFLC was used as a criterion. Item analysis data was used along with the analysis of the time required for test administration mentioned earlier to decide which items to delete. Eleven strategies for shortening the test were devised. The number of items deleted from the test in the various strategies ranged from four to sixty-seven items out of a total of 119 items on the original test. The time saved in the various strategies ranged from 3 to 49 minutes out of the 90 minutes required for the original test. The strategies that deleted more items either involved setting the minimum item-criterion correlation higher on the one hand, or deletion of whole parts of the test regardless of the intercorrelations between individual items and the criterion on the other hand. The tests were rescored according to each of the eleven strategies with the corresponding items deleted, resulting in a new distribution of test scores for each strategy. These new distributions were recorrelated with the criterion. By using the average time required to administer each type of item and the time required for instructions when whole sections of the test were deleted, the time required for test administration in each of the eleven strategies was computed. The results are at Table 2.

TABLE II

DLAB SHORTENING STRATEGIES - ITEMS DELETED AND CORRESPONDINGLY RECOMPUTED VALIDITY COEFFICIENTS AND ADMINISTRATION TIMES

ITEMS DELETED	RECOMPUTED CORRELATION WITH CRITERION	PROJECTED ADMINISTRATION TIME
Original Test	.366	90 min. 13 sec.
70,78,98,113	.390	87 min. 16 sec.
59,70,76,78,89,96,98,113	.396	84 min. 41 sec.
2,14,26,30,42,51,55,70,77, 78,83,97,98,100,101,113, 114	.452	81 min. 1 sec.
70 - 89	.383	73 min. 32 sec.
2,14,26,30,40,42,51,55, 70,77,78,83,85,97,98,100, 101,103,106,113,114	.468	78 min. 4 sec.
90 - 119	.347	65 min. 13 sec.
2,14,15,23,26,28,29,30 36,40,41,42,50,51,55,59, 63,70-89,96,97,98,100,101, 103,104,106,113,114	.480	59 min. 15 sec.
70 - 119	.353	48 min. 32 sec.
2,14,15,23,26,28,29,30 36,40,41,42,50,51,55,59,63 70,76,77,78,80,83,85,89 90-119	.4505	54 min. 11 sec.
2,14,15,23,26,28,29,30,36, 40,41,42,50,51,55,59,63, 70-119	.4225	41 min. 35 sec.

The three most promising scenarios are compared at Table II.

TABLE III

THREE SCENARIOS FOR REDUCING DLAB LENGTH

Scenario	Time	VALIDITY: Uncorrected for Range	VALIDITY: Corrected for Range
Original 119 Items	90 min. 13 sec.	.366	.555
72 Test Items	59 min. 15 sec.	.480	.684
64 Test Items	53 min. 11 sec.	.451	.650
52 Test Items	41 min. 35 sec.	.423	.620

Since item analysis data was used to decide which items to eliminate there is a high probability that some sampling error is involved. Our next step is to gather a large sample of approximately 350 answer sheets and crossvalidate our initial results. If the same shortening strategies yield the same results, we will feel justified in shortening the test.

Since DLIFLC receives the completed answer sheets of students from the AFEES and Lackland Air Force Base, and the final course grade of all past DLIFLC students are also available to us as criterion measures, there is no need for DLIFLC to conduct additional test administration or wait for a long period of time to collect criterion data when drawing a crossvalidation sample.