AD NUMBER
ADB951175
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; 09 OCT 1952. Other requests shall be referred to Office of the Adjutant General (Army), Washington, DC 20310.**
AUTHORITY
AGO per DTIC form 55

UNANNOUNCED copy 2

RESEARCH NOTES

Number 52-71

(11) 9 October 1952

(12) 11

(6) FURTHER ANALYSIS OF RESPONSES TO THE API-1
PJ-3508-10

(14) AGO-PRS-RESEARCH NOTE-52-71

PERSONNEL RESEARCH SECTION
PR AND P BR, TAGO

Research Note 52-71

FURTHER ANALYSIS OF RESPONSES TO THE API-I

BACKGROUND AND PROBLEM

In developing the revised classification testing program for Army recruits, it appeared desirable to attempt the construction of a personality test which might have usefulness as a screening instrument. The Army Personality Inventory, API-1 (WD AGO PRT 703) was constructed for this purpose. This inventory consists of 300 items taken from the Minnesota Multiphasic Personality Inventory, a widely recognized commercial test. A number of studies were then undertaken to determine the usefulness of the API-1 as a screening device.

In 1947, a study was conducted in which the API-1 was administered to approximately 11,000 men at Army Replacement Training Centers. Several item analysis scoring keys were developed against a "reason for discharge" criterion within each of two samples of the population. Upon cross-validation, the two more useful of the keys were each found to correlate .22 with the criterion.(2) The validity of these keys in combination was .26. While promising, these validities were not high enough to warrant the use of the API-1 as a screening device.

Concurrently, a different study which was concerned with the West Point Self-Description Blank (DA AGO PRT 885) had led to the preparation of a "suppressor"[1] key which raised the validity of the valid key composite from .16 to .21.(3) This constituted a gain of approximately 31% in the selective efficiency of the predictor.[2] It was felt that a similarly prepared key to an instrument of higher intrinsic validity such as the API-1 could be expected to raise the validity of this instrument by an even greater proportional amount.

The present study was undertaken to determine whether a suppressor key could be developed which would have the effect of raising the validity of the API-1 to a useful level.

Certain other findings in the original study of the API-1 items raised a second problem. Negative relationships had been found between AGCT scores and neurotic tendencies and between years of education and neurotic tendencies as

---

1/ "Suppressor" key is used herein to mean a scoring key composed of items which correlate highly with scores on the test as a whole and which correlate zero or negatively with the criterion. The effect of combining such a key with a valid predictor key is to partial out non-valid components of the predictor composite. With such effects partialled out of it, the predictor key should then show higher correlation with the criterion.

2/ The use of percentage increase in the size of validity coefficients to indicate increase in efficiency of selection has been treated in detail in a paper by Brogden.(1) In this paper he has shown that r "...indicates directly the proportion of maximum saving that is actually obtained with the use of a given test where the maximum saving is that obtained by selecting on the criterion."

measured by API-1 score. It was reasonably well established that these relationships were not solely due to difficulty of reading and comprehension of the items by the men with less education and lower AGCT scores.(4) The problem remained, however, as to whether the relationship among education, AGCT score and neurotic tendency score is curvilinear. It was therefore the second objective of this study to examine this relationship for possible curvilinearity.

METHOD

GENERAL DESIGN

The basic design of the suppressor key phase of this study consisted of three steps. First, the preparation of the suppressor keys by examination of the item analysis data for the selection and keying of items which were highly correlated with scores on the test as a whole and uncorrelated (or negatively correlated) with the criterion. Second, the determination of the validities of the newly prepared keys and the intercorrelation of all keys, old and new. Third, the computation of multiple correlations to determine the effect of the suppressor keys on the validity of the original keys.

The second phase of this study consisted of investigation of the linearity of the relationship between years of education and AGCT score on the one hand and API-1 score on the other hand. This was done by computing tetrachoric correlation coefficients at various cuts of both variables.

POPULATION

Populations A and B of the basic API-1 study(2) were used for the present study. Both populations[3] were used in the first phase while only population A was used in the second phase. Population A consisted of 1806 men who took the API-1 who remained at the training centers at least eight weeks, and upon whom follow-up information after discharge was available.[4] Following discharge they were classified into the following two criterion groups:

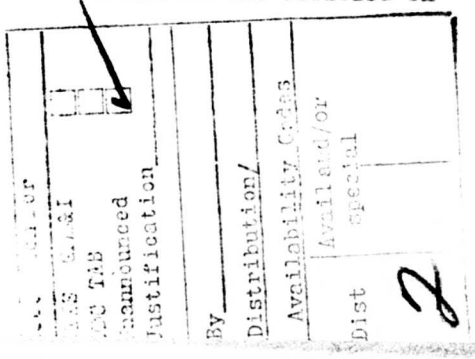
Favorable Discharge (N = 1580)

Unfavorable Discharge (N = 226)

Population B is made up of 1469 men tested with the API-1 who did not remain at the training centers long enough to provide data on their behavior at these

3/ The populations reported in the present study are somewhat smaller than those listed in the earlier report. A few cases were lost for miscellaneous reasons.

4/ This method of population breakdown was employed because of the design of the original experiment (2) in which an additional criterion was utilized in examining part of the data.



measured by API-1 score. It was reasonably well established that these relationships were not solely due to difficulty of reading and comprehension of the items by the men with less education and lower AGCT scores.(4) The problem remained, however, as to whether the relationship among education, AGCT score and neurotic tendency score is curvilinear. It was therefore the second objective of this study to examine this relationship for possible curvilinearity.

METHOD

GENERAL DESIGN

The basic design of the suppressor key phase of this study consisted of three steps. First, the preparation of the suppressor keys by examination of the item analysis data for the selection and keying of items which were highly correlated with scores on the test as a whole and uncorrelated (or negatively correlated) with the criterion. Second, the determination of the validities of the newly prepared keys and the intercorrelation of all keys, old and new. Third, the computation of multiple correlations to determine the effect of the suppressor keys on the validity of the original keys.

The second phase of this study consisted of investigation of the linearity of the relationship between years of education and AGCT score on the one hand and API-1 score on the other hand. This was done by computing tetrachoric correlation coefficients at various cuts of both variables.

POPULATION

Populations A and B of the basic API-1 study(2) were used for the present study. Both populations3/ were used in the first phase while only population A was used in the second phase. Population A consisted of 1806 men who took the API-1 who remained at the training centers at least eight weeks, and upon whom follow-up information after discharge was available.4/ Following discharge they were classified into the following two criterion groups:

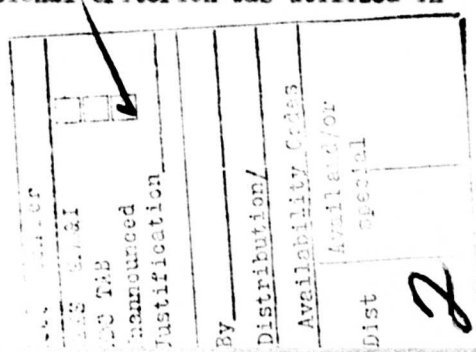
Favorable Discharge (N = 1580)

Unfavorable Discharge (N = 226)

Population B is made up of 1469 men tested with the API-1 who did not remain at the training centers long enough to provide data on their behavior at these

3/ The populations reported in the present study are somewhat smaller than those listed in the earlier report. A few cases were lost for miscellaneous reasons.

4/ This method of population breakdown was employed because of the design of the original experiment (2) in which an additional criterion was utilized in examining part of the data.



installations, but for whom follow-up information after discharge was also available. These men were also classified into two criterion groups as follows:

Favorable Discharge (N = 1205)

Unfavorable Discharge (N = 266)

VARIABLES

Criterion. The dichotomous unfavorable-favorable discharge criterion used in the earlier study was used as the criterion in the suppressor key phase of the present study. This dichotomy was based upon examination and subsequent grouping of specific causes for discharge. Favorable discharges included: Expiration of term of service; minority discharge; certificate of disability discharge; accept a commission; unlikely to render effective service upon return to duty from hospital; deceased; dependency or hardship; and re-enlisted. Unfavorable Discharges were: Dishonorable Discharge; conviction by civil court; fraudulent enlistment; physically unfit deserters and absentees; undesirable habits and traits of character; inaptness and lack of adaptability for military service; and enuresis.

In the second phase of this investigation the criterion consisted of scores on Key 10A which is listed as Predictor Variable 1, below.

The six predictor variables used in phase one of this study consist of scores obtained by use of item keys which were derived from responses of the population groups to API-1 (WD AGO PRT 703). The first four of these keys were developed in the previously mentioned original study (2) of the API. The remaining two keys were developed as a part of attaining the objective of the present study. The specific predictor variables used were:

1. Score on Key 10A, a key composed of 49 item responses which correlated with the criterion ($r_{tet} \geq .10$) and which were selected by between 15% and 85% of the subjects in Population A.

2. Score on Key 10B, a key of 53 items corresponding to Key 10A, but based upon the responses of the subjects in Population B. Keys 10A and 10B are referred to in the remainder of this paper as the "Most Valid Keys."

3. Score on Key 11A, a key consisting of 88 items having p-values from .00 to 1.00 and correlations with the criterion $r_{tet} \geq .10$, based on the responses of Population A. This key includes all of the items in Key 10A plus those of more extreme p-values.

4. Score on Key 11B, a key of 72 items corresponding to those in Key 11A, but selected on Population B. This key and Key 11A are designated the "All Valid Keys" in the remainder of this paper.

5. Score on Key 12A. This key and Key 12B are the "Suppressor keys" devised for purposes of this study. Twenty-seven items were selected because of their low correlation with the criterion and high correlation with the score on Key 10A. Items selected by between 22% and 80% of Population A were given preference. Items with more extreme p-values were selected only when their correlations with the total test were very high and their validities were .00 or, if higher, with the sign of the validity coefficient in the opposite direction to the sign of the total test score correlation. The items in this key and in Key 12B were keyed with signs opposite to their total score correlations for convenience in combining keys later.

6. Score on Key 12B, the second "Suppressor Key" consisting of 29 items selected within Population B on the same basis as those making up Key 12A.

7. The predictor variable used in the second phase of the study consisted of a composite of years of education and AGCT score.

STATISTICAL ANALYSIS

Scores on Predictor Variables 5, 6, 7, and 8 were obtained for the men in each sample. The scores were then correlated with the previously determined scores on Predictor Variables 1, 2, 3 and 4; with each other; and with the criterion. The actual values of these coefficients are given in Appendix A. They were used to derive the multiple correlations which comprise the essential results of this study.

Multiple correlation coefficients were next computed on the data for Population A. Correlations were obtained between the Most Valid Key (10A), the Suppressor Key (12A), and the criterion; and between the All Valid Key (11A), the Suppressor Key (12A), and the criterion. The weights thus derived were subsequently used for the cross-validation of similar combinations of these keys in Population B. The same procedure was used in cross-validating the weighted combinations of Keys 10B and 12B and Keys 11B and 12B.

The final step in the analysis of the first phase data was to obtain the multiple correlation of the two Most Valid Keys (10A and 10B) with the criterion when the effects of both Suppressor Keys (12A and 12B) were included in the Most Valid Key scores. The two cross-validities obtained (on Population B) for Keys 10A and 12A and those obtained (on Population A) for Keys 10B and 12B were used to derive the beta weights and multiple correlation with the discharge criterion. As there were available only the two populations in which the keys were developed and cross-validated, no completely independent measure was available of the intercorrelations of keys developed on different samples. However, because the standard deviations and the intercorrelations of all variables were equivalent in both populations the necessary use of the possibly "contaminated" intercorrelations was not considered as leading to too great an error in the estimate of the multiple correlation in an independent sample. In any case, such contamination would have the effect of increasing the intercorrelations thus lowering rather than inflating the derived coefficient.

In the second phase of this study, i.e., the investigation of the linearity of the relationship between the API-1 score (Key 10A) and the combined Education AGCT score variables, scatterplots were made and tetrachorics computed at four points of cut on each variable. The diagrams and the coefficients were then examined for evidences of curvilinearity.

RESULTS AND DISCUSSION

The multiple correlation of the two Most Valid Keys (10A and 10B) with the criterion was found to be .26. However, when the cross-validities of these keys were combined with those for the Suppressor Keys (12A and 12B), the Multiple R rose to .32 (See Table 1). This represents an increase of .06 or 23 percent[5/] over the combination of the Most Valid Keys without suppressors. Although this validity coefficient of .32 is not as high as might be desirable, it is still highly promising for future research with this instrument. This is particularly true if it is to be used in combination with other tests where its economy of administration would warrant its use to gain a relatively small contribution to a total battery.

Table 1. Weights used to determine the combined validity of the Most Valid and the Suppressor Keys.

Key	Cross Validity	Beta Weight
10A	.22	.16
12A	-.04	.03*
10B	.22	.31
12B	-.01	.22*
	R = .32	

*Positive weights resulted from the reversal of sign in scoring suppressor keys after the items had been selected.

When the two Suppressor Keys are separately combined with the Most Valid Keys (10A and 10B) and the All Valid Keys (11A and 11B), the results (Table 2) showed differing effects upon the several rights keys in the two populations.

5/ For use of percent increase in R as an indication of increase in predictive efficiency, See (1)

Table 2. Cross validation of API-1 keys. Raw score weights derived in one population applied to the scores in the second population.

Population A			Population B		
Combining Weights*	and Variables	Validity	Combining Weights*	and Variables	Validity
1.58 Key 10B	1.00 Key 12B	.28	3.13 Key 10A	1.00 Key 12A	.22
1.14 Key 11B	1.00 Key 12B	.28	1.27 Key 11A	1.00 Key 12A	.22

*Weights were obtained by dividing the beta weights for the scoring keys by the beta for the suppressor keys to indicate the effects on scoring keys with suppressors held constant.

In Population A, the Suppressor Key (12B) developed on Population B raised the cross-validity of the Most Valid Key (10B) from .22 to .28, approximately 27%. The same Suppressor key added to the cross validity of the All Valid Key (11B) by 33⅓ percent, raising it from .21 to .28.

In Population B, the Suppressor Key (12A) and the Most Valid Key (10A) had a combined cross validity of .22, exactly the same as that computed for the rights key alone. The combination of the same Suppressor Key with the All Valid Key (11A) also had a cross validity of .22--in this case 10 percent higher than the All Valid Key alone.

It appears from the above that the Suppressor Key developed on Population B (12B) was the more effective in increasing the cross-validity of the rights keys as far as these populations are concerned. Examination of the beta weights in Table 1 indicates that this key also contributed more to the combined correlation than did the Suppressor key developed on Population A. The latter key (12A) appears to have been of little value in increasing the validities of the rights keys either singly or in combination in these populations.

It is interesting to note that while Keys 12A and 12B were designed to give maximum prediction of Keys 10A and 10B respectively, they were found to have considerably higher correlation with Keys 11A and 11B. A similar phenomenon shows in the intercorrelations of the suppressor keys which are .87 in population A and .89 in population B, although the keys which they were designed to predict (10A and 10B) intercorrelate only .45 and .47. It might also be noted that the extreme p-value keys (11A and 11B) intercorrelate .66 and .67.

It seems reasonable to infer that in selecting suppressor items by total score correlations to predict Keys 10A and 10B, that content more general than that of these two keys was included. This is especially the case since the suppressor items were also chosen for their low correlation with the criterion.

All of the above findings can be explained if the presence of a general distortion factor is assumed--a factor consisting of test score variance which is not correlated with the criterion but is highly intercorrelated among the items. The higher correlations between Keys 11A and 11B (which included the items of 10A and 10B) and between the Suppressor Keys and Keys 11A and 11B follow since it is known that, by and large, items with extreme p-values involve more distortion than those with median p-values.

Many of the increases in cross validity appearing after the Suppressor Keys are added to the rights keys do not represent statistically significant differences. It might, therefore, be felt that comparisons of proportionate changes are unwarranted. Nevertheless, in view of the interest in compiling data on the usefulness of suppressor keys in different situations, a comparison of the results obtained with such keys in the present study and in an earlier study is presented in Table 3 below.

Table 3. Absolute and proportional gains in validity resulting from the addition of suppressor key scores to scores on item analysis keys.

(Presented in order of decreasing intrinsic validity)

Key	Cross-Validities	Cross-Validity When Combined With Suppressor Key	Net Gain	Proportional Gain
API 10A+10B	.26	.32	.06	23%
API 10A	.22	.22	.00	0%
API 10B	.22	.28	.06	27%
API 11B	.21	.28	.07	33%
API 11A	.20	.22	.02	10%
W.P.S.D.B.*	.16	.21	.05	31%

*West Point Self-Description Blank.(3)

Although only one of the API-1 Suppressor-valid key combinations produced a proportional increase in cross validity comparable to that found with the West Point Self Description Blank, all but one of the API-1 keys did show some increase when combined with a suppressor key. The absolute gain obtained for the two Most Valid Keys and the two Suppressor Keys combined was enough to raise the cross validity of this API-1 scoring combination to a level which gives it promise of usefulness for further research.

In the second phase of this study, the correlations and scatterplots obtained from the data on education and AGCT Score showed no evidence of curvilinearity. Therefore, on the basis of the available data, it can be assumed that the relationship between these two variables and API-1 scores is essentially linear.

CONCLUSIONS

1. Two scoring keys consisting of API-1 items selected for their "suppressor" characteristics increased the combined validity of the two Most Valid scoring keys to a point which gave the API-1 considerable promise of usefulness for future research. The proportional increase for this combination was slightly less than the increases found for certain single keys.

2. When the effects of the two suppressor keys on the various scoring keys were considered separately, only one of the two appeared to be effective in materially increasing the validity of the scoring keys for these populations. It is possible, however, that with different populations both keys might have different effects.

3. No clear evidence of curvilinearity was found in the relationship between the Education-AGCT Score combination and the API-1 Scores.

PERSONNEL

Program Coordinator: Lindsey R. Harmon

Project Director: Hubert E. Brogden

Statistical Adviser: Laverne K. Burke

Research Associate: Walter A. Klieger

Statistical Associate: Claire T. Machlin

COLLECTION OF DATA: May, 1947

PREPARATION OF REPORT: 15 February 1952

REFERENCES

1. BROGDEN, HUBERT E., On the interpretation of the correlation coefficient as a measure of predictive efficiency. J. educ. Psychol., 1946, 65-76.

Reports of the Personnel Research Section, Personnel Research and Procedures Branch, The Adjutant General's Office, Department of the Army.

2. PRS Report 865, The Army Personality Inventory, 14 July 1952.

3. PRS Report 872, Validation of the West Point Self Description Blank, 17 November 1950.

4. PRS Report 879, A Simplified Vocabulary Revision of the Army Personality Inventory, 2 January 1951.

APPENDIX A

Table A-1. Biserial correlations between various API-1 scores and discharge. Means, standard deviations, and intercorrelations of keys.

(Population A; N = 1807)

Mean	Standard Deviation	Description of Variable	Variable Number	1	2	3	4	5	6
32.2	4.0	Score on Key 10A	1						
33.1	5.4	Score on Key 10B	2	.45					
66.3	6.6	Score on Key 11A	3	.88	.56				
49.9	7.0	Score on Key 11B	4	.49	.96	.66			
6.7	4.9	Score on Key 12A	5	-.35	-.67	-.55	-.73		
8.8	5.6	Score on Key 12B	6	-.30	-.66	-.48	-.70	.87	
		Unfavorable-Favorable Discharge Criterion	7	.40	.22	.36	.21	.02	-.01

Table A-2. Biserial correlations between various API-1 scores and discharge. Means, standard deviations, and intercorrelations of keys.

(Population B; N = 1469)

Mean	Standard Deviation	Description of Variable	Variable Number	1	2	3	4	5	6
31.5	3.9	Score on Key 10A	1						
32.2	5.7	Score on Key 10B	2	.47					
65.4	6.7	Score on Key 11A	3	.87	.57				
49.0	7.3	Score on Key 11B	4	.52	.97	.67			
6.6	5.1	Score on Key 12A	5	-.38	-.67	-.59	-.73		
8.5	5.9	Score on Key 12B	6	-.35	-.66	-.53	-.70	.89	
		Unfavorable-Favorable Discharge Criterion	7	.22	.38	.20	.35	-.04	.00