

1946

The Form of Items and the Distribution of False Positive Scores in a Neurotic Inventory

Wm. A. Owens
State College of Iowa

Copyright © Copyright 1946 by the Iowa Academy of Science, Inc.
Follow this and additional works at: <https://scholarworks.uni.edu/pias>

Recommended Citation

Owens, Wm. A. (1946) "The Form of Items and the Distribution of False Positive Scores in a Neurotic Inventory," *Proceedings of the Iowa Academy of Science*: Vol. 53: No. 1 , Article 37.
Available at: <https://scholarworks.uni.edu/pias/vol53/iss1/37>

This Research is brought to you for free and open access by UNI ScholarWorks. It has been accepted for inclusion in Proceedings of the Iowa Academy of Science by an authorized editor of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.

THE FORM OF ITEMS AND THE DISTRIBUTION OF FALSE POSITIVE SCORES ON A NEUROTIC INVENTORY

WM. A. OWENS

I. History.

During July, 1945, the Test and Research Section of The Bureau of Naval Personnel was directed to construct a pre-psychiatric screening device which would function to separate neurotics in general and "combat fatigue" cases in particular from the total populations passing through Receiving Stations.

With this end in view, four experimental measuring instruments were taken to the Receiving Station, Treasure Island, San Francisco, where the four approaches represented were tentatively evaluated and the most promising one was subjected to more intensive study.

The questionnaire adopted and called the Experience Comparison Index (E.C.I.) involved the symptomatic approach and allowed a "yes or no" response. In its original form it had 65 items, and after two item analyses and two revisions it had 30 items.

II. Problem.

It was found at Treasure Island that approximately 10% of the men passing through the Receiving Station could routinely be seen by the two psychiatrists present. A heavier load than this allowed the psychiatrists, who did a splendid job, to gain only a fleeting impression which was diagnostically unsatisfactory.

When scores on the Experience Comparison Index became the basis for referral, a specific problem, which follows, obtruded itself. The partial distribution of scores made by the men diagnosed as maladjusted was, as desired, sharply negatively skewed. However, the tail of the distribution of scores obtained by the undiagnosed or "normal" men referred was not curtailed just above the cutting score as it had been hoped it would be. On the contrary, these "false positives" reached the ceiling on the questionnaire—30 symptomatic responses out of 30. The problem faced then seemed to be to force the aforementioned false positives down out of the upper ranges of the distribution.

III. Method.

Since two item analyses with attendant reductions in the number of items plus numerous attempts at rewording had failed to accomplish this objective, it was decided to change the format of the questionnaire entirely.

The paramount change made was to substitute a pair of statements, on the same continuum, for the single statement with a "yes or no" answer. The subject then was only required to affirm one statement of the pair. This new dual statement form was designated the Personal Check List.

Specifically and in detail it differed from The Experience Comparison Index as follows:

- (1) The directions and the context were slightly altered. The E.C.I. statements had read "*These men are* fairly even tempered," and the directions had in effect said, are you like them—"yes or no"? In the P.C.L. this statement appeared as "*I am* fairly even tempered", and the paired statement read, "I lose my temper a lot oftener than I used to."
- (2) In making pairs of statements these actual changes occurred: Sixteen of the original statements were left intact and the only alteration was in the addition of a paired statement. Four of the original statements were qualified by the addition of one or two words and a paired statement was then added. Eight of the original statements were paired up to form four new items. Six of the statements were so altered as to be properly considered new content.

IV. Results.

When the new P.C.L. was administered to the same type of Receiving Station population that had taken the E.C.I., a change was noted in the distribution of the false positive scores. This change is evident in what follows.

First, taking each questionnaire over four separate periods of one day each, there was a drop in the mean percent of undiagnosed cases referred to psychiatry when the P.C.L. was used. This drop was from 54.25 percent to 41.75 percent and was significant at the 5% level. The inverse sign transformation was applied to each of the four pairs of percents, and the "t" test was employed to obtain an estimate of the significance of the difference between the means of the two series.

Second, separately plotting the partial distributions of diagnosed and undiagnosed cases revealed that the apex of the "false positive" distribution which had been at 30 had shifted downward to 22, whereas the apex of the distribution of diagnosed cases still lay at 29, one point lower than previously.

Since the representativeness of this result may be questioned, two facts seem worthy of note. First, that the E.C.I. was independently discarded in another investigation because the false positive rate was too high. Second, that administration of a revised P.C.L. to 1598 men revealed *no* cases in which the highest false positive score was within 6 points of the highest true positive score.

Finally, since the effect of altering a statement or of adding new

material is doubtful, it seems important to mention that the discrimination and validity of the twenty items which were *not* so dealt with was superior to that of the remaining ten. This would tend to indicate that the superiority of the P.C.L. was *not* due to improved wording or content.

V. Discussion.

The following hypotheses are offered in explanation of these results.

- (1) It is possible that the context change from "These men" to "I am" influenced the distribution of false positive scores. This seems improbable in that such a change would be unlikely to produce a *differential* effect. Moreover, diagnosed and undiagnosed men alike, when interviewed, seemed equally aware that they were making self-ratings in responding to either form of statement.
- (2) It is possible that rewording or minor content changes were responsible for the improvement in discrimination. This also seems unlikely in that two major revisions in the old format had not produced this result, and in that the evidence from item analyses previously mentioned does not suggest it.
- (3) It seems more likely that the improved discrimination resulting from a lower incidence of false positives was conditioned by the change from single statement to paired statement items. These latter seem to the writer to possess two distinct advantages over the former.

First, they are better adapted for use in a test the function of which is to isolate a segment-like 10% of a population. This follows from the fact that a "yes-no" form of response would appear to suggest a central cleavage in the continuum, whereas paired statements may represent adjacent positions on a scale. For example—"I feel tired and played out most of the time now"—yes or no—a negative response to this statement would seem to indicate the absence of any conspicuous feeling of fatigue. Few men who had spent a considerable period at sea could honestly answer in this fashion, and the item in this form actually did *not* differentiate sharply. However, if the paired statement—"I get tired a little easier now, but it never bothers in my work"—is added, many normals could select this alternative and the discrimination of the item would be improved—as, in fact, it was.

Apparently a truism in test construction is involved—*It is most efficient to divide a population on the items into the same proportions as it is to be divided on the total test.*

Second, a negative response to an asymptomatic statement in "yes-no" form has many potential meanings, and the conditions suggested are undoubtedly of *unequal* value in the differentiation of groups. For example, "I find it easy to concentrate—"yes or no".

A denial here might indicate fatigue, home-sickness, boredom, a toxic condition, daydreaming or what you will; and, accordingly, the item did not work best in this form. However, when the paired statement, "I often find myself daydreaming", was added, the item was bettered. Apparently this second statement is related to one of the *more* differential conditions suggested in denying the first statement, and thereby definitively improved the operation of the item.

As an extension of this second point, it could be suggested that the use of only symptomatic "yes-no" statements might result in as great effectiveness as the use of paired statements. Aside from the previously mentioned "distribution hypothesis," it would seem that such a technic would be obvious enough to encourage malingering, and that the single statements unless narrowly factual would still tend to lack the clarity and definitiveness of the paired statements.

VI. Conclusions.

(1) In the construction of this particular neurotic inventory it was found that a shift of a format from single statement, "yes-no" items to paired statement items was associated with a decrease in the incidence of false positive scores.

(2) Since these data were obtained incidental to the solution of a practical problem and *not* experimentally, the necessity for further research along this line is strongly indicated.

STATE COLLEGE OF IOWA.