*Multivariate Experimental Clinical Research, Volume 4, Number 4, 1979, 125-131*

# THE DEVELOPMENT OF A VALIDITY SCALE FOR THE CLINICAL ANALYSIS QUESTIONNAIRE

Samuel E. Krug
Institute for Personality and Ability Testing

## ABSTRACT

Following the procedures similar to those used by the MMPI authors in the development of the *F* (validity) scale, a self-contained validity index was developed from the item pool of the *Clinical Analysis Questionnaire* (CAQ). Ten (10) items were found to have very infrequently endorsed alternatives in a derivation sample of 300 normal college students. These were checked in an independent sample of 200 college students and in a sample of 200 clinically diagnosed adults, all of whom were tested as part of the test's standardization. The resulting scale effectively differentiated acceptable from randomly generated profiles with a satisfactorily high level of validity (.61) and classification accuracy (83%), yet did not differentiate normal from clinical profiles, as is frequently the case with *F*. Because of its brevity and convenience, it is likely to become an important element of the CAQ profile and provide the examiner with valuable evidence on test-taking attitude.

## INTRODUCTION

In 1971, Delhees and Cattell produced an important new instrument, the *Clinical Analysis Questionnaire* (CAQ). The test has 28 scales, 16 of which are the normal personality traits previously included in Cattell's *16 Personality Factor Questionnaire* (16 PF), seven of which measure various primary manifestations of depression, and five of which were developed to measure factor-analytically identified clinical traits discovered in the MMPI item pool. Consequently, the questionnaire provides evidence of pathology along with information on personality structure, thereby serving both diagnostic and treatment-planning functions. Unfortunately, the questionnaire was developed without a validity scale, no doubt due in part to Cattell's contention that such scales represent only a "simplified compromise with the real complexity of the mechanisms involved" (Cattell, Eber, & Tatsuoka, 1970, p. 55). His position has been that the issue of distortion in personality measurement needs to be considered from the broader perspective of "perturbation theory" (Cattell, 1968, 1977), which proposes that the difference between an individual's true and obtained trait score is partly a function of other personality characteristics as well as certain role factors. The basic model has been empirically tested and found to be valid (Krug & Cattell, 1971). However, the approach has a number of important practical drawbacks. For example, the extent to which the individual has adopted influencing roles must be assessed at the same time as the personality variables. This requires additional testing time and may frequently involve the examiner in test construction if no appropriate role measure is available. Furthermore, the necessary calculations assume the availability of electronic computing facilities. Rather than a "simplified compromise," then, self-contained validity scales represent a practical compromise for most psychological examiners.

The MMPI's *F* scale is one early attempt at a self-contained validity indicator. The 64 items that constitute this scale were endorsed in the keyed direction by no more than 10% of the normative group. This led MMPI developers to conclude that elevated *F* scores represented serious departures from normal test-taking attitudes and indicated an invalid profile (Dahlstrom & Welsh, 1960). Since the criterion for item selection was infrequency in a relatively normal population, perhaps it was not too surprising to find markedly elevated *F* scores accompanying serious elevation on the clinical scales as well, notably Scales 6 (Pa) and 8 (Sc). Butcher (1969) has noted that *F* correlates with severity of illness in a clinical population and that *T* scores as high as 80 do not necessarily indicate an invalid profile. Gynther (1961) has also pointed out the clinical relationships of elevated *F* scores.

This approach to the development of a validity scale is not unique to the MMPI and has been followed with minor variations elsewhere, fairly recently by O'Dell (1971) in constructing a "random" scale for the 16 PF. Judging by the content of the 31 items O'Dell selected, it seems that this scale is likely to suffer from similar problems. More than half of the items fall on factors known to be a part of the 16 PF second-order anxiety pattern or on the intelligence factor. Consequently, people might score high on O'Dell's index because they are anxious, because they are not able to concentrate sufficiently on those items that require a correct answer, because of a combination of these two factors as in clinical depression, or because the items really were answered carelessly.

Still, the method by which the MMPI *F* scale and the 16 PF random scale were developed is appealing. It seemed that with certain modifications to the basic procedure a self-contained scale could be developed for the CAQ which would reliably differentiate valid protocols from those which could be generated by a chance response pattern, yet not constitute simply a pathology scale that would differentiate essentially normal from clinical profiles.

## METHOD

Although the CAQ contains 272 items, the initial selection was limited to those 120 that were not part of the depression, pathology, or intelligence factors. The items excluded by this criterion were those likely to show unusual endorsement patterns among normal subjects but primarily because they deal with clinically relevant content. The response patterns of 300 male and female undergraduates, sampled in roughly equal proportions—henceforth designated the derivation group—were obtained under normal testing conditions. Subjects averaged 19 years of age. All were enrolled in a psychology course at the time of testing, but not all were psychology majors. Testing took place at three public universities located in Illinois, Florida, and California. Fourteen (14) items satisfied the initial condition: endorsement frequencies of 10% or less. Parenthetically, this selection ratio (14/120 items) is almost precisely equivalent to that encountered in the construction of the *F* scale (64/550).

Next, a random sample of 200 clinically diagnosed adults was drawn from nearly 2,000 protocols that had been obtained during the development and standardization of the CAQ. This group of men and women, sampled in equal

proportions, averaged 34 years of age (range 15-71). Subjects were tested at approximately 60 locations throughout the United States and Canada. On the whole, this was a slightly more urban sample and had slightly lower minority representation than the college students, but neither difference was of major significance. These were to serve as a comparison group so that items which provided significant clinical/normal differentiation could be identified. Rather than use the derivation sample and run the risk of confounding the results with regression effects, a second, independent sample of 100 undergraduate males and 100 undergraduate females was randomly drawn from among 888 students who had participated in the CAQ standardization and used as the contrast group. Again, this group averaged 19 years old. More than 5% of the variation in four of the 14 items was related to the clinical-normal dichotomy, and these were eliminated from further consideration. The means for the three groups on each of the remaining 10 items are reported in Table 1.

**Table 1**

**ITEM COMPOSITION AND MEANS ON THE CAQ V SCALE**

| Item | Keyed Alternative | Means | | |
|---|---|---|---|---|
| | | Deviation Sample | Clinical Sample | College Sample |
| 4 | B | .060 | .050 | .145 |
| 18 | B | .073 | .060 | .080 |
| 24 | B | .090 | .085 | .180 |
| 39 | B | .040 | .110 | .080 |
| 67 | B | .097 | .130 | .095 |
| 75 | B | .080 | .065 | .070 |
| 81 | B | .097 | .105 | .105 |
| 90 | C | .083 | .155 | .110 |
| 110 | C | .070 | .050 | .070 |
| 120 | A | .077 | .080 | .110 |
| Scale Mean | | .767 | .875 | 1.045 |

*Note:* In contrast with the trait scales of the CAQ, the V scale items scored dichotomously.

Next, 200 protocols were obtained by using a computer to generate random response patterns. Scores on the 10-item validity (V) scale were calculated and comparisons made among the clinical, college, and random samples. In both clinical and college samples, the modal score on V was 0. Using a cutting score of 2 on the scale, i.e., classifying protocols with a score of 3 or higher as valid, appeared to be optimum and correctly classified 83% of the 600 profiles. A test for the validity of the classification resulted in $\chi^2$ of 222, which was significant far beyond all conventional levels. Reexpressing these findings in terms of the strength of association rather than hit rate, the V scale is found to have a validity of .61 against the desired criterion. Using this rule, only 9.8% of the clinical and college cases were incorrectly identified as invalid. Depending on which type of classification error the examiner wishes to minimize, other cutoffs may be more appropriate. For this reason, distributions are reported in Table 2.

Table 2

DISTRIBUTION OF SCORES ON THE V SCALE

| Score | Percentage of Sample Obtaining Score | | |
|---|---|---|---|
| | Clinical | College | Random |
| 0 | 49 | 41 | 1 |
| 1 | 28 | 33 | 7 |
| 2 | 14 | 16 | 23 |
| 3 | 5 | 6 | 24 |
| 4 | 2 | 3 | 27 |
| 5 | 2 | 1 | 13 |
| 6 | — | — | 3 |
| 7 | — | — | 2 |

Correlations between V and the 28 trait scales of the CAQ were calculated. Items which are part of V were not scored on their usual trait scale, to eliminate spurious correlation based on overlap of items. These values are shown in Table 3.

A final check was made on the reliability of this 10-item V scale, which was found to be .44 for the sample of 200 clinical cases as calculated by Kuder-Richardson Formula 20.

## DISCUSSION

The primary aim of the research—to construct a scale from the item pool of the *Clinical Analysis Questionnaire* (CAQ) that would effectively differentiate random from normal profiles but that was not directly tied to pathology—seems essentially satisfied. The scale has 10 items which are very infrequently endorsed in the keyed direction by either normal or clinical subjects. Seven of the 10 items are keyed for the "in between" or "uncertain" response. Only one of the items is taken from a scale that contributes importantly to the anxiety factor. Three come from Factor A, warmth. For V, however, one is keyed in the A+ direction, one in the A— direction, and one is neutral. For five of the items (4, 24, 67, 90, 93), the other two response alternatives are endorsed with roughly equal frequency. On the remaining five items, usually only one non-keyed alternative is heavily endorsed. Thus, the scale incorporates items that have two "attractive" response alternatives as well as those that have only one. Because the items themselves are protected by copyright, they cannot be reproduced here. However, in general, higher scores indicate that the individual is uncertain about the type of material he prefers to read, what kind of animal he prefers, whether he likes to work alone, or which of two jobs he would enjoy more. On the other hand, he is definitely too busy to do charity work and would definitely not like to be a lion tamer.

Despite the fact that no attempt was made to develop a homogeneous scale, and considering its brevity, the internal consistency of V is reasonable. When adjusted by a Spearman-Brown formula to the 64-item length of *F*, the reliability jumps to .83, which is quite close to what has been reported for *F* (Dahlstrom & Welsh, 1960).

**Table 3**

**CORRELATIONS OF V WITH CAQ PRIMARY FACTORS**

| | Primary Trait | Clinical Sample | College Sample |
|---|---|---|---|
| | ***Normal Primaries*** | | |
| A | Warm-hearted | −07 | −09 |
| B | Intelligent | −10 | −30** |
| C | Emotionally stable | −03 | 02 |
| E | Assertive | −06 | −03 |
| F | Enthusiastic | −07 | −07 |
| G | Conscientious | −17* | −31** |
| H | Venturesome | −13 | −11 |
| I | Sensitive | 11 | −06 |
| L | Suspicious | −07 | −10 |
| M | Imaginative | 06 | −02 |
| N | Shrewd | −03 | −08 |
| O | Worried | 07 | −01 |
| $Q_1$ | Experimenting | 10 | 04 |
| $Q_2$ | Self-sufficient | 24** | 04 |
| $Q_3$ | Controlled | −16 | −07 |
| $Q_4$ | Tense | 09 | 03 |
| | ***Depression Primaries*** | | |
| $D_1$ | Hypochondriasis | 16* | 20** |
| $D_2$ | Suicidal depression | 17* | 24** |
| $D_3$ | Agitated depression | 04 | 05 |
| $D_4$ | Anxious depression | 10 | 04 |
| $D_5$ | Low energy | 14* | 06 |
| $D_6$ | Guilt and resentment | 03 | 08 |
| $D_7$ | Bored depression | 15 | 39** |
| | ***Clinical Primaries*** | | |
| Pa | Paranoia | 21** | 28** |
| Pp | Psychopathic deviation | −17 | −05 |
| Sc | Schizophrenia | 20** | 26** |
| As | Asthenia | 06 | 07 |
| Ps | General psychosis | 09 | 28** |

* $p < .05$

** $p < .01$

*Note:* Decimal points have been omitted.

Using an optimum cutoff score, 83% of all profiles were correctly classified, and a validity coefficient of .61 was reported. In one sense this should be considered cross-validation evidence on the scale, since the sample used in initial item selection was not used in subsequent research on its classification effectiveness. In terms of how scores relate to pathology, Table 1 shows that the clinical group actually scored lower than the college sample on V, though not significantly lower.

In terms of the relation of the validity scores to the primaries on the questionnaire, it is found to correlate negatively with both B (intelligence) and G (conscientiousness) in the college group. Both correlations are in the same direction for the clinical group, but only that for G is statistically significant. This relationship to intelligence, one of the highest found, is surprising in light of the fact that V contains no intelligence items, but is consistent with the aim of developing a scale that would measure the examinee's tendency to answer unreliably and inconsistently. The correlation of .39 with the CAQ Bored Depression scale is also consistent with this interpretation.

When we look at the relationship of V to broader second-order factors that have been identified among the CAQ primaries (Krug & Laughlin, 1977), the significant correlations are —.26 with Extraversion, .25 with Depression, .25 with Psychoticism, and —.20 with Superego Strength. The patterns were highly congruent across clinical and normal subjects. In short, despite all efforts to eliminate pathological elements from V, there are indications that individuals who score higher on V are somewhat more disturbed than low-scoring individuals, though these tendencies are slight. Perhaps as a byproduct of measuring consistent item patterns V picks up some elements of the cognitive confusion very central to depressive disorders (Beck, 1967).

Nevertheless, based on the evidence reported here regarding its effectiveness and since it is short and convenient to score, the CAQ V scale is likely to become an important element of the test profile and provide the examiner with valuable evidence on test-taking attitude.

## REFERENCES

1. Beck, A. T. *Depression: Causes and treatment.* Philadelphia: University of Pennsylvania Press, 1967.
2. Butcher, J. N. *MMPI: Research developments and clinical applications.* Minneapolis: University of Minnesota Press, 1969.
3. Cattell, R. B. Trait-view theory of perturbations in ratings and self-ratings (L (BR- and Q-data)): Its applications to obtaining pure trait score estimates in questionnaires. *Psychological Review,* 1968, *75,* 96-113.
4. Cattell, R. B. A more sophisticated look at structure: Perturbation, sampling, role, and observer trait-view theories. In R. B. Cattell & R. M. Dreger (Eds), *Handbook of modern personality theory.* New York: Wiley, 1977.
5. Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. *Handbook for the 16 Personality Factor Questionnaire.* Champaign, Ill.: Institute for Personality and Ability Testing, 1970.

6. Dahlstrom, W. G., & Welsh, G. S. *An MMPI handbook: A guide to use in clinical practice and research.* Minneapolis: University of Minnesota Press, 1960.
7. Delhees, K. H., & Cattell, R. B. *Handbook for the Clinical Analysis Questionnaire.* Champaign, Ill.: Institute for Personality and Ability Testing, 1971.
8. Gynther, J. D. The clinical utility of "invalid" MMPI *F* scores. *Journal of Consulting Psychology,* 1961, *25,* 540-542.
9. Krug, S. E., & Cattell, R. B. A test of the trait-view theory of distortion in measurement of personality by questionnaire. *Educational and Psychological Measurement,* 1971, *31,* 721-734.
10. Krug, S. E., & Laughlin, J. E. *Handbook for the IPAT Depression Scale.* Champaign, Ill.: Institute for Personality and Ability Testing, 1976.
11. Krug, S. E., & Laughlin, J. E. Second-order factors among normal and pathological primary personality traits. *Journal of Consulting and Clinical Psychology,* 1977, *45,* 575-582.
12. O'Dell, J. W. Method for detecting random answers on personality questionnaires. *Journal of Applied Psychology,* 1971, *55,* 380-383.

*Request for reprints should be sent to:*

*Samuel E. Krug*
*1. P.A.T.*
*1602 Coronado Dr.*
*Champaign, IL 61820*