A Reply to the American Psychiatric Association Task Force Report on Megavitamin and Orthomolecular Therapy in Psychiatry: The HOD Test

Harold Kelm, Ph.D. 1

A Task Force of the American Psychiatric Association has issued a report titled Mega-vitamin and Orthomolecular Therapy in Psychiatry (1973), one-third of which is devoted to the Hoffer-Osmond Diagnostic Test (HOD). Pauling (1974) and Hoffer and Osmond (1976) have written detailed replies to the other two-thirds of the report, and it would seem that the section concerned with the HOD also deserves a review, especially since it contains errors and omissions.

The Task Force refers to data from only a few studies, some of which it reports in error, and then questions the reliability and validity of the HOD, making no reference whatever to most of the data concerned with reliability and validity. The following reply reveals some of these errors and refers briefly to data in some of the major studies of the HOD which the Task Force has failed to mention and which do not support its conclusions.

DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF SASKATCHEWAN, SASKATOON, SASKATCHEWAN S7N 0W0.

RELIABILITY

The Task Force questions the reliability of the HOD and asserts that the originators of the Test have failed "... to cross validate its reliability..." (p. 34). Data in two reports co-authored by one of the originators and not mentioned by the Task Force provide information on the reliability of the Test.

One study provides an estimate of reliability based upon two separate administrations: all the test-items and the Short Form scale (Kelm, Callbeck, and Hoffer, 1967). In two groups of patients the correlations between the whole test and this short form scale were .84 and .85, and .92 and .94 between this short scale administered alone and as part of the entire test. In another study the Test was administered on two separate occasions to 358 patients on a psychiatric ward of a general hospital, a second group of 25 psychiatric patients at another hospital, and to a small group of college students (Kelm, Hoffer, and Hall, 1967). The retest correlations for the Total Score of the Test were .88, .81, and .82 for these three
groups, respectively. Most of the correlations for three sub-scores were in the upper .70s and higher. The split-half approach in 1,143 psychiatric patients and 1,252 students yielded correlations ranging from .87 to .99. A recent study, unfortunately not available when the Task Force prepared its report, made a detailed examination of the reliability of the HOD scales in 882 psychiatric patients and found that the internal consistency coefficients ranged from .83 to .96 (Chattier, 1977).

These data indicate that the reliability of the HOD compares very well with and even surpasses other established psychological tests.

VALIDITY

Originators of the Test

The Task Force states that the originators of the HOD have not cross validated their Test. Again, published data show the Task Force to be wrong. For example, two studies of over 1,000 patients showed that the Test could make the originally claimed schizophrenic-nonschizophrenic discrimination (p values ranged from .014 to < .00003; Kelm and Hoffer, 1965, 1967). Other studies confirm the original report (Hoffer and Osmond, 1961) that the HOD has potential for "... making some prognostic estimate" (Hoffer and Osmond, p. 328, 1961; Kelm et al., 1968). Also, the developers of the HOD have provided detailed test-data on 1,644 psychiatric patients tested at four centers and on 1,252 students, showing among other things the different distributions of scores for schizophrenic and various non-schizophrenic subjects (Kelm, Hoffer, and Osmond, 1967). None of these data is mentioned by the Task Force.

Two "Negative" Studies

The Task Force apparently examined the data from two studies and concluded that they "... have led to conflicting results" (p. 33). Of the first they write: "One study (Stewart and Mahood) noted that the test scores were significantly different between schizophrenics, neurotics and character disorders but not between schizophrenics and other psychotics" (p. 33). The Task Force is again in error; no such findings are noted by these investigators (Stewart and Mahood, 1963). Where did the Task Force get such results? A later study which re-examined the data from Stewart and Mahood's patients showed that two of the HOD scales discriminated schizophrenic from nonschizophrenic patients (Perceptual p = .01, Ratio p = .0005; Kelm et al., 1965). In addition, a later item analysis of the Test resulted in the introduction of another scale (Short Form) which in Stewart and Mahood's sample also significantly differentiated between schizophrenic and other psychotic, schizophrenic and psychoneurotic, and schizophrenic and character-behavior disorder patients (p = .04, .002 and .02, respectively; Kelm et al., 1966). Thus, a study which the Task Force would have the reader believe does not support the diagnostic claims of the originators of the Test does, in fact, provide data strongly supporting the validity of the HOD.

The second study to which the Task Force refers (Craig, unpublished 1971; summary published by Craig, 1972) has several deficiencies. Firstly, without any explanation over 35 percent of the nonschizophrenic patients (four psychotic depressions, one manic depressive, and 14 neurotic depressions) were excluded in the analysis comparing schizophrenic and non-schizophrenic patients. From the unpublished report it is possible to calculate the mean HOD scores of this excluded group which are shown in Table 1. These data reveal that the 19 excluded patients score lower than both the schizophrenic and selected nonschizophrenic groups on all the HOD scales. No justification is given for excluding patients with the lowest HOD scores. Secondly, one of the most sensitive HOD scales for differentiating between schizophrenic and nonschizophrenic patients (Ratio Score; Kelm and Hoffer, 1965; Kelm et al., 1965) and which has relatively low correlations with the other scales (Kelm, Hoffer, and Osmond, 1967;
Kelm and Osmond, 1975), was also excluded in the analysis, without offering any explanation for this omission. Thirdly, in other studies of the HOD, including Stewart and Mahood's, nonparametric statistics were used because the parametric were not the most appropriate. Craig, however, used parametric statistics throughout his report.

It seems highly probable that if the 19 subjects excluded from the nonschizophrenic group and the Ratio Score were included using nonparametric statistics, the HOD would, as in other studies, be able to differentiate between the schizophrenic and nonschizophrenic patients in Craig's sample. Naturally, the best way to determine this would be to do such an analysis. Repeated attempts have been made to obtain a copy of the data so that such an analysis could be done, but it has been consistently withheld. The reader can judge why an investigator would violate scientific procedures by not using the lowest test scores obtained in an experiment, leave out one of the most sensitive scales in a test, fail to replicate statistical analyses, and then not permit another investigator to check his data. Despite these deficiencies, Craig found that both Wittenborn Psychiatric Scales and HOD scales differentiated between delusional and nondelusional patients, and concluded that "The HOD was a valid measure of delusional thinking" (p. 8). No patients, however, were excluded in the analysis of the data which led to this conclusion.

### Other Studies Not Mentioned by the Task Force

Sugerman and Williams (1965) reported changes in HOD scores in response to treatment (p values ranged from .02 to .001). Another study based upon 218 psychiatric patients found that the HOD again differentiated between schizophrenic and nonschizophrenic patients (most p values < .0003; Kelm et al., 1965). Still another investigation analyzing each of the 145 items of the Test on 818 psychiatric patients resulted in the selection of 17 items designed to more sharply discriminate between schizophrenic and nonschizophrenic subjects (Kelm et al., 1966). This group of 17 items, called the Short Form scale, was cross validated on two mental hospital samples, one of them including Stewart and Mahood's patients referred to earlier. In both samples this score was found to differentiate between schizophrenic and various nonschizophrenic patients (p values ranged from .02 to .00003). Yet another study reported similar results (p < .005; Silzer et al., 1972).

The validity of the HOD may also be examined by determining its relationship to other established tests which are known to differentiate between schizophrenic and nonschizophrenic subjects. One such test is the visual figural after-effect (FAE) which is also an accurate laboratory measure of perceptual distortion (Kelm, 1962; 1968; Wertheimer, 1954; Wertheimer and Jackson, 1957). In a study of the HOD and FAE in a

### Table 1

Mean HOD Scores of Craig's Schizophrenic (N = 31), Selected Nonschizophrenic (N = 32), and Excluded Nonschizophrenic (N = 19) Patients

<table>
<thead>
<tr>
<th>HOD</th>
<th>Schizophrenic</th>
<th>Selected Nonschizophrenic</th>
<th>Excluded Nonschizophrenic</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>36.2</td>
<td>31.8</td>
<td>23.0</td>
</tr>
<tr>
<td>PerS</td>
<td>6.8</td>
<td>4.6</td>
<td>3.3</td>
</tr>
<tr>
<td>PS</td>
<td>2.4</td>
<td>2.2</td>
<td>1.8</td>
</tr>
<tr>
<td>DS</td>
<td>6.4</td>
<td>7.2</td>
<td>5.8</td>
</tr>
<tr>
<td>SF</td>
<td>2.4</td>
<td>1.8</td>
<td>1.1</td>
</tr>
</tbody>
</table>
group of psychiatric patients it was found that there were significant correlations between these two tests (p < .05 to < .025; Kelm and Hall, 1967). Listing the primary and secondary symptoms of schizophrenia, the Task Force asserts that "Psychological tests such as the MMPI or the Wittenborn Scaies for the diagnosis of schizophrenia take these symptoms into consideration" (pp. 23-24). It is indeed interesting that several studies have consistently reported significant relationships between the MMPI and HOD (Chartier, 1977; Hawkins and Neziroglu, 1975; von Hilsheimer et al., 1977; Milner et al., 1973; Neziroglu, 1975; Njaa, 1972). Unfortunately, all of these data were not available when the Task Force prepared its report. The combined use of the Wittenborn and HOD was mentioned earlier.

Finally, one of the most sophisticated studies of the HOD again showed that it was able to discriminate significantly between schizophrenic and nonschizophrenic patients (Silzer et al., 1972). These investigators also compared the HOD with the Inpatient Multidimensional Psychiatric Scale (IMPS) and reported that their "Canonical correlation indicated that the two tests appeared to reflect some similarities in dimensions of psychopathology" (p. 359). These researchers also factored the two tests and found that both were well represented in the first two factors which accounted for 55 percent of the total variance. They also suggested that using the HOD and IMPS together would be more powerful than using only the IMPS.

Other studies cross validating the HOD and supporting the initial claims made by the originators of the Test may be cited; however, the above will suffice to show the nature and extent of the Task Force's examination of the available data.

CONCLUSIONS

Although the Task Force devotes approximately one-third of its report to the HOD, it refers to data from only three studies, two of which actually support the Test, while the third may show similar results if proper scientific procedures are followed, and concludes that the inventory lacks sufficient reliability and validity. No reference is made whatever to most of the data concerned with reliability and validity, giving the reader the impression that such data do not exist. Also, the Task Force's summary of the data from the only published "negative" study mentioned in its report is in error.

From its partial review and even incorrect summary of available data, The Task Force arrives at some erroneous conclusions which, whether by design or not, may effectively mislead the reader.

It is apparent that serious errors and omissions are not confined to the one-third of the Task Force report devoted to the HOD, but are also found in the remaining two-thirds devoted largely to niacin therapy as Pauling (1974) and Hoffer and Osmond (1976) have shown.

REFERENCES


NEZIROGLU, F.: The relationships among the Hoffer-Osmond Diagnostic Test, the Minnesota Multiphasic Personality Inventory, and independent clinical diagnoses. Journal of Clinical Psychology 31,430-433,1975.


The Journal of Orthomolecular Psychiatry is an educational journal published quarterly by the Academy of Orthomolecular Psychiatry as a service to the public. Certain classes of memberships to the Canadian Schizophrenia Foundation and the Huxley Institute for Biosocial Research include subscriptions to the Journal and a quarterly Newsletter. To become a member, write to the Canadian Schizophrenia Foundation, 2231 Broad Street, Regina, Saskatchewan, Canada S4P 1Y7, or to the Huxley Institute, 1114 First Avenue, New York, N.Y. 10021. Membership contributions are income tax deductible. Members of the Academy will receive the Journal.

Annual subscription rate for the Journal of Orthomolecular Psychiatry to non-members is $22, and to libraries $30.

Copyright 1978 by the Academy of Orthomolecular Psychiatry. All rights reserved. Publisher's written permission must accompany papers published previously in another journal.

The Journal, its editors, and the publishers do not take responsibility for opinions expressed by the authors. The Journal board and editors do not necessarily agree with all the articles that are published.

Send correspondence relating to published issues and annual subscription orders to the Canadian Schizophrenia Foundation, 2231 Broad Street, Regina, Saskatchewan, Canada S4P 1Y7.