

# Scoring Errors in the Hoffer-Osmond Diagnostic Test

Joseph A. Mitchell<sup>1</sup>

## Abstract

*While creating a Microsoft Excel® spreadsheet to aid in scoring the results of the Hoffer-Osmond Diagnostic assessments gathered from a forensic population in Mid-Michigan, the author uncovered several numerical errors in the calculations of the published scoring. One error resulted from an error on the Key B scoring protocol itself, that leads to incorrect Key B, and subsequent total score results. The second problem is in the range-score potential of the ratio scores. The ratio score problem develops partly from the Key B protocol error, but also in part by an error reported for the range-score potentials as cited in the instrument's manual. The following paper discusses these errors in detail, the appropriate method to correct the errors, and includes a short discussion regarding previous assessment results obtained by the use of the instrument.*

## Introduction

The utility of the 145-item Hoffer-Osmond Diagnostic (HOD) has long been established. It is easy to administer, quick to score, and there are few psychological assessment instruments with the sensitivity and specificity of the HOD regarding schizophrenia related symptoms, or associated perceptual anomalies. The HOD provides the test administrator with a total score (TS), and five subscales that define an individual's perceptual score (PerS), paranoid score (PS), depression score (DS), ratio score (RS), and short form (SF) score. These scores are then compared to potential range scores as reported in the manual (Kelm, Hoffer & Osmond, 1981) to establish the extent to which the testee reports

experiencing the symptoms. Armed with this information, the doctor, therapist, or other mental health practitioner can then easily develop the proper treatment regimen to ameliorate the client's symptoms.

However, there are several problems related to those reported range scores that will be addressed here. The TS range as reported in the manual (p. 6) is impossible to achieve, even if the testee were to positively endorse every symptom in the test. This is due to the Key B template excluding test item 42. It also makes the range score of Key B impossible to achieve. Consequently, the RS range is also affected.

## The Key B Template Error

As stated above, Key B template excludes item 42, "I sometimes feel my stomach is dead." This item was initially intended to be included in the Key B template (A. Hoffer, personal communication, January 27, 2003), but appears to have been excluded in a publisher or printer's error, and remained undiscovered until now. The HOD's potential range for the TS is currently gleaned from adding the values of the two Key A scores ([26 items x five points] + [4 items x two points]), plus the Key B score (104 items x one point). The final equation for TS then would be:  $(130 + 8) + 104 = \text{Incorrect TS (242)}$

This would result in a possible or potential range for the TS of 0 to 242, as opposed to the manual's published range potential peak of 243.

This discrepancy led to the evaluation of scale loading for each of the items on the HOD. It was clear after this evaluation that all items on the test were included in the TS with the exception of items 83, 85, 87, 89, 91, 93, 95, 97, 99, and 101 (the logical connection items that were never intended to be included in the TS), and the

1. Department of Psychological Research and Counseling, Mitchell and Associates, PO Box 258, Midland, Michigan, 48640

mentioned missing item 42. Including item 42 in Key B, as was the test author's intent, fixes the Key B error, and thus the TS error.

Correcting the Key B template for HOD users is much easier than recalculating potential numerical ranges, and is a fairly simple process. Using a blank HOD test score sheet, overlay the PerS template, and beside item 42 place a light pencil mark on the template to signify item 42. Second, place the Pers template over the Key B template and gently trace the box signified as item 42 by the identifying mark. Using an Exacto knife or similar implement carefully cut out the traced chad. Now, with the inclusion of item 42 to Key B in subsequent assessments, the equation adds the potential TS correctly as zero to 243 reported by the manual:  $(130 + 8) + 105 = \text{Corrected TS (243)}$  However, this correction does little to fix the problem regarding potential ratio scores, which will be addressed in the next section.

### The Ratio Score Errors

Ratio scores (RS) for the HOD are calculated by, "dividing DS into TS, and when DS is zero, multiplying TS by 2" (Kelm, Hoffer & Osmond, 1981, p. 8). The RS subscale of the HOD is invaluable with regard to insight into the client's treatment progress, and helps the test administrator differentiate various psychiatric presentations. While it is rare to get a zero score on the RS it is possible to endorse none of the TS items on the HOD as true.

To determine the potential range of possible scores for this subscale, it must be understood that if every item was endorsed as true on the HOD TS except the 18 DS items (which are all Key B items, incidentally), or rather, the testee endorses a zero score for the DS subscale with all other TS items endorsed as true, those 18 DS items must be subtracted from the full potential TS. With those 18 items removed from the TS (to meet the zero criteria for the DS to double the TS), that leaves only 225 remain-

ing items to be doubled. Thus:

$$1.) \text{ TS (243) - DS (18) = 225;}$$

$$2.) \text{ TS (225) x 2 (DS = 0) = Corrected (RS peak) 450}$$

The manual cites a highest possibility of RS as 452 (p. 6), when in fact the actual potential high RS peak is 450. It would be extremely rare for that combination of endorsements to occur, but it is theoretically possible. Equally rare, but also possible, would be that no TS item would be endorsed. If that were the case, the lowest possible RS score would be zero. Even one endorsed TS item though results in an RS of one, and is doubled if that item is not a DS item. If the one endorsed TS item is a DS item, then one TS (DS) item divided by one DS still equals an RS of one. Each additionally endorsed DS item would only increase the denominator as well as the numerator, because the DS is a subsection of the TS, and summarily there is no possible way to get a fractionally lower RS except to endorse no HOD item as true (with the exception of items 83, 85, 87, 89, 91, 93, 95, 97, 99, and 101, which are not included in the TS or any other subscale). Subsequently, the possible valid RS range potential is zero to 450.

### Discussion

It should be mentioned that while these errors, and resultant numerical corrections, may have no great effect on previous assessment results, there could be instances of borderline cases wherein they may have a clinical impact. The author reviewed 40 randomly selected previous HOD results, and in no case did the current numerical changes affect either the client's diagnosis, or resulting treatment modality. However, a further encouragement to immediately incorporate these changes in ongoing HOD assessment would be in the event that the clinician is called to testify within the forensic arena regarding his or her clinical decisions based on HOD assessment. In that event, correct numerical structures must be reported to allow the clinician's testimony to be viewed as credible.

### Summary

Modifying the Key B template as described above corrects the publisher or printer's error of Key B, and that of the total score, bringing possible range scores of the TS in line with those reported in the HOD manual. The ratio score error is a different matter. Prior to correcting Key B, the reported ratio scores and actual potential scores were off by four points. But even after the Key B template correction, the manual and the actual potential scores are still off by two points.

If a revision of the HOD were ever con-

sidered to add, subtract, or modify items, make changes in certain statements to reflect current vernacular, or to add a dissimulation scale, these numerical factors should also be taken into serious consideration. With doctors and clinicians now being called to testify more often than ever before in forensic issues, a sound statistical basis of assessment is required.

### References

1. Kelm, H., Hoffer, A. & Osmond, H (1981). Hoffer-Osmond Diagnostic Test Manual: Revised Edition. Behavior Science Press: Tuscaloosa, AL.