

Alternative Scoring System of Repeated Items on the MMPI: *Caveat Emptor*

Robert E. McGrath
Fairleigh Dickinson University

W. Brian O'Malley
Department of Physical Medicine
Ohio State University

Jason R. Dura
Ohio State University

There are 16 items in the standard MMPI group forms which are included twice. It was found that a number of computerized scoring services use only the first occurrence of repeated items in scoring the MMPI scales, whereas the handscoring templates use an arbitrary combination of the first and second occurrence of these items. Comparison of these conventions suggested a potential for significant differences in score, particularly on the Sc scale. Scoring a set of 126 MMPIs of chronic pain patients by both of these scoring conventions revealed differences of up to 10 T points on the Sc scale. It is recommended that a single scoring convention for the MMPI be adopted by psychologists. For several reasons we suggest that only the first occurrence of repeated items should be used for scoring purposes. In the absence of a single convention, comparisons between Sc scores on different protocols should not be made without first insuring that the protocols were scored in the same manner.

According to Greene (1980), 16 items were originally included twice on the Minnesota Multiphasic Personality Inventory (MMPI) in order to facilitate machine scoring of the group booklet form. Unfortunately, scoring systems for the MMPI have been inconsistent in choosing which occurrence of a repeated item is to be scored. Eighteen years ago, Fowler and Coyle (1968) noted that differences in the item selected for scoring resulted in raw score discrepancies among the scoring services available at the time, leading Dahlstrom, Welsh, and Dahlstrom (1972) to advocate use of a single convention for scoring the MMPI.

Despite this suggestion, varying conventions continue to be in practice. The Form R scoring templates provided by the Psychological Corporation (Psych. Corp.) consistently use the first occurrence on scales, *L*, *Hs*, *Hy*, *Pd*, *Ma*, and *Si*. On scales *F*, *D*, *Pa*, *Pt*, and *Sc* the templates use the first occurrence for some

items and the second occurrence for others (there are no repeated items on scales *K* and *MF*).

After examination of the templates, a telephone survey was conducted of those eight firms that advertised computerized MMPI scoring in the September 1984 *Monitor*. Of the six who responded to the survey, five reported using only the first occurrence of items (Caldwell Report, Integrated Professional Systems, National Computer Systems, Precision People, and Psych. Lab.). Western Psychological Services reported using the Psych. Corp. scoring instead.

The potential impact of alternative scoring conventions on raw scores is magnified by inclusion of the repeated items on multiple scales. For example, Item 8 (repeated as Item 318) appears on five scales. A 1-point difference is possible on scale *F*. Two-point differences are possible on scales *D*, *Pa*, and *Pt*. The most serious discrepancies are associated with the *Sc* scale, where an 11-point difference is possible on the basis of scoring convention.

Given that the two widely used scoring systems could potentially result in significant scoring discrepancies, the present study addressed the extent to which this issue represents a confounding element in practice. A set of MMPIs was scored according to both the first occurrence and Psych. Corp. conventions, and the results from the two systems were compared.

METHOD

Subjects

During 1982, 128 individuals were administered the MMPI as part of an evaluation for participation in a program for the behavioral management of chronic pain. Of this group, two MMPIs were deleted due to errors in the computerized scoring, leaving 126 observations in the data set. This group consisted of 56 females and 70 males with a mean age of 43 years ($SD = 11.6$).

Procedure

The MMPI protocols were sent to National Computer Systems for scoring, using the first occurrence of the repeated items. Upon return they were then re-scored according to the Psych. Corp. convention. Difference scores were then computed between the raw scores without *K* correction for scales *F*, *D*, *Pa*, *Pt*, and *Sc*.

RESULTS AND DISCUSSION

Table 1 summarizes the findings for all five scales. The number of protocols on which a difference was found varied between 10 for scale *F* (8% of the data pool) and 64 for scale *Sc* (51%). In general, the effect was small. For 135 of the 154, observed raw scores are all positive. Given the manner in which the difference

TABLE 1
Observed Differences Between Psych. Corp. and
First Occurrence

| Scale | N | M | SD | Maximum Observed Difference |
|-------|----|-----|-------|--------------------------------|
| F | 10 | .20 | 1.033 | 1 |
| D | 25 | .12 | 1.013 | 1 |
| Pa | 29 | .69 | .806 | 2 |
| Pt | 26 | .23 | .992 | 1 |
| Sc | 64 | .61 | 1.590 | 5 |

scores were computed, this suggests that when a discrepancy occurs, the score using the first occurrence tends to be higher than the Psych. Corp. score. Respondents appear more willing to endorse items in a pathological direction on first contact than they are on the second.

For nine of the ten clinical scales one need not be overly concerned about scoring discrepancies. However, for scale Sc the difficulty becomes significant. Table 2 summarizes the pattern of results for the Sc scale. Observed raw score differences varied between 1 and 5. Table 2 provides the frequency of each observed value. In addition, the range of raw scores and T scores associated with each observed value is provided. Raw scores are based on the Psych. Corp. scoring convention, which, as previously noted, is likely to produce a lower score than the first occurrence convention. T score conversions are without K correction, using tables provided by Greene (1980).

Three cases (2.4% of the sample) demonstrated a raw score difference of 4 to 5 points. In one case this corresponds to a change in T score from 73 to 78. According to Lachar (1983), with a T score of 73 "it is often difficult to judge whether general alienation or blatant psychotic content is being endorsed" (p. 22). For a T score of 78 the interpretive statement becomes "most individuals who obtain such elevations show, at minimum, a schizoid mentation" (p. 22).

TABLE 2
Summary of Scoring Convention Differences for
the Sc Scale

| Observed Difference | N | Raw Score Range | T Score Range |
|------------------------|----|--------------------|------------------|
| 1 | 46 | 5-52 | 44-102 |
| 2 | 11 | 10-53 | 50-103 |
| 3 | 4 | 18-42 | 61-94 |
| 4 | 2 | 28-29 | 73-75 |
| 5 | 1 | 22 | 67 |

Note. Ranges are based on the Psych. Corp. convention rather than the first occurrence convention. T values were computed without K correction.

The present situation unnecessarily complicates the interpretation of scale Sc scores when alternative scoring systems are employed, such as when a research unit or clinic moves from template to computerized scoring. The psychologist who is overjoyed to surrender the burdensome task of scoring MMPIs by hand may be completely unaware that the discharge protocols of his or her schizophrenia patients are not strictly comparable to their intake profiles, or that the actuarial data used to interpret these protocols were derived from research utilizing a different scoring convention.

Given this state of affairs, two conclusions are appropriate. First, we must concur with Dahlstrom et al. (1972) on the need for a consistent MMPI scoring convention. The use of only the first occurrence of repeated items suggests itself for several reasons. It is the procedure most commonly implemented by scoring services, and that most easily implemented given the published lists of scale items. It seems much less arbitrary than the Psych. Corp. convention. In addition, the present results suggest that respondents admit pathology more readily on the first occurrence of items. Ultimately, it is hoped the second occurrence of the repeated items will be removed from the MMPI test booklet.

It is very possible that our first recommendation will have little more impact on standard scoring practices than when Fowler and Coyle (1968) and Dahlstrom et al. (1972) made the same point. Until such time as a common convention is agreed upon, it is incumbent upon the MMPI researcher to inform his or her audience of the scoring convention used in a project, particularly when the project explores the characteristics of the Sc scale. Both the researcher and the clinician also should limit comparisons of the Sc scale across protocols to MMPIs which have been scored by the same method.

REFERENCES

- Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. (1972). *An MMPI handbook: Vol. 1. Clinical interpretation* (rev. ed.). Minneapolis: University of Minnesota Press.
- Fowler, R. D., & Coyle, F. A., Jr. (1968). Scoring error on the MMPI. *Journal of Clinical Psychology*, 23, 68-69.
- Green, R. L. (1980). *The MMPI: An interpretive manual*. New York: Grune & Stratton.
- Lachar, D. (1983). *The MMPI: Clinical assessment and automative interpretation*. Los Angeles: Western Psychological Services.

Robert E. McGrath
Department of Psychology
Fairleigh Dickinson University
Teaneck, NJ 07666

Received June 17, 1985
Revised January 27, 1986