

## CONTEXTUAL EFFECTS ON THE MMPI<sup>1</sup>

JULIA E. PERKINS

AND

LEWIS R. GOLDBERG

*Dammasch State Hospital*

*University of Oregon*

This study was designed to determine in what ways and to what extent MMPI scores are influenced by the ordering of their items. 12 shortened forms of the MMPI were developed, using 4 different scales (paranoia, social desirability, physical health, and acquiescence) and 3 different patterns of ordering (all items appearing first, being scattered throughout, or appearing last in the test). Ss included 473 college students, 108 neurotics, and 54 paranoid schizophrenics. Each S was administered 2 of the test forms, each of which contained a different ordering schema of the same items, with a 1-week interval between test administrations. Findings indicated no significant contextual effects on any of the scales.

Departures from standard MMPI test administration procedures are probably quite common in clinical practice. Psychologists, when under work pressure or when evaluating illiterate individuals or patients with visual defects, may administer (or read) the items included in one or two scales rather than the entire MMPI. Some subjects are given the group form of the MMPI in which the items are always presented in the same order, while other subjects take the individual form of the test in which the items may be arranged in any order. Sometimes, subjects skip various MMPI items—perhaps a large number of items throughout the entire test—later to return and respond to the skipped items. Occasionally, subjects complete only a portion of the test due to fatigue or illness and finish the remainder after a lapse of time. Scores are typically treated as if they were identical to those obtained under other testing procedures.

Over 10 years ago Cronbach (1949) suggested the existence of "a habit or momentary set which causes the subject to earn a different score from the one he would earn if the same items were presented in a different form." It is possible that the *patterning* of items in a test matrix can elicit a temporary set which

can cause a subject to respond in a different way than he would if the same items were presented in a different fashion.

In the *Technical Recommendations for Psychological Tests and Diagnostic Techniques*, APA (1954), psychologists are cautioned concerning the influence of contextual effects on test performance:

When a short form of a test is prepared by reducing the number of items or organizing a portion of the test into a separate form, new evidence should be obtained and reported for that new form of the test. . . . This is especially important for inventories, where placing items in a new context might alter the person's responses. For example, the MMPI properly retains some items which were not scored in any key, because removing those items might alter the discriminating power of the items which were scored.

However, in spite of the warning, MMPI ordering modifications are frequent and few studies have been directed toward understanding the nature of possible contextual effects.

Dahlstrom and Welsh (1960) have reviewed the research relating to item-ordering effects within the MMPI. Two groups of studies were cited, those comparing scores on the individual versus the group forms of the test, and those comparing complete versus shortened versions of the test. In general, Dahlstrom and Welsh minimized the importance of contextual effects on the MMPI and stated,

It was the feeling of Hathaway and McKinley (1940) in proposing the card form for the MMPI that in this format, items could be added to or removed from the test without any marked effect

<sup>1</sup> This report is a summary of a PhD dissertation completed at the University of Oregon by the first author under the direction of the second author. See Perkins (1962) for a more detailed presentation of the procedures and findings. Part of the data analysis was supported by funds from National Science Foundation Grant G-25123 to the second author.

on replies to the other items. Subsequent empirical work has tended to bear them out.

Research, indeed, suggests that the *removal* of unscored MMPI items seems to have little effect on scale scores (Charen, 1954; Dana & Condry, 1963; Ferguson, 1946; Holzberg & Alessi, 1949; MacDonald, 1952b; Machover & Anderson, 1953; Olson, 1961; Siegel & Feldman, 1958). While Taylor (1953) and Bechtoldt (1953) offered some evidence for contextually induced differences when her Manifest Anxiety Scale was administered as part of the MMPI and when it was administered in a shortened form, at least six subsequent studies utilizing different forms of the MAS (Bendig, 1956; Feldman & Siegel, 1958; Lebo & Noblin, 1958; McCreary & Bendig, 1954; Siegel & Feldman, 1958; Voas, 1956) have failed to confirm her findings. Also, studies comparing scores on group and individual forms of the MMPI (e.g., Wiener, 1947) found no evidence for ordering effects. However, other studies (Cottle, 1950; MacDonald, 1952a; McQuary & Truax, 1952) report some differences obtained between card-form and booklet-form scores.

In summary, while the few MMPI studies indirectly concerned with contextual effects encourage the hope that such effects are rare in the MMPI item pool, research demonstrating marked contextual effects in other inventories (e.g., Ellis, 1947) should make us pause. Moreover, since every individual form of the MMPI has some relatively *unique* ordering pattern, studies comparing group and individual forms have only tangential relevance to the problem of uncovering specific contextual effects. Consequently, it is necessary to directly and systematically challenge the MMPI item pool in a manner most

closely approximating the usual divergences from standard administration practices. The present study sought to ascertain the influence on MMPI scores of different ordering patterns of MMPI items. Both college students and two samples of psychiatric patients were used as subjects in order to allow for the emergence of any possible population interaction with contextual effects.

#### METHOD

Twelve shortened forms of the MMPI, each containing 286 items, were developed for the purpose of testing the influence on test scores of different combinations of ordering schema and item content. Four different scales (Paranoia, Physical Health, Social Desirability, and Acquiescence Response Bias) were utilized in this study, each with three different ordering schema (all items in the scale appearing together at the beginning of the test, items scattered throughout the test, and all items in the scale appearing together at the end of the test). Each of the four scales was included in 3 of the 12 test forms, and each of these 3 forms contained a different schema for the ordering of the items in the scale.

Items from two of Comrey's factors, "Paranoia" and "Poor Physical Health," were selected to be used in two scales assumed to differentiate paranoid schizophrenics and neurotics from normal subjects. Comrey carried out a series of factor analyses of items from each of the MMPI scales, excluding *Mf* and *L* (Comrey, 1957a, 1957b, 1957c, 1958). The *PA* (Paranoia) Scale contains 26 items of paranoid content which were selected from Comrey's "Paranoia" Factor. All items contained in this scale are included either on the *Pa*, *Sc*, *F*, or *Pd* scales. The *PH* (Physical Health) Scale contained 26 items of physical health content, 17 of which were drawn directly from Comrey's "Poor Physical Health" Factor. The nine remaining items, all of physical health content but not contained in Comrey's Factor, were added to increase the number of items in the scale. All of the items are scored on either the *Hs*, *D*, or *Hy* scales.

Wiggins' *Sd* (Social Desirability) Scale (Wiggins, 1959) and *Rb* (Acquiescence Response Bias) Scale (Wiggins, 1962) were selected as scales differentiating subgroups of normals while not constructed to differentiate psychiatric patients from normal groups. Wiggins' *Sd* Scale was developed by contrasting the responses of 140 college undergraduates who took the MMPI according to the standard set of instructions with responses of 178 college undergraduates who were instructed to respond in a socially desirable manner. The 40 items which discriminated (at the .001 level) the social desirability role players from the controls make up the *Sd* scale. Wiggins' *Rb* Scale was constructed as a measure of acquiescence response bias, defined by Wiggins as a "tendency to agree to a wide variety

TABLE 1  
MEAN ITEM ENDORSEMENT PERCENTAGES  
FOR FOUR SCALES

Scale	Adult sample		College sample	
	Males	Females	Males	Females
<i>PA</i>	.11	.09	.04	.04
<i>PH</i>	.18	.25	.10	.15
<i>Sd</i>	.46	.46	.37	.40
<i>Rb</i>	.56	.55	.46	.46

TABLE 2  
EXPERIMENTAL DESIGN

PA scale		PH scale		Sd scale		Rb scale	
Admin- istration A	Admin- istration B	Admin- istration A	Admin- istration B	Admin- istration A	Admin- istration B	Admin- istration A	Admin- istration B
PA-1	PA-2	PH-1	PH-2	Sd-1	Sd-2	Rb-1	Rb-2
PA-1	PA-3	PH-1	PH-3	Sd-1	Sd-3	Rb-1	Rb-3
PA-2	PA-3	PH-2	PH-3	Sd-2	Sd-3	Rb-2	Rb-3

of statements when no issue seems at stake" (Wiggins, 1962). In the construction of the *Rb* scale, 84 items of high "controversiality" (items endorsed by 40 to 60% of the subjects) were found in a sample of 190 college students who took the MMPI under standard instructions. These items were rated on a social desirability scale by a different sample of 50 students. The 27 items which fell within the "neutral" interval (3.5 to 4.5 on a 7-point scale) were selected for the *Rb* scale. All items are keyed "True." Sixteen *Rb* items are not keyed on any of the regular MMPI scales; eight are scored on the *Mf* scale, while only four items are scored on any of the MMPI clinical scales. The scale has comparatively low interitem correlations (Wiggins, 1962).

Since two of the scales (*PA* and *PH*) are relatively homogeneous in content while the other two (*Sd* and *Rb*) are relatively heterogeneous, it is possible to investigate the effect of content homogeneity upon contextual effects. Moreover, the scales vary in the controversiality of the items contained in them. *PA* contains mostly items of extreme endorsement frequencies; *PH* includes mostly items of moderately extreme endorsement frequencies; while *Sd* and *Rb* include mostly items endorsed by roughly half the normal population. Table 1 presents the mean endorsement percentages for adult and college student populations for items in these four scales. In Table 1, the samples used to compute average endorsement percentages come from studies by Hathaway and Briggs (1957) and Hathaway and McKinley (1940), as reported in Dahlstrom and Welsh (1960).

All items from Welsh's Factor Scales *A* and *R*, the validating scale *K*, and the clinical scale *D* were included as filler items in each of the 12 test forms. In addition, all items from Wiggins' *Sd* and *Rb* scales were included as filler items in all test forms in which the items did not constitute the experimental scale. Additional filler items were selected in the order of their appearance in the group form of the MMPI, in order to bring the total number of items in each test form to 286. Items appeared only once in each test form. In general, an attempt was made to distribute filler items evenly throughout those portions of the test not containing items from the experimental scale.

A total of 456 subjects (294 college students,

108 neurotics, and 54 paranoid schizophrenics) were tested. All obtained scores were used. Roughly two-thirds of the college students used as subjects were volunteers from upper division psychology courses at the University of Oregon; a third of the subjects were students in a general psychology class who participated in the study to fulfill a class requirement. The total college student sample comprised 139 males and 155 females, whose ages ranged from 17 through 63, with a mean age of 25.

The neurotics were obtained from eight hospitals and clinics in the states of Oregon and Washington. Staff psychiatrists and psychologists were referred to the description of neurotic disorders by the American Psychiatric Association (1952) to select neurotic subjects for this study. Sixty-four males and 44 females were tested. Their ages ranged from 17 through 67, with a mean age of 38. The subjects included professional people, semiskilled workers, and laborers, whose educations ranged from sixth grade through completion of several years of graduate work (the median being twelfth grade). Forty-three patients were diagnosed as "anxiety reaction," 19 as "depressive reaction," 12 as "conversion reaction," and 18 as other neurotic disorders. Admitting complaints covered symptoms including depression, anxiety, nervous tension, headaches, and fatigue.

Fifty-four paranoid schizophrenics, selected from three hospitals in Oregon, were tested. Professional staff were referred to the description of paranoid schizophrenia by the American Psychiatric Association (1952) as criteria for the selection of these subjects. The sample included 48 males and six females, whose ages ranged from 23 through 67, with a mean age of 43. Length of current hospitalization varied from 2 months to 18 years. These patients were generally characterized as chronic cases, hospitalized for custodial care and not receiving treatment.

Neurotic and paranoid schizophrenic subjects were tested individually or in small groups, while college students were tested in large groups. Test instructions were given orally by the test administrator at the same time as they were read by the subject. With slight modifications, test instructions were the same as those that appear in the booklet form of the MMPI. The first testing session lasted approximately 75 minutes. It terminated after each

TABLE 3  
MEAN SCORES (AND STANDARD DEVIATIONS) FOR EIGHT  
SCALES IN THREE SAMPLES

Scale	College students ( <i>N</i> = 294)	Neurotics ( <i>N</i> = 108)	Paranoid schizo- phrenics ( <i>N</i> = 54)	<i>F</i>	<i>t</i>
<i>PA</i>	1.3 <sup>a</sup> (1.9)		5.6 (5.4)		3.9**
<i>PH</i>	3.8 <sup>b</sup> (3.5)	16.1 <sup>c</sup> (6.1)			11.2**
<i>Sd</i>	13.1 (4.1)	13.8 (3.8)	17.4 (4.4)	24.9**	
<i>Rb</i>	14.0 (3.0)	13.7 (3.1)	12.9 (3.8)	3.4*	
<i>A</i>	11.8 (7.9)	23.5 (9.6)	12.4 (10.0)	76.8**	
<i>R</i>	16.6 (5.3)	18.0 (5.3)	19.4 (5.9)	7.4**	
<i>K</i>	15.8 (4.0)	12.2 (5.0)	17.2 (6.2)	30.3**	
<i>D</i>	20.4 (4.7)	31.1 (6.8)	24.3 (7.2)	148.1**	

\* *p* < .05.  
\*\* *p* < .01.  
<sup>a</sup> *N* = 60  
<sup>b</sup> *N* = 78  
<sup>c</sup> *N* = 36

subject had finished answering the 286 items in one test booklet.

Approximately one week later, each subject was retested on a second test booklet, containing a different ordering schema of the same items. The procedures of the previous session were repeated. In addition, after completing the test, subjects filled out an Information Sheet in which they described their mood during both testing sessions and indicated whether they felt the two test forms seemed different. The subjects generally completed the second testing period in approximately one hour. These features of the design made possible supplementary analyses of test-retest effects as well as any possible attitudinal changes resulting from the retest situation.

The experimental design (Table 2) contained 24 cells, with normal subjects distributed throughout all cells and with neurotic subjects and paranoid schizophrenic subjects distributed throughout a selected number of cells.

Each cell with *PA* test forms contained 10 normals and 9 paranoid schizophrenics; each cell with *PH* test forms, 13 normals and 6 neurotics; each cell with *Sd* test forms, 13 normals and 6 neurotics; and each cell with *Rb* test forms, 13 normals and 6 neurotics. For one-half of the subjects, Administration B preceded Administration A.

## RESULTS

### *Differentiation between Samples*

Table 3 indicates the mean scores on eight scales for college students, neurotics, and paranoid schizophrenics. Except on the *Rb* scale, all analyses of variance were significant beyond the .01 level.

### *Ordering Effects on Four Experimental Scales*

Two types of statistical tests were employed to test for ordering effects upon the four experimental scales: (1) *t* tests for differences between the two administrations of the same items, each with a different ordering schema, for the same subject; and (2) analyses of variance for all subjects for first administration scores only.

Table 4 presents the findings from the first analysis. Only one of the matched pair *t* tests was significant, that being a comparison between the first and second ordering schemas for the *Sd* scale on college students.

Since so many comparisons were made, it seemed quite reasonable that this one significant finding may have arisen simply by chance. Consequently, a replication study using 26 new subjects (18 male and eight female General Psychology students) was carried out. Forms *Sd*-1 and *Sd*-2 were administered in counterbalanced order, with

TABLE 4  
*T* TESTS FOR ORDERING EFFECTS ON FOUR EXPERIMENTAL SCALES IN THREE SAMPLES (TEST-RETEST)

Scale	Sample	Number	Comparison	<i>t</i>
<i>PA</i>	College students	20	First versus scat.	.5
			First versus last	.9
			Scat. versus last	1.0
	Paranoid schizophrenics	18	First versus scat.	.3
			First versus last	.6
			Scat. versus last	.6
<i>PH</i>	College students	26	First versus scat.	1.8
			First versus last	.6
			Scat. versus last	.3
	Neurotics	12	First versus scat.	.0
			First versus last	.7
			Scat. versus last	.5
<i>Sd</i>	College students	26	First versus scat.	2.1*
			First versus last	.9
			Scat. versus last	.3
	Neurotics	12	First versus scat.	1.3
			First versus last	.7
			Scat. versus last	.1
<i>Rb</i>	College students	26	First versus scat.	.9
			First versus last	.6
			Scat. versus last	.8
	Neurotics	12	First versus scat.	.6
			First versus last	1.1
			Scat. versus last	2.0
<i>Sd</i>	College students	26	Replication study (see text) First versus scat.	.6

\* *p* < .05.

TABLE 5  
ANALYSES OF VARIANCE FOR FOUR EXPERIMENTAL SCALES IN THREE SAMPLES  
(FIRST ADMINISTRATION ONLY)

Scale	Sample	Number	Ordering schema				F
			First	Scat.	Last		
PA	College students	20	1.6	1.6	.7	Subject effects	32.4**
	Paranoid schizophrenics	18	6.8	3.8	6.1	Ordering effects	1.3
						Interaction effects	1.9
PH	College students	26	2.8	4.5	4.1	Subject effects	189.0**
	Neurotics	12	17.0	17.3	14.1	Ordering effects	1.4
						Interaction effects	1.9
Sd	College students	26	12.8	14.2	13.7	Subject effects	1.0
	Neurotics	12	13.5	12.8	16.8	Ordering effects	2.2
						Interaction effects	2.2
Rb	College students	26	13.8	13.8	14.3	Subject effects	.6
	Neurotics	12	15.4	13.1	14.8	Ordering effects	1.5
						Interaction effects	1.3

\*\*  $p < .01$ .

a one-week interval between tests. The findings from the replication can be found in the last row of Table 4. As suspected, the original finding was not replicated.

Table 5 presents the findings from four analyses of variance of first administration test forms. The PA scale significantly differentiates college students from paranoid schizophrenics and the PH scale dramatically differentiates college students from neurotics; Sd and Rb do not differentiate the college student and neurotic groups. Again, however, there are no significant ordering effects.

*Analysis of Contextual Effects on Filler Items*

Since the items from Welsh's A scale were included as filler items scattered throughout all of the test forms, it was possible to test for any contextual effects on A scores from the introduction of different experimental scales ordered in different ways. The A scale, composed of items tapping content similar to many MMPI clinical scales and loading quite highly on the MMPI factor accounting for most of the test's common variance, provides a useful reference scale for many other scales not included in this investigation. Extended analyses of the effects on A scale scores of the items from the four different experimental scales appearing before, scattered among, or

following the A scale items revealed no significant contextual effects.

*Analysis of Test-Retest Effects*

T tests were carried out on first versus second administration scores from each sample for the four experimental scales as well as the scales embedded as filler items. Since contextual effects had been shown to be insignificant, all test forms were used. Table 6 summarizes these findings. Since extensive data bearing on test-retest effects have already been gathered (Windle, 1954 and 1955), a prediction regarding the direction of test-retest changes can be made. If retest scores are assumed to change in the socially desirable direction, Sd and K scores should increase and scores on PA, PH, A and D should decrease; changes on Rb and R are unpredictable. Although most of the changes were not statistically significant, all 16 test-retest comparisons for Sd, K, PA, PH, A, and D were in the predicted direction.

*Analysis of the Information Sheets*

Seventy-nine per cent of the subjects (229 college students, 91 neurotics, and 39 paranoid schizophrenics) stated they felt that the two test-forms they had filled out were identical. The 97 subjects in the total group

who indicated any awareness of differences between their two test-forms described many types of differences. Only 8% of the total group (50 college students, six neurotics and one paranoid schizophrenic) correctly identified item ordering as a variable differentiating the two test forms, and no subject correctly described the precise nature of the ordering differences.

Fifty-nine per cent (162 college students, 65 neurotics and 43 paranoid schizophrenics) reported their belief that their performance on the two tests was identical. Of those who noted some differences in their test-retest performance, college students most often believed that there had been a change in feeling, that there had been a change in interpretation of the meaning of some items, that they had answered items more "correctly" the second time, that they had felt better on retest, or that different responses had been made after reconsidering items during the intervening week. The neurotics most often felt that the

change in performance was due to a shift in mood or "a clearer mind." The paranoid schizophrenics offered a variety of explanations.

## DISCUSSION

### *Contextual Effects*

That three different ordering schemata had no significant effects on four very different types of MMPI scales in three quite distinct subject groups should come as a heartening finding to MMPI users. Moreover, the particular ordering schemata employed in this study allow some generalization to common ordering deviations from standard group-form administration. Schema 1, in which all of the items in the scale were placed together at the beginning of the test, should correspond to the situation where one MMPI scale is administered singly, instead of the entire MMPI. Findings from this study would suggest that such a procedure would yield mean scores equivalent to those obtained

TABLE 6  
TEST-RETEST EFFECTS ON THREE SAMPLES

Sample	Scale	N	Administration 1		Administration 2		t
			$\bar{X}$	$\sigma$	$\bar{X}$	$\sigma$	
College students	PA	60	1.3	1.9	.9	1.2	1.6
	PH	78	3.8	3.5	3.2	2.8	1.3
	Sd	294	13.1	4.1	13.7	4.3	1.5
	Rb	294	14.0	3.0	14.3	3.3	1.2
	A	294	11.8	7.9	10.5	8.2	2.0*
	R	294	16.6	5.3	16.4	4.4	.5
	K	294	15.8	4.0	16.4	5.1	1.6
Neurotics	D	294	20.4	4.7	19.1	4.8	3.3**
	PH	36	16.1	6.1	14.8	6.9	.9
	Sd	108	13.8	3.8	14.3	4.3	.9
	Rb	108	13.7	3.1	13.9	3.0	.6
	A	108	23.5	9.6	22.8	10.7	.4
	R	108	18.0	5.3	18.3	5.2	.4
	K	108	12.2	5.0	12.5	4.9	.5
Paranoid schizophrenics	D	108	31.1	6.8	30.1	7.6	1.1
	PA	54	5.6	5.4	4.1	4.4	1.6
	Sd	54	17.4	4.4	18.0	4.4	.2
	Rb	54	12.9	3.8	12.7	3.9	.2
	A	54	12.4	10.0	10.6	10.3	.9
	R	54	19.4	5.9	19.7	6.1	.3
	K	54	17.2	6.2	18.1	6.2	.8
D	54	24.3	7.2	23.9	7.2	.3	

\*  $p < .05$ .

\*\*  $p < .01$ .

when the scale is embedded in the MMPI item matrix (Schema 2) or at the end of the MMPI (Schema 3).

### *Test-Retest Effects*

The findings from this study provided evidence that (a) changes in test scores on retesting tend to be in the socially desirable direction but (b) that the magnitude of these changes is typically quite small.

Windle (1954) discussed 41 studies, most of them using college students as subjects, in which 62 mean "adjustment" scores were reported for test and retest. By Windle's classification, on retest 15 of these mean scores showed "definite" increased adjustment, 16 "indefinite" increased adjustment, 24 essentially no change, 5 "indefinite" decreased adjustment, and two "definite" decreased adjustment. A test-retest positive effect occurred most often when retest was scheduled within 1 month, with the effect decreasing over time beyond a month. A large proportion of the positive studies discussed by Windle involved intervals of less than 2 months, while many of the negative studies used longer time intervals.

In a study by Windle (1955), 55 female college students were given two administrations of the MMPI with a 1-week interval between test and retest. Mean scores on five MMPI scales, *D*, *Pt*, *Hs*, *A*, and *Pd*, were found to increase significantly on retest. The present study replicates Windle's findings for the *A* and *D* scales.

Rosen (1953) measured MMPI score changes over a brief test-retest interval for a male psychiatric population (25 neurotics, 11 psychotics, two alcoholics, one inadequate personality, and one convulsive disorder). The 40 patients were tested within 3 days after hospitalization and retested approximately four days later, before "receiving much, if any, therapy of a formal nature." Five significant test-retest mean scale differences emerged: *K* increased and *Pd*, *Pa*, *Pt*, and *Sc* decreased. This trend is described by Rosen as "movement in the 'normal' direction with an improvement in defensive structure." In the present study, the performance of 108 neurotics and 54 paranoid schizophrenics did not differ significantly from test

to retest on *A*, *R*, *K*, and *D*, but all changes were in the predicted direction.

Windle (1955) considered various rationales for the test-retest effect. He listed reasons suggested by others to account for this trend: arousal and loss of interest (McGeoch & Whitely, 1927), pull of social norms with gaining of greater insight into the nature of the test (Eisenberg & Wesman, 1941), increased relaxation or hostility (Ruch, 1941), test wiseness (Kaufmann, 1950), and regression in the direction of "normality" (Schofield, 1950). Windle's (1955) 55 subjects were interviewed after the administration of the second test to determine attitudinal variables influencing performance changes. Windle judged that changes in test performance on *D*, *Pt*, *Hs*, and *A* scales were primarily related to changes in "test-taking anxiety."

The 294 college students tested in the present study were asked to indicate their test-taking moods on an Information Sheet. A qualitative analysis of this data suggested that these subjects seemed to be less interested, enthusiastic, and intellectually stimulated, and more bored and indifferent on retesting, a finding which seems similar to the loss of interest on retest as discussed by McGeoch and Whitely (1927). This finding, coupled with that of Gordon (1952) who found that subjects are more likely to respond in the socially desirable direction as they progress towards the end of a long inventory, suggests that the test-retest effect may occur because some subjects on retest become bored with the task and respond more to the social desirability of the content of the items than they did upon the item's initial presentation.

### REFERENCES

- AMERICAN PSYCHIATRIC ASSOCIATION, Committee on Nomenclature and Statistics. *Diagnostic and statistical manual of mental disorders*. Washington, D. C.: American Psychiatric Association Mental Hospital Service, 1952.
- AMERICAN PSYCHOLOGICAL ASSOCIATION. *Technical recommendations for psychological tests and diagnostic techniques*. Suppl. to *Psychol. Bull.*, 1954, 51.
- BECHTOLDT, H. P. Response defined anxiety and MMPI variables. *Proc. Iowa Acad. Sci.*, 1953, 60, 495-499.

- BENDIG, A. W. The development of a short form of the Manifest Anxiety Scale. *J. consult. Psychol.*, 1956, 20, 384.
- CHAREN, S. A note on the use of a paper-and-pencil form of the MMPI *Hs* scale for hospital use. *J. consult. Psychol.*, 1954, 18, 344.
- COMREY, A. L. A factor analysis of items on the MMPI hypochondriasis scale. *Educ. psychol. Measmt.*, 1957, 17, 568-577. (a)
- COMREY, A. L. A factor analysis of items on the MMPI depression scale. *Educ. psychol. Measmt.*, 1957, 17, 578-585. (b)
- COMREY, A. L. A factor analysis of items on the MMPI hysteria scale. *Educ. psychol. Measmt.*, 1957, 17, 586-592. (c)
- COMREY, A. L. A factor analysis of items on the MMPI paranoia scale. *Educ. psychol. Measmt.*, 1958, 18, 99-107.
- COTTLE, W. C. Card versus booklet form of the MMPI. *J. appl. Psychol.*, 1950, 34, 255-259.
- CRONBACH, L. J. *Essentials of psychological testing*. New York: Harper, 1949.
- DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook: A guide to use in clinical practice and research*. Minneapolis: Univer. Minnesota Press, 1960.
- DANA, R. H., & CONDRY, J. C. MMPI retest results: Context, order, practice, and test-taking anxiety. *Psychol. Rep.*, 1963, 12, 147-152.
- EISENBERG, P., & WESMAN, A. G. Consistency in response and logical interpretation of psychoneurotic inventory items. *J. educ. Psychol.*, 1941, 32, 321-338.
- ELLIS, A. A comparison of the use of direct and indirect phrasing in personality questionnaires. *Psychol. Monogr.*, 1947, 61(3, Whole No. 284).
- FELDMAN, M. J., & SIEGEL, S. M. The effect on self description of combining anxiety and hostility items in a single scale. *J. clin. Psychol.*, 1958, 14, 74-77.
- FERGUSON, R. G. A useful adjunct to the MMPI scoring and analysis. *J. clin. Psychol.*, 1946, 2, 248-253.
- GORDON, L. V. The effect of position on the preference value of personality items. *Educ. psychol. Measmt.*, 1952, 12, 669-676.
- HATHAWAY, S. R., & BRIGGS, P. F. Some normative data on new MMPI scales. *J. clin. Psychol.*, 1957, 13, 364-368.
- HATHAWAY, S. R., & MCKINLEY, J. C. A multiphasic personality schedule (Minnesota): I. Construction of the schedule. *J. Psychol.*, 1940, 10, 249-254.
- HOLZBERG, J. D., & ALESSI, S. Reliability of the shortened MMPI. *J. consult. Psychol.*, 1949, 13, 288-292.
- KAUFMANN, P. Changes in the Minnesota Multiphasic Personality Inventory as a function of psychiatric therapy. *J. consult. Psychol.*, 1950, 14, 458-464.
- LEBO, D., & NOBLIN, C. Buffer item variation and schizophrenic anxiety scores. *J. clin. Psychol.*, 1958, 14, 209-210.
- MACDONALD, G. L. A study of the shortened group and individual forms of the MMPI. *J. clin. Psychol.*, 1952, 8, 309-311. (a)
- MACDONALD, G. L. Effect of test-retest interval and item arrangement on the shortened form of the MMPI. *J. clin. Psychol.*, 1952, 8, 408-410. (b)
- MACHOVER, S., & ANDERSON, H. J. Validity of a paper-and-pencil form of the MMPI Psychopathic Deviate scale. *J. consult. Psychol.*, 1953, 17, 459-461.
- MCCREARY, J. B., & BENDIG, A. W. Comparison of two forms of the Manifest Anxiety Scale. *J. consult. Psychol.*, 1954, 18, 206.
- MCGEOCH, J. A., & WHITELEY, P. L. The reliability of the Pressey X-O Tests for investigating the emotions. *Ped. Sem. & J. genet. Psychol.*, 1927, 34, 255-270.
- MCQUARY, J. P., & TRUAX, W. E. A comparison of the group and individual forms of the MMPI. *J. educ. Res.*, 1952, 45, 609-614.
- OLSON, G. W. The influence of context on the depression scale of the MMPI in a psychotic population. *J. consult. Psychol.*, 1961, 25, 178-179.
- PERKINS, J. E. Contextual effects on the MMPI. Unpublished doctoral dissertation. University of Oregon, 1962.
- ROSEN, A. Test-retest stability of MMPI scales for a psychiatric population. *J. consult. Psychol.*, 1953, 17, 217-221.
- RUCH, F. L. Effects of repeated interviewing on the respondent's answers. *J. consult. Psychol.*, 1941, 5, 179-182.
- SCHOFIELD, W. Changes in responses to the Minnesota Multiphasic Inventory following certain therapies. *Psychol. Monogr.*, 1950, 64(5, Whole No. 311).
- SIEGEL, S. M., & FELDMAN, M. J. A note on the effect on self description of combining anxiety and hostility items on a single scale. *J. clin. Psychol.*, 1958, 14, 389-390.
- TAYLOR, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.
- VOAS, R. B. Comparison of the Taylor Anxiety Scale administered separately and within the MMPI. *Psychol. Rep.*, 1956, 2, 373-376.
- WIENER, D. N. Differences between the individual and group forms of the MMPI. *J. consult. Psychol.*, 1947, 11, 104-106.
- WIGGINS, J. S. Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *J. consult. Psychol.*, 1959, 23, 419-427.
- WIGGINS, J. S. Strategic, method and stylistic variance in the MMPI. *Psychol. Bull.*, 1962, 59, 224-242.
- WINDLE, C. Test-retest effect on personality questionnaires. *Educ. psychol. Measmt.*, 1954, 14, 617-633.
- WINDLE, C. Further studies of test-retest effect on personality questionnaires. *Educ. psychol. Measmt.*, 1955, 15, 246-253.

(Received July 23, 1963)