Sum of MMPI Scales F, 4, and 9 as a Measure of Aggression

L. Rowell Huesmann University of Illinois at Chicago Circle Monroe M. Lefkowitz

New York State Department of Mental Hygiene Albany, New York

Leonard D. Eron University of Illinois at Chicago Circle

The relationship between scores on the Minnesota Multiphasic Personality Inventory and both concurrent and prior aggression was examined for a sample of 426 19-year-olds from the general population. Aggression was measured through peer nominations obtained concurrently and 10 years earlier. Correlation and regression analysis indicated that the sum of T scores for Scales F, 4, and 9 was a valid measure of aggression. The composite was also shown to have a higher reliability than its component scales. Using an additional 283 subjects from delinquent populations, it was demonstrated that the composite was an excellent discriminator between delinquent and general populations of males and females even when intelligence and social status were controlled.

Elevation of the T scores on both Scales 4 and 9 of the Minnesota Multiphasic Personality Inventory (MMPI) has been thought to form the profile characteristic of the male juvenile delinquent (Dahlstrom & Welsh, 1960; Dahlstrom, Welsh, & Dahlstrom, 1972; Hathaway & Monachesi, 1953). Scale 4 (Psychopathic Deviate) by itself has been used to measure levels of social deviance or antisocial behavior (Elion & Megargee, 1975; Hathaway & Monachesi, 1953; Megargee & Mendelsohn, 1962). Dahlstrom et al. (1972) noted that marked elevations on Scale 4 can be observed in prison groups. These authors also noted that Scale 9 (Hypomania) appears to energize the pattern related to Scale 4. Scale 9 was viewed by Hathaway and Monachesi (1953) as an exciter that in combination with

Requests for reprints should be sent to L. Rowell Huesmann, Department of Psychology, University of Illinois at Chicago Circle, Box 4348, Chicago, Illinois 60680. Scale 4 produces rebellious and excitable behavior in high-delinquent children. A later analysis of their data (Monachesi & Hathaway, 1969) showed that the highest rates of delinquency for both boys and girls were associated with deviance in MMPI Scales 4, 8, and 9.

In a study of juvenile delinquents who succeeded or failed in adjustment to institutionalization, Lefkowitz (1966) found that the mean score on Scale 9 was significantly higher for the failures. Similarly, in a follow-up study of parolled prisoners whose postrelease behavior was classified as acceptable or unacceptable, the MMPI 49 code type was heavily represented among the unacceptable group (Jacobson & Wirt, 1969). Butcher (1965) found that highly aggressive boys (based on peer nominations) had significant elevations on Scales 4 and 9. These boys responded in a rebellious and excitable manner in interpersonal situations, whereas the lowaggressive boys tended to internalize their conflicts, which were then manifested in hypochondriacal symptoms and withdrawal.

There has been a proliferation of attempts to validate not only certain of the clinical scales as aggression measures (see, e.g., Megargee & Mendelsohn, 1962; Shipman, 1965) but also a number of special scales developed from the MMPI. At this writing,

Copyright 1978 by the American Psychological Association, Inc. 0022-006X/78/4605-1071\$00.75

This study was supported by Grant M1726 from the National Institute of Mental Health, by Contract HSM 42-70-6 from the Surgeon General's Scientific Advisory Committee on Television and Social Behavior, and by the New York State Department of Mental Hygiene. The data analysis was performed with the assistance of the computation center of the University of Illinois at Chicago Circle. For Part 1 of this study, Leopold Walder was a co-investigator.

as many as 25 scales of the MMPI have been proposed as measures of hostility or aggression (Deiker, 1974). Thus, the present study is an attempt to simplify this area of measurement by providing a reliable and valid single score of aggressive behavior obtained by summing the T scores of a few common clinical scales. The virtue of a single score over a profile of scale scores is that the former is readily usable as a continuous variable, whereas the latter is a dichotomous variable and therefore in the main limited to frequency analysis.

A two-step procedure was followed in this study. First, using a sample of subjects from a general population, we attempted to identify a valid and reliable composite of T scores for measuring aggression. Then using a sample of subjects from a delinquent population, we attempted to validate the composite as a discriminator of delinquent adolescents.

Study 1

Method

Table 1

Subjects. The subjects were part of a larger longitudinal research project on aggressive behavior reported elsewhere (Eron, Huesmann, Lefkowitz, & Walder, 1972; Lefkowitz, Eron, Walder, & Huesmann, 1977). In the first wave, data were gathered during 1959-1960 from the entire 3rd-grade population of 875 boys and girls residing in a semirural county in New York State. Ten years later in the second wave, termed the "13th grade," data were collected from 426 of these subjects (211 boys and 215 girls who could be located at that time; Lefkowitz et al., 1977). In the 3rd and 13th grades, the model ages of this sample of 426 were 8 and 19 years, respectively. The mean IQ of this sample in the 3rd grade was 107.10 ± 13.66 . Based on fathers' occupation, the sample can be described as predominantly middle class.

Procedure. In the 3rd and 13th grades, aggression scores were obtained by a peer nomination questionnaire comprised of items describing aggression behavior. For example, each subject was asked to name anyone in the classroom "Who starts a fight over nothing?" The peer nomination technique also included items intended to measure aggression avoidance; for example, "Who will never fight even when picked on?" The reliability (r > .85) and validity of this aggression measure and its the procedures for administration have been discussed elsewhere (Butcher, 1965; Eron, Walder, & Lefkowitz, 1971; Walder, Abelson, Eron, Banta, & Laulicht, 1961). In the 3rd grade the peer nomination technique, including 10 aggression items, was administered to classroom groups. In the 13th grade the technique, using 9 of the 10 aggression items from the 3rd grade, was administered individually to subjects as part of a larger 2-hour interview procedure. At the time of the 3rd grade interview, IQ (Sullivan, Clark, & Tiegs, 1957) and father's occupational status (U.S. Bureau of Census, 1960) were recorded for most subjects. During the 13th grade interview, subjects also completed self-report inventories on aggressive behavior and were administered the MMPI. The K-corrected T scores on the MMPI were used as the potential predictors of aggression in this study. The self-report measure of aggression, termed total aggressive environment, was comprised of 43 items divided into five subscales: (a) respondent as a victim of aggression, (b) respondent as a witness of aggression, (c) respondent's aggressive habits, (d) respondent's antisocial behavior, and (e) respondent's aggressive feelings. The exact composition of this scale and the procedures used to administer all of the 13th grade measures have been reported elsewhere (Lefkowitz et al., 1977).

Results

The intercorrelations between the various measures of aggression are shown in Table 1 for all 426 subjects. Since 13th grade peer-rated aggression (Peer Agg 13) correlated well with all the other measures and was measured at the same time that the MMPI was admin-

Measure	1	2	3	4	5	
1. Peer Agg 13						
2. Peer Avoid Agg 13	364					
3. Self Rep Agg 13	.520	357				
4. Peer Agg 3	.420	284	.200			
5. Peer Avoid Agg 3	236	.298	233	381		
6. Sex ^a	.346	080	.390	.199	139	

Correlations	Between	the	Criterion	Measures	of	Aggression	for	All 42	6 Subjects

Note. $r_{.999} = .15$. Peer = peer nominated; Agg = aggression; Avoid = avoidance of; Rep = reported; 3 refers to Grade 3; 13 refers to Grade 13. ^a Female = 0; male = 1.

1072

I abic 2	Ta	ble	2
----------	----	-----	---

		Concurrent peer-rated aggression			
	M	alesª	l'emales ^b		
MMPI scale	Ÿ	Raw regression coefficient	ŗ	Raw regression coefficient	
L	.064		.087		
F	.322***	1.675*	.260***	1.302***	
K	134		056		
1	.052		.049		
2	.077		.005		
3	.001		.090		
4	.339***	2.469***	.327***	1.073***	
5	129	-1.799**	.182**	.561*	
6	.180**		.121		
7	.131		017		
8	.232***		.111		
9	.332***	1.351*	.146*		
10	132		150*	674*	

Correlations and Significant Multiple Regression Coefficients (MRs) for Predicting Concurrent Aggression from Minnesota Multiphasic Personality Inventory (MMPI) Scales

^a For males, n = 211, $R = .471^{***}$.

^b For females, n = 215, $R = .478^{***}$.

istered, it was decided to use it as the primary criterion measure. Since sex was significantly related to aggression, all further analysis were done separately for males and females.

Table 2 shows the correlations between Peer Agg 13 and each of the MMPI scales as well as the results of a multiple regression analysis predicting Peer Agg 13 from the MMPI scales. For males, as expected, Scales 4 and 9 were highly significant correlates of aggression, but Scale F was also highly correlated. Scale 8 correlated less highly though significantly. For females, the two best unvariate predictors were Scale 4 and the F scale. Scales 5, 9, and 10 also were significantly correlated with aggressiveness but not as highly. Since the Tscores for Scales 4, 7, 8, and 9 were K corrected and since K scores were slightly negatively correlated with aggressiveness, the observed correlations between aggressiveness and 4, 7, 8, and 9 may have been slightly reduced by the K correction. However, the reduction could not be significant considering the small size of the correlation between K and Peer Agg 13.

The multiple regression analysis shows the

best weighted composite of the T scores for measuring aggression.¹ As expected, the regression coefficients reveal that Scale 4 was the most important predictor of aggression across sexes. Surprisingly, however, Scale F was as important a predictor for females and equal with Scales 5 and 9 in predicting aggression in males. Scale 5 relates negatively to aggression in males and positively in females, indicating the relation between aggression and sex role behaviors. Scale 9 made an important contribution for males but not for females, for whom Scale 10 appeared as an inverse predictor.

Discussion

From these results it appears that a composite to be used as a general measure of

^{*} p < .05.

^{**} p < .01. *** p < .001.

¹ The raw regression coefficients are presented because the objective is to develop a predictive composite. Even though the intercorrelations between the MMPI scales are mostly significant, they only range from -.32 to .52 for the significant scales in the equation, so multicolinearily should not be a serious problem.

					Sca	le				
		Ma	les $(n = 211)$				Feme	ales $(n = 21)$	5)	
Aggression measure	F + 4 + 9	F + 4	F	4	6	$F + \frac{1}{4} + 9$	F + 4	F	4	6
Concurrent Peer Age 13	414***	.377***	.322***	.339***	.332***	.313***	.343***	.257***	.327***	.146*
Peer Avoid Agg 13	314**	274***	252***	227***	271***	257***	246***	179^{**}		177**
Self-Kep Agg 13	.495***	.433***	.359***	.401***	.426***	.408***	.354***	.200**	.393***	.342***
10 years prior Peer Age 3	210***	218***	170*	.212***	.144*	137*	.107	101	.086	134*
Peer Avoid Agg 3	311***	303***	.248***	.284***		079	085	100	052	038
Note. Peer = peer nomin	ated , Agg = ag	gression; Avo	oid + avoida	nce of ; Rep	= reported, 3	refers to Grade	: 3; 13 refers t	co Grade 13.		

aggression across sexes without contamination by sexual stereotyping should be composed of Scales F and 4 and either Scale 9 or scale 10. Of course, a weighted linear composite involving every scale would have the highest correlation with aggression. However, a measure consisting of the sum of only two or three scales has the advantage of simplicity and ease of administration. Furthermore, as Wainer (1976) and others have argued, the simple sum of significant predictor variables is often almost as good a predictor in the population as the entire multiple regression equation.

Scale 5, although a significant predictor, is not a desirable variable to include in a general measure of aggression, because it represents that component of aggression associated with sexual stereotyping. Since it was included in the regression equation, however, the effects of sexual stereotyping have been partially controlled in determining the other significant predictors.

On a priori grounds, Scale 9 would seem to be a more appropriate element to include in a measure of aggression than Scale 10. It is intended to measure a propensity for manic behavior, so it certainly has face validity as a component of aggression. Furthermore, as reported in the Introduction, Scale 9 has been found to be significantly higher in several high-aggressive, albeit male, populations. On the other hand, Scale 10 has to our knowledge never been proposed as a discriminator of delinguents. In addition, the utility of Scale 10 in this study could possible be an artifact of the peer nomination procedure. A female scoring high on Scale 10 is scoring high on social introversion. Therefore, she might be less likely to be nominated by her peers on any scale resulting in negative correlations. The available data are consistent with this interpretation. Even though Scale 10 correlated negatively with peer-rated aggression for females (r = -.15, p < .03), it did not correlate significantly with self-rated aggression (r = .04), though one would usually expect a higher correlation between two self-ratings. Also, peer-rated popularity, which correlated negatively with peer-rated aggression (r = -.28, p < .001), had exactly the same correlation with Scale 10 (r = -.15) as did peer-rated aggression. Thus, it seems that

Table 3

0000

, d d * * * Table 4

Scale 10 is measuring a component of peer rating more than aggression. Scale 9, on the other hand, correlated more significantly with self-rated aggression for girls (r = .34, p < .001) than with peer-rated aggression (r = .15, p < .03). Therefore, a composite measure of aggression was constructed consisting of the sum of Scales F, 4, and 9.

Table 3 displays the correlations between this composite measure of aggression and each of the other variables measuring aggression. It also contains the correlations between the aggression measures and the individual scales. One can see that for males the sum of Scales F, 4, and 9 is a valid measure of concurrent aggression and is significantly related to the subject's aggression to years earlier. The composite correlated more strongly with every measure than any of its components or F + 4. For females, the composite also correlated significantly with every measure of concurrent aggression, though less strongly than for males. Scale 4 alone and F + 4 are just as strongly or slightly more strongly related to peer ratings than the full composite, but the full composite did better on the other measures. The composite also significantly correlated with aggression 10 years earlier.

Reliability

The reliability of any linear combination of variables can be computed from its variance and the variances and reliabilities of its components. For the current sample, however, MMPI item scores were not available to us, so reliability data from comparable samples had to be used. The largest study of reliability in a college-age population of normals appears to have been conducted by Mauger in 1972 (cited in Dahlstrom et al., 1972). Test-retest correlations were computed for 490 subjects over an 8-month lag. The average correlations of males and females were .56 on Scale F, .57 on Scale 4, and .62 on Scale 9. These relatively long-term stability coefficients can be viewed as lower bounds on internal consistency reliabilities (i.e., coefficient alpha). Dahlstrom et al. (1972) did not report any studies involving a substantial number of normal college-age subjects for which coefficient alpha was calculated, so these stability coefficients

ex and Delinqu	lency		
Population	Males	Females	Total
General	183.3	174.6	178.9
п	210	215	
Delinquent	217.4	237.7	227.2
n	147	136	
Total	197.3	199.0	198.2

Means for F + 4 + 9 as a Function of

will have to serve as our conservative estimates of reliabilities. We can now show that even assuming such conservative estimates, the

reliability of F + 4 + 9 is acceptable. From Nunnally (1967, p. 229), we can write: Let

$$y = \text{MMPI}_F + \text{MMPI}_4 + \text{MMPI}_9$$

$$r_{yy} = (\sigma_F^2 + \sigma_4^2 + \sigma_9^2) - r_{FF}\sigma_F^2 - r_{44}\sigma_4^2 - r_{99}\sigma_9^2$$

 σ_y^2 Using our sample of 427 subjects to estimate the standard deviations gives

$$\sigma_{u} = 26.52 \ \sigma_{F} = 10.82 \ \sigma_{4} = 11.15 \ \sigma_{9} = 11.50,$$

and using Mauger's stability coefficients as conservative estimates of reliability gives

$$r_{FF} = .56$$
, $r_{44} = .57$, and $r_{99} = .62$.

Therefore,

$$r_{yy} = 1 - \frac{155.23}{703.31} = .78.$$

Thus the reliability of F + 4 + 9 is substantially higher than the reliabilities of its components and sufficient for its use as a measure of aggression. Furthermore, using less conservative estimates of the reliabilities of the scales based on 1-week stabilities (Dahlstrom et al., 1972), the reliability of F + 4 + 9 is almost .87.

Study 2

To further validate F + 4 + 9 as a measure of aggression, a second analysis was undertaken to compare a sample from a known population of delinquents with the previously studied sample from a normal population.

Source	df	SS	MS	F	
Covariates					
IQ	1	157,298.44	157,298.44	218.27*	
Father's occupation	1	13,772.81	13,772.81	19.11*	
Effects			,		
Sex (A)	1	10.496.69	10.496.69	14.57*	
Population $(B)^n$	1	153,635,63	153,635,63	213.19*	
$A \times B$	1	17.628.37	17.628.37	24.46*	
Residual	621	447.522.25	720.65		
Total	626	800.354.19	120100		

Table 5

Analysis of Covariance for F + 4 + 9 as a Function of Sex and Delinquency

Note. The total sample size for the analysis of covariance was only 627 because IQ or father's occupation was missing for 81 subjects.

^a Delinquent versus general.

* p < .001.

M ethod

Subjects. The delinquency sample consisted of 136 females institutionalized at a facility of the New York State Division for Youth and 147 males from a privately operated institution for delinquent boys in New York.² All subjects had been sent to the institutions by family and children's courts. The subjects ranged in age from 12.9 to 17.1, with a median of 14.9. The mean IQ for girls was only 77.3 but was 100.3 for boys. The entire population of the institutions minus a few subjects for whom usable data could not be obtained constituted the sample. The offenses committed by the delinquents ran the gamut from incorrigibility, shoplifting, petty larceny, and prostitution to the car theft, breaking and entering, and assault.

Procedure. Form R of the MMPI was group administered to 3–5 subjects at a time. The 399 items were read aloud twice on a tape while the subjects, closely monitored, read along silently in their booklets. The IQs and fathers' occupations were obtained from the subjects' case histories.

Results

The means and analysis of covariance shown in Tables 4 and 5 indicate that F + 4 + 9is an excellent discriminator between populations known to vary in their levels of aggression. The analysis of covariance was performed in a hierarchical manner, so the F value for delinquent versus nondelinquent populations, $F(1, 621) = 244.1, \ p < .001$, represents the discriminative strength of F + 4 + 9 after the effects of IQ, father's occupation, and sex have been partialed out. Since the standard deviation of the composite was about 25, the table of means indicates that delinquent boys scored about $1\frac{1}{2}$ standard deviations higher than normals, and delinquent girls scored about $2\frac{1}{2}$ standard deviations higher. This greater difference for females is reflected in the significant Sex × Population interaction F(1, 621) = 25.83, p < .001. Although males scored higher than females on F + 4 + 9 in the normal population, t(423) = 3.42, p < .001, females scored higher in the delinquent population, t(281) = 5.98, p < .001.

Discussion

These results provide construct validity for the sum of MMPI Scales F, 4, and 9 as a measure of aggression. The sum of these scales was an excellent discriminator between delinquents high in aggression and youth from a normal population even when IQ and social class were controlled. Furthermore, as one would predict for a measure of aggression, nondelinquent males scored significantly higher on F + 4 + 9 than nondelinquent females. Within the delinquent group, however, females, contrary to the hypothesis, scored significantly higher than males. Speculatively, this unpredicted result of higher scores for delinquent females than delinquent males on this MMPI measure may be due to an artifact of selection. Traditionally, females have encountered far fewer difficulties with law-enforcing agencies

² The authors wish to express their thanks for the cooperation of the staffs of the Hudson School for Girls and the Berkshire Farm Institute for Training and Research.

and/or have been treated more leniently by these agencies than males, particularly juveniles (Monachesi & Hathaway, 1969). Thus, in order for a female to be considered delinguent and institutionalized, the level of antisocial behavior may have to be excessive as compared to males. Consequently, the present group of institutionalized delinquent females may be unrepresentative of delinquent females generally, in that their level of aggressiveness may be inordinately high. The unusually low mean IQ for the female delinguents supports this interpretation. If this post hoc hypothesis is valid, this MMPI measure should be able to discriminate among degrees of delinquency in a population of delinquent females.

Although the normals in the present study were approximately 5 years older than the delinquents, it is unlikely that this age difference affected the results. Hathaway and Monachesi (1963) provided MMPI norms for 4,944 boys and 5,207 girls in the ninth grade whose average ages, respectively, were 15.1 and 14.9 years. The sum of the T scores on Scales F, 4, and 9 were 176.3 for boys and 170.4 for girls. These values are markedly below those of delinquents of the same age (as shown in Table 4) but approximately the same as those of the normal subjects in the present study.

Summary

The results of the present study indicate that the sum of the T scores on MMPI Scales F, 4, and 9 serves as a reliable and valid unitary measure of aggression. Highly significant correlations obtained between F + 4 + 9and both peer-rating and self-rating of aggression. Added strength is lent to the validity by the fact that scores on the measure related back 10 years in time to nominations on aggressive behavior that the subjects received from their peers in a classroom setting. Moreover, as evidenced by the multiple regression analysis, the composite is valid for both males and females and is not simply a sex-typing measure. In general, the sum of these scales appeared more valid than any of its components. Other support for the validity of this MMPI measure was derived from its ability

to distinguish between delinquent and nondelinquent populations known to differ in aggressiveness. Finally, F + 4 + 9 was shown to have a reliability greater than that of any of its components.

Although the intent was to validate this measure on a noninstitutionalized population of normal subjects, the fact that the measure discriminates so well between institutionalized and noninstitutionalized populations suggests that it may also be useful in criminal justice settings in which the MMPI is so widely used. The ease with which this measure is obtained from the MMPI makes it a potentially useful tool for screening and perhaps program placement in such settings. Also, the measure meets the demand for a valid paper-and-pencil measure of aggressive behavior for research with normal subjects.

References

- Butcher, J. Manifest aggression: MMPI correlates in normal boys. Journal of Consulting Psychology, 1965, 29, 446-454.
- Dahlstrom, W. G., & Welsh, G. S. An MMPI handbook. Minneapolis: University of Minnesota Press, 1960.
- Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. An MMPI handbook. Minneapolis: University of Minnesota Press, 1972.
- Deiker, T. E. A cross-validation of MMPI scales of aggression on male criminal criterion groups. *Journal* of Consulting and Clinical Psychology, 1974, 42, 196–202.
- Elion, V. H., & Mcgargee, E. I. Validity of the MMPI Pd scale among black males. Journal of Consulting and Clinical Psychology, 1975, 43, 166–172.
- Eron, L. D., Huesmann, L. R., Lefkowitz, M. M., & Walder, L. O. Does television violence cause aggression? American Psychologist, 1972, 27, 253-263.
- Eron, L. D., Walder, L. O., & Lefkowitz, M. M. Learning of aggression in children, Boston: Little, Brown, 1971.
- Hathaway, S. R., & Monachesi, E. D. Analyzing and predicting juvenile delinquency with the MMPI. Minneapolis: University of Minnesota Press, 1953.
- Hathaway, S. R., & Monachesi, E. D. Adolescent personality and behavior: MMPI patterns of normal, delinquent, dropout, and other outcomes. Minneapolis: University of Minnesota Press, 1963.
- Jacobson, J. L., & Wirt, R. D. MMPI profiles associated with outcomes of group psychotherapy with prisoners. In J. N. Butcher (Ed.), *MMPI: research develop*ments and clinical applications. New York: McGraw-Hill, 1969.
- Lefkowitz, M. M. MMPI scores of juvenile delinquents adjusting to institutionalization. *Psychological Reports*, 1966, 19, 911-914.

- Lefkowitz, M. M., Eron, L. D., Walder, L. O., & Huesmann, L. R. Growing up to be violent: A longitudinal study of the development of aggression. New York: Pergamon Press, 1977.
- Megargee, E. I., & Mendelsohn, G. A. A cross-validation of twelve MMPI indices of hostility and control. *Journal of Abnormal and Social Psychology*, 1962, 65, 431-438.
- Monachesi, E. D., & Hathaway, S. R. The personality of delinquents. In J. N. Butcher (Ed.), *MMPI: Research developments and clinical applications*. New York: McGraw-Hill, 1969.
- Nunnally, J. C. *Psychometric theory*. New York: McGraw-Hill, 1967.
- Shipman, W. G. The validity of MMPI hostility scales. Journal of Clinical Psychology, 1965, 21, 186-190.

- Sullivan, E. T., Clark, W. W., & Tiegs, E. W. California short-form test of mental maturity. Los Angleles: California Test Bureau, 1957.
- United States Bureau of Census. Classified index of occupations and industries. Washington, D.C.: U. S. Government Printing Office, 1960.
- Wainer, H. Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin*, 1976, 83, 213-217.
- Walder, L. O., Abelson R., Eron, L. D., Banta, T. J., & Laulicht, J. H. Development of a peer-rating measure of aggression. *Psychological Reports*, 1961, 9, 497-556.

Received August 19, 1977