Inter-Inventory Predictability and Content Overlap of the 16 PF and the CPI

John B. Campbell and Ki-Taek Chun
Institute for Social Research
The University of Michigan

Each scale on the 16 PF and the CPI was predicted from the scales on the other inventory using both standard and stepwise multiple regression ($N = 241$ undergraduates). The discrepancy between the predictabilities obtained by these two methods was minimal. The cross-validational shrinkage of the stepwise regression predictabilities, when examined by both the conventional and the McNemar methods, was quite small. The mean predictability was .63 for the 16 PF and .64 for the CPI. Four 16 PF and five CPI scales were "highly predictable," while four 16 PF and three CPI scales were essentially non-predictable. Seven scales from each inventory appeared to have counterparts in the other inventory. Thus, despite major differences in philosophy and strategy of construction, the overall predictability remained the same whether the 16 PF scales were predicted from the CPI scales or vice versa. Furthermore, the pattern of predictabilities suggested a substantial overlap between the 16 PF "Adjustment vs. Anxiety" and CPI "Adjustment" factors, and between the 16 PF "Introversion vs. Extroversion" and CPI "Extroversion" factors.

The number of psychological inventories has increased steadily in recent years (e.g., Goldberg, 1971). This proliferation would be a welcome development, were it not for the lack of information regarding the relationship among inventories and the consequent inability to integrate findings based on different inventories. The need to examine the degree of redundancy or independence between inventories is obvious; in fact, Hundleby and Connor consider it to be "a generally accepted maxim of the testing movement" (1968, p. 152) that the inter-inventory relationships should be examined. However, this "maxim" is more often acknowledged than heeded. Table 1, which lists 23 published studies investigating relationships among prominent inventories, reveals how incomplete the information on inter-inventory relationships is. Eleven of the 28 cells are completely empty, indicating that 11 of the 28 inter-inventory relationships have never been examined in the literature. Seven more of the 28 cells have one entry only. Of the 28 inter-relationships, only those six among the MMPI, CPI, 16 PF, and EPPS[1] have been examined in three or more studies. However, those studies investigating the same inter-relationship often have used different analytic techniques without explicating the nature of the relationships among the techniques (see Footnotes a through f, Table 1). Consequently, even the information on these six relationships is in-

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 1, No. 1 Winter 1977 pp. 51-63
© *Copyright 1977 West Publishing Co.*

[1]The following abbreviations have been used in the present study: 16 PF (Sixteen Personality Factor Questionnaire), CPI (California Psychological Inventory), CPS (Comrey Personality Scales), EdPI (Edwards Personality Inventory), EPPS (Edwards Personal Preference Schedule), EyPI (Eysenck Personality Inventory), MMPI (Minnesota Multiphasic Personality Inventory), and PRF (Personality Research Form).

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

Table 1

Published Studies of Inter-inventory Relationships among Major Inventories

Inventory	MMPI	CPI	16 PF	EyPI	CPS	EPPS	EdPI
CPI	Gough (1960)[a] Rodgers (1966)[e] Megargee (1966)[e] Wiggins, Goldberg, & Appelbaum (1971)[a]						
16 PF	Karson & Pool (1957)[a] LaForge (1962)[abc] Hundleby & Connor (1968)[a] Cattell & Bolton (1969)[abdf] O'Dell & Karson (1969)[bd] Cattell, Eber, & Tatsuoka (1970)[f]	Gough (1960)[a] Mitchell (1963)[ab] Edwards & Abbott (1973b)[b]					
EyPI	Hundleby & Connor (1968)[ag]		Comrey & Duffy (1968)[b] Hundleby & Connor (1968)[a]				
CPS			Comrey & Duffy (1968)[b]	Comrey & Duffy (1968)[b]			
EPPS	Merrill & Heathers (1956)[a] Allen (1957)[a] Wiggins et al. (1971)[a]	Dunnette, Kirchner, & DeGidio (1958)[a] Gough (1960)[a] Wiggins et al. (1971)[a] Edwards & Abbott (1973b)[b]	Borgatta (1962)[a] Lemke & Kirchner (1971)[b] Edwards & Abbott (1973b)[b]				
EdPI		Abbott (1971)[ab] Edwards & Abbott (1973b)[b]	Edwards & Abbott (1973b)[b]			Edwards & Abbott (1973a)[b] Edwards & Abbott (1973b)[b]	
PRF	Trott & Morf (1972)[ab]	Jackson (1967)[a]				Edwards, Abbott, & Klockars (1972)[ab] Edwards & Abbott (1973a)[b]	Edwards & Abbott (1973a)[b]

[a,b,c,d,e,f]These superscripts refer to the type of analytic technique(s) used in individual studies: (a) inter-inventory correlation coefficients, (b) factor analysis of an inter-inventory correlation matrix, (c) separate factor analyses of intra-inventory correlation matrices, (d) inter-inventory canonical correlation, (e) inter-inventory prediction based on item overlap, (f) inter-inventory prediction based on multiple regression.

[g]Hundleby & Connor actually analyzed the Maudsley Personality Inventory, the predecessor to the EyPI.

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

complete and fragmented, and as a result, the findings of these studies are not readily amenable to integration.

Although inventory users generally are interested in knowing the degree of inter-inventory predictability, only 4 of the 23 studies listed in Table 1 deal with the predictability of scale scores of one inventory from those of another (those four studies are superscripted with "e" or "f"). The inattention to this pragmatic aspect of inter-inventory relationships, i.e., predictability, also can be recognized in another way. Nearly half of those studies listed in Table 1 (i.e., those with superscript "a") relied on correlations between pairs of scales from two inventories as an index of inter-inventory relationship. Since the intercorrelations already were computed, a minimum of additional work would have provided information on the predictability of scale scores from one inventory to the other.

The few studies which have focused on inter-inventory predictability provide a useful beginning in this area. For instance, Rodgers (1966) capitalized on the item overlap between the CPI and MMPI to assess the degree of CPI-MMPI predictability. (This study indeed may have been the first to examine the inter-inventory predictability between any inventories.) While Rodgers was interested in estimating MMPI scale scores from the CPI, Megargee (1966) was concerned with predicting CPI scale scores from the MMPI. In these two studies, the overlapping items (i.e., those present in both the predictor and predicted inventories) were scored in the predictor inventory, pro-rated for length, and used as an estimate of the to-be-predicted scale score. However, since the item overlap between most inventories is not as extensive as that between the CPI and the MMPI, the general applicability of this technique is necessarily limited.

As Cattell and Bolton (1969) suggest, a more general method of assessing inter-inventory predictability would be that of either multiple regression or canonical correlation. As Weiss (1972) points out, there is a close relationship between multiple correlation and canonical correlation. Canonical correlation analysis, however, is based on linear combinations of scales (i.e., canonical variates) in both the predictor and the predicted inventories. Consequently, its interpretation in terms of the predictability of individual scales is not as straightforward as that provided by multiple regression. The R^2 generated by multiple regression estimates the proportion of variance of individual scale scores in one inventory that can be predicted from those in another inventory. Although this index can be derived from canonical correlation, the prediction of individual scale scores and its interpretation in terms of individual predictor scales is possible only with multiple regression. Therefore, given the purpose of the present paper, multiple regression seems the more practical of the two methods.

The multiple regression approach has been used to predict the MMPI scales from the 16 PF scales (Cattell & Bolton, 1969), the MMPI and Guiford-Zimmerman Temperment Survey scale scores from the 16 PF scale scores (Cattell, Eber, & Tatsuoka, 1970, pp. 43-47), and the scores on the Comrey Personality Scales from the California Life Goals Evaluation Schedules, as well as the scores on the Evaluation Schedules from the Comrey Scales (Backer, Comrey, & Hahn, 1971). Although these studies make it possible to assess the overall degree of inter-inventory predictability, they can be criticized for their failure both to cross-validate and to identify the significant predictor scales.

The present study uses the multiple regression approach to assess the feasibility of reciprocal prediction between the 16 PF scales and the CPI scales (i.e., the prediction of each 16 PF scale from the CPI scales and of each CPI scale from the 16 PF scales). The study specifically: 1) examines the utility of using a subset rather than the entire set of scales from the predictor inventory; 2) establishes the degree of cross-validated predictability; and 3) evaluates the predictability of individual scales against their test-retest reliabilities. Based upon the predictability data, this study then examines the content overlap between the two inventories.

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

Method

As part of a larger study, 247 undergraduates at Cornell University were paid to complete the 16 PF (Form A + B) and the CPI. The order in which the inventories were taken was left to the individual. The protocols of six students were discarded prior to analysis because of failure to follow instructions, resulting in a final sample size of 241. Analysis was carried out in the following three stages.

Entire Set of Predictors

The goal here was to estimate the population predictability of each 16 PF and CPI scale from the entire set of scales on the other inventory. The generally accepted means of obtaining such estimates is to split one sample into halves and then to employ cross-validation or double cross-validation procedures (Mosier, 1951; Norman, 1965). However, recent evidence (as summarized in Wiggins, 1973, pp. 47-48) suggests that predictability estimates, obtained by correcting for shrinkage the values of R based on the total sample, are as accurate as estimates obtained by splitting the sample and applying cross-validation procedures. Furthermore, estimates based on the total sample should be more stable than estimates based on half-samples, as they will not be susceptible to possible sampling biases introduced by arbitrarily splitting the sample into halves. Consequently, predictabilities for the 16 PF and CPI scales, based on the entire set of predictors (i.e., the 18 scales from the CPI and the 16 scales from the 16 PF), were estimated by correcting for shrinkage the values of R obtained by computing standard multiple regression equations on the entire sample.

Selected Predictors

Not all the scales from one inventory will be of equal importance in the prediction of a given scale from another inventory, and inclusion of these marginal scales in the prediction equation will only increase the possibility of capitalizing on chance variation. Therefore, each scale on the 16 PF and CPI also was predicted using a forward stepwise regression algorithm which included as predictors only those scales with significant sequential ($p < .05$) and partial ($p < .10$) F ratios (e.g., see Draper & Smith, 1966, pp. 171-172). The predictabilities generated from the entire set of predictor scales obviously will serve as an upper bound on the predictabilities obtained using only selected predictors. When the discrepancy between the two sets of predictabilities is minimal, the use of the selected set is preferable primarily because of interpretive simplicity.

Selecting predictor variables on the basis of their ability to predict the criterion in a given sample is susceptible to capitalization on chance variation peculiar to the given sample. As a result, some predictors may be selected on fortuitous grounds. Since shrinkage formulas do not take account of this source of error, some form of cross-validation must be employed (see McNemar, 1969, pp. 208-209; Wiggins, 1973, pp. 48-49). There is, however, no general agreement regarding what cross-validation techniques are most appropriate in this situation. For comparative purposes, therefore, two different techniques of cross-validation were employed in the present study: 1) The sample was split into odd and even halves (based on student identification numbers), and the stepwise regressions were computed for each 16 PF and CPI scale within both half-samples. For each scale, the regression equation (i.e., the weights for the selected predictors) developed in the odd half-sample then was applied in the even half-sample, and conversely, the regression equation developed in the even half-sample was applied in the odd half-sample. The two resulting values of R for each scale were then taken as estimates of the predictability of that scale. 2) Extending McNemar's recommendation for cross-validation (1969, pp. 208-209), stepwise regressions were computed for each 16 PF and CPI scale within both half-samples. Using only those scales identified as significant predictors in one half-sample, standard multiple regression equations were developed in the other half-sample. This procedure was repeated from odd to even and

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction
requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

from even to odd halves. For each scale, the two resulting values of R, corrected for shrinkage, were taken as estimates of the predictability of that scale.

A minimal drop in R upon cross-validation would indicate that the two half-samples do not have large proportions of idiosyncratic or chance variance, thereby suggesting that the selection of predictors in the total sample is not likely to have been determined by characteristics unique to the half-samples. Furthermore, except in studies with a very large N, the sets of predictors selected on the basis of the total sample will be more stable than the sets selected in either half-sample (see Horst, 1966, pp. 139-140, for a criticism of cross-validation based on this fact). Therefore, when the drop in R is minimal, the beta weights and shrunken (i.e., corrected for shrinkage) Rs derived in the total sample are preferable as population estimates to those values derived from the half-sample cross-validation.

In anticipation of a small drop in R upon double cross-validation, we recomputed a stepwise regression for each 16 PF and CPI scale based on the total sample. This analysis has two additional advantages. First, it provides a single estimate of the predictability for each scale (since different predictors may be selected for a given scale in the two half-samples, the two cross-validated Rs for a given scale cannot be averaged, as they can in double cross-validation of regressions based on an entire set of predictors). Second, it permits a further check on the role of chance variation, in that the sets of significant predictors derived in the total sample can be compared with those derived in the two half-samples.

Predictability Compared with Reliability

The optimal predictor of a given scale is the scale itself. Since the test-retest reliability measures a scale's predictability from itself, the adequacy of prediction was evaluated against the test-retest reliability reported in the respective manuals for each 16 PF and CPI scale.[2]

Results and Discussion

Entire Set of Predictors

The first and second columns in Table 2 contain the raw and shrunken Rs ($\hat{R}$ indicates a shrunken R) obtained by regressing each scale from the 16 PF and the CPI on the entire set of scales from the other inventory. The mean[3] of the raw Rs was .66 for the 16 PF (range = .39 to .84) and .67 for the CPI (range = .44 to .79), while the mean of the shrunken Rs was .62 for the 16 PF (range = .29 to .83) and .64 for the CPI (range = .37 to .77). The mean drop in R after correcting for shrinkage was .04 (range = .01 to .10) for the 16 PF and .03 (range = .02 to .07) for the CPI. Thus, on the average, approximately 38 percent of the variance of individual 16 PF scales and approximately 41 percent of the variance of individual CPI scales were predictable from the full set of scales on the other inventory. The estimated predictability (i.e., $\hat{R}$) of individual scales varied considerably. The predictability was .70 or greater for four 16 PF scales (*C, H, O*, Q_4) and for five CPI scales (*Do, Sy, Sp, Sc, Gi*), while it was .50 or less for four 16 PF scales (*A, B*, Q_1, Q_2) and for three CPI scales (*Re, So, Cm*).

Selected Predictors

In going from the odd to the even half-sample, the mean of the cross-validated Rs across both inventories was .59 (range = .20 to .82) for the standard double cross-validation and .60 (range = .18 to .83) for the McNemar double cross-vali-

[2]There is a problem with this plan of comparison. The 16 PF and CPI manuals report reliabilities over different intervals: a two-month interval for the 16 PF (Cattell et al., 1970, pp. 29-31) and a twelve-month interval for the CPI (Gough, 1960, pp. 19-20). Consequently, the reported reliabilities are lower for the CPI as a result of the longer interval. This discrepancy reflects the lack of consensus regarding the number and length of retest intervals over which test-retest reliabilities should be reported, which in turn results from insufficient attention by test constructors to the temporal decay function of test-retest reliabilities.

[3]All computations involving R were based on Fisher's r to z transformation.

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

Table 2

Inter-Inventory Predictability of the 16 PF and the CPI Based on the Total Sample

Predicted Scale	Standard Regression R	Standard Regression $\hat{R}$	Stepwise Regression R	Stepwise Regression $\hat{R}$	Test-Retest Reliability	Significant Predictor Scales[c]
16 PF: *A*	52	46	49	47	85[a]	*Sa, Fe, So, Py, Do*
B	39	29	31	29	63	*Ai, Cs*
C	79	77	78	78	75	*Wb, Fe, Gi, Sp, Ac, Sc**
E	68	65	64	63	85	*Do, Re, Fe, Sc, Py**
F	72	69	70	69	78	*Sp, Sc, Sy, Py, So, Ac, Re*
G	67	64	65	64	84	*Fx, Re, Sc, Fe, Cm, Ie*
H	84	83	84	83	88	*Sy, Do, Gi, Sa, Ai, Sp, Ie**
I	60	56	58	57	87	*Fe, Ai, Gi, Ac,* Cm**
L	69	66	68	67	76	*Sc, Fx, Gi, Wb, Do, Cs**
M	57	52	55	54	71	*Fx, Do, Ai,* Ac,* Gi**
N	56	51	54	52	74	*Sp, Ac, Do,* Fx,* Sc*
O	83	81	81	81	77	*Fe, Sp, Sc, Do,* Gi*
Q_1	49	42	43	41	83	*Py, Ac, Do, Fe*
Q_2	42	33	38	36	81	*Sy,* Ai, So, Fx**
Q_3	70	67	67	67	70	*Sc, Fx, Fe, Ac**
Q_4	80	78	79	78	78	*Gi, Fe, Ac, Ai, Py, Sc,* Sp**
Mean	66	62	64	63	78	
CPI: *Do*	76	74	74	74	68[b]	*H, E, G, M*
Cs	64	61	63	62	65	*H, L, M, B, E*
Sy	77	75	76	76	70	*H, F, C, M, G**
Sp	79	77	78	77	62	*H, O, F, G, N, B**
Sa	66	63	65	64	69	*H*, Q_4,* *E,* O**
Wb	71	68	70	69	72	*C, O, L, I*
Re	53	48	50	48	69	*G, C, B, A, F, M**
So	50	44	41	40	67	*C, E, A*
Sc	75	73	74	73	72	Q_4, *F, L*, Q_3, *I, C, E*, Q_1*
To	63	59	59	58	66	*L, C, I, B*
Gi	77	75	77	76	69	Q_4, *I, H, L*, Q_3, *M, E,*
Cm	44	37	40	39	41	*G, I, F**
Ac	62	58	59	57	67	*C, G, B, N, O,* A, F**
Ai	61	57	55	54	60	*L, M, B, G*
Ie	68	65	65	65	76	*O, C*, Q_4, *L, B*
Py	66	63	62	61	49	*O, C*, Q_1, *L, A, E**
Fx	70	67	67	66	64	*G, L*, Q_3, *M**
Fe	71	68	68	68	62	*I, O, A, E*, Q_4
Mean	67	64	65	64	65	

Note. - Decimal points omitted, All analyses based on total sample.

[a]Two-month interval, N = 132 (Cattell et al., 1970, p. 31).

[b]One-year interval, mean of 101 male and 125 female high school students (Gough, 1960, p. 19).

[c]Listed in order of extraction. Zero-order correlations with the predicted scales are available from the authors.

*Predictor scales not selected in either half-sample.

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

Table 3
Results of the Cross-Validations

Predicted Scale		Odd-Half to Even-Half			Even-Half to Odd-Half		
		$R_{(odd)}$[a]	$R_{St(even)}$[b]	$\hat{R}_{Mc(even)}$[c]	$R_{(even)}$[a]	$R_{St(odd)}$[b]	$\hat{R}_{Mc(odd)}$[c]
16 PF:	*A*	52	38	40	52	33	30
	B	37	20	18	23	25	24
	C	75	78	78	81	72	72
	E	66	53	52	63	58	62
	F	70	66	67	73	58	57
	G	65	55	54	65	57	62
	H	80	80	83	86	80	80
	I	54	53	52	58	49	49
	L	66	65	65	66	59	57
	M	63	37	36	49	46	59
	N	48	38	36	52	34	32
	O	79	82	82	84	76	75
	Q_1	39	32	29	51	27	26
	Q_2	33	21	18	37	12	19
	Q_3	65	66	66	73	47	61
	Q_4	74	77	77	81	72	72
Mean		62	57	57	65	53	55
CPI:	*Do*	73	74	76	76	68	68
	Cs	61	55	54	63	59	59
	Sy	72	78	78	81	69	70
	Sp	77	78	79	79	67	68
	Sa	62	61	61	66	57	58
	Wb	69	70	69	72	67	67
	Re	53	42	41	46	39	37
	So	42	31	32	40	27	24
	Sc	67	76	76	77	66	65
	To	55	52	56	72	46	47
	Gi	72	76	77	81	70	70
	Cm	36	32	31	43	31	34
	Ac	57	51	54	58	41	42
	Ai	61	50	50	61	47	51
	Ie	57	60	61	70	55	54
	Py	63	51	54	63	52	51
	Fx	76	57	55	55	73	73
	Fe	72	64	67	70	59	61
Mean		64	61	62	67	57	57

Note. - Decimal points omitted.

[a] *R*s generated by the stepwise regressions in the original half-sample.

[b] *R*s generated through the standard cross-validation technique.

[c] $\hat{R}$s generated through McNemar's cross-validation procedure.

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

dation (see the second and third columns in Table 3). Similarly, in going from the even to the odd half-sample, the mean of the cross-validated *R*s was .55 (range = .12 to .80) for the standard and .56 (range = .19 to .80) for the McNemar cross-validations (see the fifth and sixth columns in Table 3). Thus, the results from the two techniques are highly comparable, and this similarity is indicated further by the fact that, across all 68 pairs (i.e., twice the total of 34 pairs) of cross-validated *R*s, each technique generated the higher of the two *R*s an equal number of times. Furthermore, in only 4 of the 68 possible comparisons did the discrepancy between the two techniques exceed .04. Since the results of the two different double cross-validation techniques were generally quite similar, it can be said that from an empirical standpoint, the choice between the two techniques is of little consequence.

When the cross-validational data were averaged across both techniques and directions (i.e., odd to even and even to odd) of cross-validation, the mean drop in *R* due to cross-validation was .06 (i.e., .64 to .58).[4] This drop seemed reasonably small, although no statistical test was applied. We conclude that the values obtained by computing the stepwise regressions on the total sample would be appropriate estimates of predictabilities (see the Method section for the rationale).

Accordingly, the raw and shrunken *R*s were computed for the total sample from the stepwise regression of each 16 PF and CPI scale on the scales from the other inventory. These values are listed in the third and fourth columns of Table 2. The mean of the raw *R*s was .64 for the 16 PF (range = .31 to .84) and .65 for the CPI (range = .40 to .78), while the mean of the *R*s was .63 for the 16 PF (range = .29 to .83) and .64 for the CPI (range = .39 to .77). The mean raw *R* for both the 16 PF and the CPI was only .02 less than the corresponding mean value based on the full set of predictors, and the $\hat{R}$ was virtually identical, in mean and range, to the corresponding mean value based on the full set of predictors. After correcting for shrinkage, the discrepancy between the predictabilities based on the entire and the selected sets of predictors exceeded .02 for only 3 of the 34 scales. From these considerations, we conclude that the predictabilities based on the selected sets of predictors were comparable to those obtained with the entire sets of predictors. There is no reason to rely on the full set of predictor scales when selected sets of predictors lead to virtually identical conclusions. Therefore, we subsequently will consider only results based on the selected sets of predictors.

On the average, approximately 40 percent of the variance of individual 16 PF scales and approximately 41 percent of the variance of individual CPI scales were predictable from the selected sets of scales on the other inventory. As with the full set of predictors, estimated predictabilities (i.e., $\hat{R}$s) were .70 or greater for four 16 PF (*C, H, O,* $\bar{Q}_4$) and five CPI (*Do, Sy, Sp, Sc, Gi*) scales, while $\hat{R}$ was .50 or less for four 16 PF scales (*A, B,* Q_1, $\bar{Q}_2$) and three CPI scales (*Re, So, Cm*).

Table 2 also lists the significant predictor scales for each 16 PF and CPI scale, as identified by the stepwise equations based on the total sample. For each predicted scale, every predictor which was significant in both half-samples was also significant in the total sample. Furthermore, there were only a few scales which were significant primary predictors in the total sample but not in either of the half-samples (see those scales with an asterisk in Table 2). These findings support the contention that the sets of significant predictors identified in the total sample will be more stable than those identified in the half-samples. However, the sets of significant predictors identified within the odd and even half-samples were rather discrepant; in fact, there was no scale for which exactly the same set of predictors was identified in the two

[4]Unless a total sample is unusually homogeneous, one of its half-samples will tend to be more predictable than the other. In the present study, the even half-sample was slightly more predictable than the odd half-sample. However, the discrepancies were minimal, and we attributed the differences between the half-samples to chance variation.

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

Table 4
Predictability of the 16 PF and CPI Scales as a Function of the Number of Significant Predictor Scales Common to Both Half-Samples

Number of Significant Common Predictors	16 PF		CPI	
	Scale(s)	Median $\hat{R}$[a]	Scale(s)	Median $\hat{R}$[a]
0	*B* Q_1 Q_2	.36	None	
1	*A I L*	.57	*Sa So To Cm Ae Ie Py*	.58
2	*E F G M N* Q_3 Q_4	.64	*Do Cs Sy Sp Wb Re Ai Fe*	.68
3	*C H*	.80	*Gi Fx*	.71
4	*O*	.81	*Sc*	.73

[a]Derived from the stepwise regressions based on the total sample.

half-samples. When several scales from one inventory are substantially intercorrelated, the selection of any particular one of them as a predictor in a half-sample will be influenced by marginal characteristics of the half-sample. Indeed, approximately one-sixth of the values in the 16 PF intercorrelation matrix (Cattell et al., 1970, p. 113) and approximately one-third of the values in the CPI intercorrelation matrix (Gough, 1960, p. 40) equal or exceed .40. Thus, the half-sample differences in the sets of significant predictors might have been expected.

As shown in Table 4, the predictability of scales increased as the number of significant predictors common to the two half-samples increased. However, since the rank-order correlations between the predictability of a scale and the total number of significant predictors for that scale were only .48 for the 16 PF and .36 for the CPI, this relationship is not simply a function of an increasing number of predictors. Rather, the number of predictors common to both half-samples seems to reflect the extent of overlap between the predicted scale and the predictor inventory.

Predictability Compared with Reliability

As shown in the fourth and fifth columns of Table 2, the $\hat{R}$ from the stepwise regressions based on the total sample equalled or exceeded the reported test-retest reliability for three 16 PF scales (*C, O,* Q_4) and for eight CPI scales (*Do, Sy, Sp, Sc, Gi, Py, Fx, Fe*). In addition, the $\hat{R}$ approached the reliability (i.e., discrepancy $\leq$ .05) for two more 16 PF (*H,* Q_3) and four more CPI (*Cs, Sa, Wb, Cm*) scales. These data do not suggest that the CPI scales are predicted more adequately than the 16 PF scales, since the mean predictabilities of the 16 PF and the CPI are virtually identical. The apparent discrepancy between these two sets of findings is due to the longer test-retest interval, and consequent lower reliabilities, for the CPI (i.e., a twelve-month interval for the CPI vs. a two-month interval for the 16 PF). If a scale is designated arbitrarily as "highly predictable" when its $\hat{R}$ equals or exceeds .70 and also approaches or exceeds its test-retest reliability, then four 16 PF scales (*C, H, O,* Q_4) and five CPI scales (*Do, Sy, Sp, Sc, Gi*) can be considered highly predictable.[5]

[5]The correlation between $\hat{R}$ and test-retest reliability was .23 (n.s.) for the 16 PF and .46 ($p < .05$) for the CPI. Thus, predictability is related to stability for the CPI, but not for the 16 PF. On the other hand, the correlation between $\hat{R}$ and coefficient alpha (calculated on the present sample) was .70 ($p < .01$) for the 16 PF and .37 (n.s.) for the CPI. Thus, there is a very strong relationship between predictability and internal consistency for the 16 PF scales, but not for the CPI scales. When one considers that the CPI scales were constructed on the basis of external validity rather than internal consistency, while the reverse was true for the 16 PF scales, it is striking that the mean of coefficient alphas for the CPI scales was .66 compared to .65 for the 16 PF scales. (Respective ranges were .39 to .80 and .30 to .88.)

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

Overlap between 16 PF and CPI

Having examined various aspects of the predictability findings, we now proceed to consider the problem of the common and unique coverage of the 16 PF and the CPI. The degree of content overlap between the two inventories will be examined at the level of both individual scales and their underlying factors.

Overlap at the scale level. There were five instances in which a 16 PF scale and a CPI scale were the "primary predictors" of one another (i.e., the first predictor selected in the stepwise regression): 16 PF-*C* (Emotionally stable) and CPI-*Wb* (Sense of well-being), 16 PF-*G* (Conscientious) and CPI-*Fx* (Flexibility), 16 PF-*H* (Venturesome) and CPI-*Sy* (Sociability), 16 PF-*I* (Tenderminded) and CPI-*Fe* (Femininity), 16 PF-Q_4 (Tense) and CPI-*Gi* (Good impression). In addition, nearly perfect reciprocal predictability (i.e., one scale was the first predictor for another scale, while the latter was the second best predictor for the former scale) was present in four more instances: 16 PF-*E* (Assertive) and CPI-*Do* (Dominance), 16 PF-*H* (Venturesome) and CPI-*Do* (Dominance), 16 PF-*O* (Apprehensive) and CPI-*Fe* (Femininity), 16 PF-*G* (Conscientious) and CPI-*Re* (Responsibility). Seven scales in each inventory (i.e., *C, E, G, H, I, O,* and Q_4 for 16 PF and *Do, Sy, Wb, Re, Gi, Fx,* and *Fe* for CPI) then were reciprocally predictable. Since the scales in each of these nine pairs tend to refer to similar attributes, the demonstrated reciprocal predictability is hardly surprising. Thus, the correspondence between the 16 PF and the CPI at the individual scales level was considerable, though not all the anticipated reciprocal relationships emerged. For instance, 16 PF-*B* (Bright) and CPI-*Ie* (Intellectual efficiency) showed no systematic relationship.

The factor analytic study by Edwards and Abbott (1973b) provides corroboration for the pairwise reciprocal predictabilities described above. For each of the 22 factors extracted from the combined correlation matrix of the 16 PF, CPI, EPI, and EPPS, these investigators reported the scale with the highest loading from each of the four inventories. On four of these 22 factors, the 16 PF and CPI marker scales had loadings of .40 or greater. Three of these four pairs of marker scales are reciprocally predictable in the present study (i.e., *G* and *Fx, I* and *Fe,* and *H* and *Do*). In addition, Edwards and Abbott identified one factor specific to the CPI (defined by *Cm*) and two factors specific to the 16 PF (defined by *B* and by *N*). In the present study, the definers of these inventory-specific factors, as expected, were among the least predictable scales. That is, *Cm* was the least predictable of the CPI scales ($\hat{R}$ = .39), while *B* and *N* had the lowest and the fifth lowest predictabilities among the 16 PF scales ($\hat{R}$ = .29 and .52, respectively).

Overlap at the factor level. The degree of overlap between the 16 PF and the CPI also is considerable at the factor level. Two predominant factors have emerged consistently in factor analyses of the 16 PF scales (Cattell et al., 1970, pp. 115-121) and of the CPI scales (Megargee, 1972, pp. 110-115). For the 16 PF, these two factors are "Introversion vs. Extroversion" (defined by *A, E, F, H,* and Q_4) and "Adjustment vs. Anxiety" (defined by *C, H, L, O,* Q_3 and Q_4). For the CPI, the two factors are "Extroversion" (defined by *Do, Cs, Sy, Sp,* and *Sa*) and "Adjustment" (defined by *Sc, Gi, Wb, To,* and *Ac*). Of the five "highly predictable" CPI scales, three scales (*Do, Sy,* and *Sp*) are markers for the CPI "Extroversion" factor, while the remaining two scales (*Sc* and *Gi*) are the predominant markers for the "Adjustment" factor. Similarly, all four of the "highly predictable" 16 PF scales (*C, H, O,* and Q_4) load highly on the "Adjustment vs. Anxiety" factor, while *H* is also a marker for "Introversion vs. Extroversion." The fact that the highly predictable scales are definers of the two major factors of each inventory suggests an extensive overlap in the coverage of the two inventories. Based on these observations, one might predict a substantial correlation between the "Introversion vs. Extroversion" and the "Extroversion" factor scores, and likewise between the "Adjustment vs. Anxiety" and the "Adjustment" factor scores.

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction
requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

The overlap between the major factors of the two inventories is also apparent when one examines the extent to which the markers of a factor serve as the predictors for the markers of the corresponding factor in the other inventory. That is, for all five markers of the CPI "Adjustment" factor, the primary predictor (i.e., the first predictor in the stepwise regression equation) and at least two additional predictors come from the marker scales of the 16 PF "Adjustment vs. Anxiety" factor (see Table 2). Likewise, the primary predictor for all six markers of the 16 PF "Adjustment vs. Anxiety" factor except *H* and *O* comes from the marker scales for the CPI "Adjustment" factor, while at least two predictors for all six markers of the 16 PF "Adjustment vs. Anxiety" factor except *H* come from among the marker scales for the CPI "Adjustment" factor.

In a similar vein, the 16 PF "Introversion vs. Extroversion" factor overlaps with the CPI "Extroversion" factor. For all five markers of the CPI "Extroversion" factor the primary predictor and one additional predictor come from among the markers of the 16 PF "Introversion vs. Extroversion" factor. For all five markers of the 16 PF "Introversion vs. Extroversion" factor the primary predictor comes from among the markers of the CPI "Extroversion" factor. However, note that 16 PF-*H* is the primary predictor for all five markers of CPI "Extroversion" but for no other CPI scale, and that four of the five markers of CPI "Extroversion" are predictors of 16 PF-*H*. It appears that the close linkage of 16 PF-*H* and CPI "Extroversion" is responsible for the overlap between the 16 PF "Introversion vs. Extroversion" and the CPI "Extroversion" factors. This finding, together with the discrepancy between *H* and the remaining markers of 16 PF "Adjustment vs. Anxiety" noted in the preceding paragraph, suggests that *H* should be considered primarily as a marker of "Introversion vs. Extroversion," rather than as a marker of both 16 PF factors. (This suggestion is quite consistent with the higher loading of *H* on the "Introversion vs. Extroversion" factor than on the "Adjustment vs. Anxiety" factor. See Cattell et al., 1970, p. 121.) Therefore, one might conclude that the 16 PF "Introversion vs. Extroversion" factor is tapping a similar but somewhat broader domain than the CPI "Extroversion" factor.

Mitchell's (1963) joint factor analysis of the 16 PF and CPI scales also indicates a correspondence between the two major factors of the 16 PF and of the CPI (see Mitchell's Table 2, p. 159). The five CPI and five 16 PF scales with the highest loadings on his Factor I ("General Adjustment") are the scales which define the CPI "Adjustment" and the 16 PF "Adjustment vs. Anxiety" factors. (However, note that 16 PF-*H* loads only .29 on Factor I, compared to .79 on Factor II. This provides additional support for our suggestion that *H* is a marker only for 16 PF "Introversion vs. Extroversion.") Similarly, the five CPI and the five 16 PF scales with the highest loadings on his Factor II ("Extroversion") are those scales which define the CPI "Extroversion" and 16 PF "Introversion vs. Extroversion" factors. In addition, the loadings of the CPI markers on his Factor II are generally higher than the loadings of the 16 PF markers with the exception of scale *H*. This fact supports our interpretation that the CPI "Extroversion" factor is linked to the 16 PF "Introversion vs. Extroversion" factor largely through 16 PF-*H*. It is interesting to note that 16 PF-*I* and CPI-*Fe* have the highest loadings (.71 and .69) on Mitchell's Factor IV, while 16 PF-*G* and CPI-*Fx* have the highest loadings (.65 and .70) on his Factor V. (Note that reciprocal predictability of these pairs previously has been demonstrated in the "overlap at the scale level" section).

Conclusion

The general findings from the present study can be summarized as follows. 1) The selected sets of predictors (i.e., the significant predictors identified in the stepwise regressions) were essentially as powerful as the entire sets of predictors. 2) Despite the differences in their underlying logic, the extended McNemar and the standard double cross-validation techniques generated very similar estimates of predictabilities. 3)

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction
requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

Within both inventories, the predictability of individual scales varied considerably: four 16 PF scales and five CPI scales were "highly predictable," while four 16 PF scales and three CPI scales were non-predictable ($\hat{R} < .50$). The mean predictability was .63 for the 16 PF and .64 for the CPI. The fact that the two inventories can be predicted equally well from one another is at variance with the assertion that the 16 PF, in contrast to other inventories, measures the basic, hence the "source," traits of personality. Rather, it seems that the traits measured by the CPI scales are as basic as those measured by the 16 PF scales. 4) Seven scales in each inventory are reciprocally predictable and hence can be said to have counterparts in the other inventory. 5) There is a substantial content overlap between the 16 PF "Adjustment vs. Anxiety" and the CPI "Adjustment" factors. The overlap between the 16 PF "Introversion vs. Extroversion" and the CPI "Extroversion" factors is also considerable but appears to be due to the close linkage of the 16 PF-*H* and the CPI "Extroversion" factor.

The feasibility of reciprocal prediction between the 16 PF scales and the CPI scales is demonstrated well in the present study. It now remains for this study to be replicated with sex-specific samples, and for the resulting prediction equations to be cross-validated. The present study examined reciprocal prediction at the group level, and as such it needs to be supplemented by studies of predictability at the level of individual profiles. Finally, the multiple regression approach used in the present study needs to be compared with alternative techniques of inter-inventory prediction.

References

Abbott, R. D. A factor analysis of the CPI and EPI. *Educational and Psychological Measurement,* 1971, *31,* 549-553.

Allen, R. M. The relationship between the Edwards Personal Preference Schedule variables and the Minnesota Multiphasic Personality Inventory scales. *Journal of Applied Psychology,* 1957, *41,* 307-311.

Backer, T. E., Comrey, A. L., & Hahn, M. E. Comparison of the Comrey Personality Scales and the California Life Goals Evaluation Schedules. *Psychological Reports,* 1971, *29,* 513-514.

Borgatta, E. F. The coincidence of subtests in four personality inventories. *Journal of Social Psychology,* 1962, *56,* 227-244.

Cattell, R. B., & Bolton, L. S. What pathological dimensions lie beyond the normal dimensions of the 16 PF? A comparison of MMPI and 16 PF factor domains. *Journal of Consulting Psychology,* 1969, *33,* 18-27.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. *Handbook for the sixteen personality factor questionnaire.* Champaign, Ill.: Institute for Personality and Ability Testing, 1970.

Comrey, A. L., & Duffy, K. E. Cattell and Eysenck factor scores related to Comrey personality factors. *Multivariate Behavioral Research,* 1968, *3,* 379-392.

Draper, N. R., & Smith, H. *Applied regression analysis.* New York: Wiley, 1966.

Dunnette, M. D., Kirchner, W. K., & DeGidio, J. Relations among scores on Edwards Personal Preference Schedule, California Psychological Inventory, and Strong Vocational Interest Blank for an industrial sample. *Journal of Applied Psychology,* 1958, *42,* 178-181.

Edwards, A. L., & Abbott, R. D. Relations among the Edwards Personality Inventory, the Edwards Personal Preference Schedule, and the Personality Research Form. *Journal of Consulting and Clinical Psychology,* 1973, *40,* 27-32. (a)

Edwards, A. L., & Abbott, R. D. Relationships between the EPI scales and the 16 PF, CPI, and EPPS scales. *Educational and Psychological Measurement,* 1973, *33,* 231-238. (b)

Edwards, A. L., Abbott, R. D., & Klockars, A. J. A factor analysis of the EPPS and PRF personality inventories. *Educational and Psychological Measurement,* 1972, *32,* 23-29.

Goldberg, L. R. A historical review of personality scales and inventories. In P. McReynolds (Ed.), *Advances in psychological assessment: Volume two.* Palo Alto: Science and Behavior Books, 1971.

Gough, H. G. *Manual for the California Psychological Inventory.* Palo Alto, Calif.: Consulting Psychologist Press, Inc., 1960.

Horst, P. An overview of the essentials of multivariate analysis methods. In R. B. Cattell (Ed.). *Handbook of multivariate experimental psychology.* Chicago: Rand McNally, 1966.

Hundleby, J. D., & Connor, W. H. Interrelationships between personality inventories: The 16 PF, the MMPI, and the MPI. *Journal of Consulting and Clinical Psychology,* 1968, *32,* 152-157.

Jackson, D. N. *Personality Research Form manual.* Goshen, N. Y.: Research Psychologists Press, Inc.,

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/

1967.
Karson, S., & Pool, K. B. The construct validity of the Sixteen Personality Factors test. *Journal of Clinical Psychology,* 1957, *13,* 245-252.
LaForge, R. A correlational study of two personality tests: The MMPI and Cattell 16 PF. *Journal of Consulting Psychology,* 1962, *26,* 406-411.
Lemke, E. A., & Kirchner, J. H. A multivariate study of handwriting, intelligence and personality correlates. *Journal of Personality Assessment,* 1971, *35,* 584-594.
McNemar, Q. *Psychological statistics.* (4th ed.) New York: Wiley, 1969.
Megargee, E. I. Estimation of CPI scores from MMPI protocols. *Journal of Clinical Psychology,* 1966, *22,* 456-458.
Megargee, E. I. *The California Psychological Inventory handbook.* San Francisco: Jossey-Bass, Inc, 1972.
Merrill, R. M., & Heathers, L. B. The relation of the MMPI to the Edwards Personal Preference Schedule on a college counseling center sample. *Journal of Consulting Psychology,* 1956, *20,* 310-314.
Mitchell, J. V. A comparison of the first and second order dimensions of the 16 PF and CPI inventories. *Journal of Social Psychology,* 1963, *61,* 151-166.
Mosier, C. I. Problems and designs of cross-validation. *Educational and Psychological Measurement,* 1951, *11,* 5-11.
Norman, W. T. Double-split cross-validation: An extension of Mosier's design, two undesirable alternatives, and some enigmatic results. *Journal of Applied Psychology,* 1965, *49,* 348-357.
O'Dell, J. W., & Karson, S. Some relationships between the MMPI and 16 PF. *Journal of Clinical Psychology,* 1969, *25,* 279-283.
Rodgers D. A. Estimation of MMPI profiles from CPI data. *Journal of Consulting Psychology,* 1966, *30,* 89.
Trott, D. M., & Morf, M. E. A multimethod factor analysis of the Differential Personality Inventory, Personality Research Form, and Minnesota Multiphasic Personality Inventory. *Journal of Counseling Psychology,* 1972, *19,* 94-103.
Weiss, D. J. Canonical correlation analysis in counseling psychology research. *Journal of Counseling Psychology.* 1972, *19,* 241-252.
Wiggins, J. S. *Personality and prediction: Principles of personality assessment.* Reading, Mass.: Addison-Wesley, 1973.
Wiggins, J. S., Goldberg, J. R., & Appelbaum, M. MMPI content scales: Interpretive norms and correlation with other scales. *Journal of Consulting and Clinical Psychology,* 1971, *37,* 403-410.

Acknowledgements

The 16 PF and CPI data were made available by Dr. Henry A. Alker and prepared for analysis by Drs. William R. Revelle and James A. Kulik. Herbert Z. Wong collaborated in the early stages of this study. This project was supported in part by the National Institute for Mental Health Grant 1 R01 MH16974-04 (S. Cobb, K. T. Chun, and J. R. P. French, Jr.). We are grateful for the assistance and comments provided by Sidney Cobb and John R. P. French, Jr.

Author's Address

Ki-Taek Chun, U. S. Commission on Civil Rights, Office of Research, 1121 Vermont Avenue, N. W., Washington, D. C. 20425.

Downloaded from the Digital Conservancy at the University of Minnesota, http://purl.umn.edu/93227.
May be reproduced with no cost by students and faculty for academic use. Non-academic reproduction requires payment of royalties through the Copyright Clearance Center, http://www.copyright.com/