INITIAL PSYCHOMETRIC EVALUATION OF URDU VERSION OF CALIFORNIA PSYCHOLOGICAL INVENTORY (CPI)

Iftikhar Ahmed

National Institute of Psychology, Islamabad

Adapted and translated in Urdu, the California Psychological Inventory (CPI) was examined on a number of psychometric dimensions. Bilingual subjects were administered the Urdu and English versions to study their equivalence. The results show the Urdu version to be sufficiently similar to the English version. The differences between the means were statistically significant ($p < .05$) on six scales: Wb, Re, Sc, Gi, To, and Le. Relevant explanations were advanced to account for these differences. On the Urdu version, males and females had comparable scores except on the Femininity scale. Means for the Urdu version were generally distinctly lower than for the American norms. This reflects on the cultural differences affecting response rates, and indicates the need to norm the test in Pakistan. KR-20 indices of reliability of the scales ranged from .44 to .93, with a median of .68, which is fairly comparable and even better than values reported for American samples. The matrix of interscale correlations was found to be theoretically consistent and meaningful, verifying that on the Pakistani sample the CPI assesses a set of positively related variables. The overall psychometric evaluation lends reasonably sufficient credence to further use of the CPI in Pakistan and to research on predictive and classificatory issues.

California Psychological Inventory (Gough, 1956) is one of the most widely used objective type personality tests. The volume of its published research literature is second only to MMPI and possibly 16 PF. As an inventory of quite a sufficient bandwidth, it has 18 bipolar scales (see table 3) of which eleven have been empirically developed through comparison of defined criterion groups, four were developed rationally and were further refined through the consistency response analysis and the remaining three scales were derived to detect the tendency to dissimulate or respond randomly. Largely an empirically developed instrument, it has been reasonably successful in its use in every day practical situations.

Harrison Gough, the author of CPI takes this inventory as a measure of variables important to social living, interaction, and adjustment. He emphasizes the need for having tests which are relevant to actual life situations across cultures and claims CPI to be the one which taps qualities that are cross-culturally relevant and universally recognized. He calls CPI variables 'Folk Concepts', that is, terms which are extensively applicable to normal human beings everywhere. The use of CPI across nations and cultures has attested to the universality of the folk-constructs employed in the test. Since these folk concepts emerge from social interaction and social living,

they have a kind of immediate relevance which enhances their attractiveness and social diagnostic value.

CPI has been translated in many languages including French, German, Greek, Italian, Japanese, Mandarin, and Spanish, making available volumes of cross-cultural research literature. This tempted this investigator to assess the scope of the application of the CPI in Pakistani society also. Obviously use of CPI as a test in English language would be limited only to urban and educated population here, therefore, the test had to be adapted and translaed into Urdu, which is the national language of Pakistan. The permission of the author and publisher of the test was therefore sought for making an Urdu version of the test and using it in Pakistan for research purposes.

There is no standard prescription for the task of translating and adapting verbal material, specially for the affective measures, nor is there any objective gauge to assess the equivalence of the translated version. In fact, the whole set of linguistic idiosyncracies and cultural differences between the source and the target language plus the measurement problem in the test itself complicates and confounds the cross-national use of personality questionnaires. Each such task, therefore, becomes a unique experiment and a lot of exploratory work is required to think about the possible nature of problems involved in it. Nonethless, the test was first translated by bilingual committee approach following the guidelines of Brislin (1973) and Butcher and Pancheri (1976). Emphasis was laid on conceptual equivalence in order to provide for common meanings and legitimate comparisons between the original and the target texts. Subsequently, back translation procedure was also used to improve the equivalence of the Urdu version with the original test. Finally, Urdu CPI was field-tested and indices of empirical equivalence between the two versions were estimated. The differences in endorsement of an item as "true", on the two versions, were found to be statistically significant (at .01 level) for only 25 percent of the items of the inventory. A difference was considered significant when it was 20 percent. For details see Ahmad (1985).

While reading CPI, one forms the impression that its contents sampled middle class thinking, social values and outlook in a largely Westernised life style.However, most of the items were observed to be etic and so cross-culturally relevant here also, at least to middle class educated population, who by virtue of their Europian-like system of education and its spirit would understand and appreciate the contents of the test. The few items which had foreign cultural contents nevertheless demand quite some sophistication to understand them and it is intuitively felt that the test could only be used with subjects who have had at least ten years of education.

The specific purpose of this study was to estimate the psychometric properties of the Urdu CPI and to assess whether it qualifies as a test of some

potential use in the population of our interest. The salience of this test would depend upon its goodness vis-a-vis the response data characetristic of the subjects of this study. Standard psychometric procedures have been employed here in evaluating the Urdu CPI in order to determine its usefulness as an objectively scorable personality test, to be used in Pakistan.

METHOD

Sample

An adhoc (nonrandom) group of 76 college students (14th year in education) were initially employed for this study. Of these, 70 (37 boys, 33 girls) completed the work. Primarliy, the subjects had been selected as bilinguals in English and Urdu. The procedure of determining the status of the subjects as bilinguals could not be very rigorous, as proficiency tests were not available in the two languages to select any traditionally defined bilingual gruop, however, the subjects employed, had qualified Higher Secondary School Examination in which both English and Urdu languages are compulsory subjects. They had opted for English as their medium of instruction at the college level and had Urdu as their first language. This was held suffice for the bilingual requirements of the task.

Procedure

Subjects were administered the test at their respective colleges during the class hours. It was difficult to have test-retest arrangement for the administration of both the versions on the same subjects because of students' time constraint. Thus alternate CPI versions were administered to subjects. They were first briefed about the research study with regards to its objectives. The protocols were subsequently scored after the standard key for both the versions and appropriate procedures were adopted to carry out the analysis of the data on the parameters important in evaluating the test. A summary of some of the findings is given in the following section.

RESULTS AND DISCUSSION

Validity of the Pakistani data

In a test alien to Pakistani society, one major concern in the use of the tests would be to see whether or not the test contents are properly understood and responded to; subjects know how to take the test and write answers in the required style; have proper test-taking attitude, etc. Data were explored to assess some of these matters. Rate of omission of response was tabulated (see table 1) which was found to be just negligible indicating that the testees did respond to the test contents and they tended to answer the questionnaire categorically and the items seemed to be working well in this group of subjects.

Next, table 2, presents the rate of "true" response given to each of the statements in the test. It shows that the obtained response pattern does not smack of any response set operative in the data.

Table 1

Frequency of omission of responses for Urdu (CPI) items (N=37)

Frequency of Omissions	Frequency of Items
0	245
1	112
2	78
3	25
4	11
5	4
6	3
7	2

Table 2

Percentage of subjects giving response "true" to items on Urdu CPI

Percentage "True"	Frequency of items
96 – 100	16
91 – 95	20
86 – 90	25
81 – 85	24
76 – 80	29
71 – 75	24
66 – 70	29
61 – 65	27
56 – 60	25
51 – 55	29
46 – 50	25
41 – 45	30
36 – 40	28
31 – 35	30
26 – 30	19
21 – 25	16
16 – 20	21
11 – 15	23
6 – 10	14
4 – 5	9

Scalar Equivalence between the English and the Urdu Versions

To assess metric equivalence at the level of scales, mean scores were calculated which were found to be quite comparable between the two versions, except on six scales, namely, Wb, Re, Sc, To, Gi and Ie (see table 3), where differences between these indices were found to be statistically significant ($P<.05$). This corroborates the item-analysis done on the two versions in that the items of these six scales showed clear differences in response rates. It is interesting to note that this set of scales formed factor 1 in most of the factor analytic studies (Mitchell and Pierce-Jones, 1960; Bouchard, 1969; Nichols and Schnell,1963). This factor has been referred to as a measure of 'general adjustment', as a measure of 'intrapersonal beliefs', and as a 'means of cultural values'. Hence, it is not surprising that these are the scales showing major differnces between the two versions.

Problems of this kind (linguistic and value differences) are in fact inherent in the bilingual research method (Bond and Yang, 1980).

As against the English version, the scores on the Urdu version were moderately enhanced on almost all the scales. The average scores on the scales of the Urdu CPI adequately corresponded to the mid-value of the number of items set for the scales (see table 3) which means that Urdu version would have pretty fair discriminative ability in this population.

Sex-Differences and Test Scores

Scores of male and female subjects were also well comparable on the Urdu version (see table 3). The difference between the mean scores of the sexes was statistically significant on none of the scales, except, as expected, on 'Femininity' scale ($P<.01$).This attests to the validity of the said scale which seeks to differentiate between males and females, and to define a personological syndrome that can be properly conceptualized as 'feminine' at one pole and 'masculine' at the other. The validity of the 'Fe' sclae has, therefore, been confirmed here also, as in several other cross cultural investigations (see Gough et al. 1968; Levin & Darani, 1971; Nishilyama, 1975; Pitariu, 1981).

Reliability Evidence

As reliability is one of the most important properties of an objective personality test, KR-20 estimates were calculated to assess the internal consistency and homogeneity of the CPI scales (table 4). The obtained indices of reliability estimates ranged from .44 to .93 with a median value of 68, which is satisfactory for CPI as a largely externally criterioned test.

Table 3

Mean, Standard Deviation & T-Test for Scores on CPI.

	n	VERSION Urdu (N=37)	VERSION English (N=37)	t	URDU VERSION Female (N=19)	URDU VERSION Male (N=18)	t
Do	46	23.81 (6.16)	22.15 (5.11)	1.20	24.06 (6.13)	23.57 (6.19)	.23
Cs	32	10 (4.16)	10.03 (4.28)	.02	9.72 (3.82)	10.79 (3.87)	.82
Sy	36	21.24 (4.89)	20.18 (4.31)	.09	21.28 (5.32)	21.22 (4.48)	.04
Sp	56	25.83 (5.56)	25.33 (5.32)	.37	25.11 (4.81)	26.53 (6.08)	.76
Sa	34	17.18 (3.51)	17.06 (3.26)	.14	16.94 (3.75)	17.42 (3.27)	.40
Wb	44	30.16 (6.11)	26.3 (6.24)	2.57**	30.83 (5.69)	29.53 (6.41)	.63
Re	42	26.86 (3.89)	22.51 (4.72)	4.16**	27.22 (2.57)	26.53 (4.87)	.53
So	54	33.75 (5.29)	32.60 (4.15)	.98	34.67 (5.44)	32.89 (4.87)	1.04
Sc	50	26.72 (9.01)	23.00 (5.96)	2.00*	28.22 (8.30)	25.32 (9.20)	1.97
To	32	15.59 (3.98)	12.78 (4.80)	2.66**	15.22 (3.69)	15.42 (4.33)	1.47
Gi	40	20.16 (5.99)	15.69 (4.97)	3.37**	20.56 (5.09)	19.79 (6.58)	.38
Cm	28	22.05 (2.62)	20.70 (3.03)	1.67	22.72 (1.88)	21.32 (3.10)	1.60
Ac	38	24.56 (4.73)	22.75 (3.27)	1.79	26.00 (3.87)	23.21 (4.95)	1.85
Ai	32	13.78 (8.01)	12.63 (3.52)	.69	13.67 (2.85)	13.89 (3.23)	.21
Ie	52	32.29 (5.73)	27.09 (5.70)	3.79**	32.89 (5.47)	31.74 (5.76)	.60
Py	32	9.32 (2.67)	8.06 (2.31)	.40	9.72 (2.18)	8.95 (2.96)	.87
Fx	22	5.40 (3.45)	5.48 (2.38)	.11	4.83 (2.54)	5.95 (3.98)	.08
Fe	38	18.08 (3.80)	18.03 (4.51)	.68	19.05 (4.18)	16.21 (2.95)	2.70**

Standard deviations are given in parenthesis. $*P < .05$; $**P < .01$

Table 4

Internal Consistency Coefficient of 18 scales of CPI

	Freshmen Women[a] (N = 179)	High School Boys[a] (N = 3572)	High School Girls[b] (N = 4056)	Pakistani results College Boys & Girls[c] (N = 37)
Do	.80	.70	.71	.77
Cs	.59	.61	.68	.66
Sy	.80	.74	.75	.69
Sp	—	.74	.75	.66
Sa	—	.51	.58	.44
Wb	—	.76	.79	.83
Re	.67	.72	.70	.51
So	.78	.68	.67	.66
Sc	—	.82	.85	.89
To	.74	.74	.75	.65
Gi	—	.77	.77	.79
Cm	—	.70	.52	.51
Ac	.65	.69	.94	.69
Ai	.63	.54	.56	.93
Ie	.72	.81	.74	.72
Py	.44	.22	.23	.44
Fx	—	.56	.51	.72
Fe	.30	.62	.29	.47

[a] Data from Hase and Goldberg (1967) using KR-20 reliability estimates.
[b] KR-20 estimates computed by Edwin I. Megargee (1977) from data in the CPI manual.
[c] Computed from the present data using KR-20 estimates.

These estimates are also fairly comparable with the American data (see table 4). Interestingly, KR-20 index of homogeneity of most of the externally criterioned scales was as good and oddly enough, even better than that of the scales which were developed by internal consistency technique. As an explanation for this observation, Farly and Cohen (1980), can well be mentioned who in a similar investigation of CPI found common items between the scales to have specially contributed to the internal consistency of all the scales of this multidimensional test.

Cross-cultural Validity

CPI has been stipulated to be positively related set of qualities important for adjustment and social living. This claim is first to be verified by Pakistani

data also in order to assess the applicability of this rationale of the test here. This will reflect on the cross-cultural implication of the construct of these traits. Pakistani data were, therefore, intercorrelated on all the scales, on both the versions separately (see tables 5 & 6). The fact that most of the indices so obtained in the correlational matrix were significantly positive upheld the rationale that the test assesses a set of related vectors.

Next, Pakistani data were compared with the American data on English version taken from the California Psychological Inventory Handbook (Megargee, 1977) to see how close or otherwise they are vis-a-vis the two national samples (see table 5). The inter-scale correlation between the two national data as well as between the two versions on the Pakistani data was computed as .099 as against the same estiamtes of .299 between the two versions (see table 6). The smaller difference between the data of the cross-national groups supports the cross-cultural validity of the "Folk Concept" construct logic of the test and endorses the claim that CPI is applicable across borders. Relatively larger index of difference between the two versions, however, reflects on the current degree of equivalence between them.

Cultural Differences and Response Rates

Pakistani data of the Urdu version were compared to that of the modal American profile given in the CPI manual (Gough, 1957) for cross-national normative comparison (see fig. 1) where-upon the former was found to be much lower than the latter in terms of the base rate on most of the scales including two of the three validity scales. This is indicative of cultural differences which affected the frequency of response-rate in the two samples, that is, the Pakistani subjects did not respond to the items in the keyed direction as frequently as the Americans did. This refers to the measurement problem and so makes an ipso facto case for renorming of the test on Pakistani subjects. Gynther (1972), in his review of the literature, while comparing the performance of the Whites and the Blacks, has stated similar views about MMPI. Revalidation of the scales may also be necessary, if subsequent research shows them to lack predictive and classificatory validity.

Aslo, in the context of Pakistan, Todd (1983) and Mirza (1983) found MMPI scores obtained by the Pakistani subjects to be significantly higher than those of the Americans, which leads to their elevated profile indicating strong abnormal tendencies. The profile of even the normal Pakistani group. on MMPI appeared quite higher than that of the Americans as a matter of cultural differences in thinking and feeling. Mirza (1983) attributed it to the tendency of Pakistani subjects, as in the case of Japanese, to feel more strongly and to think and act in a way that would appear to the Americans as rather severe and abnormal behaviour. The performance of the general
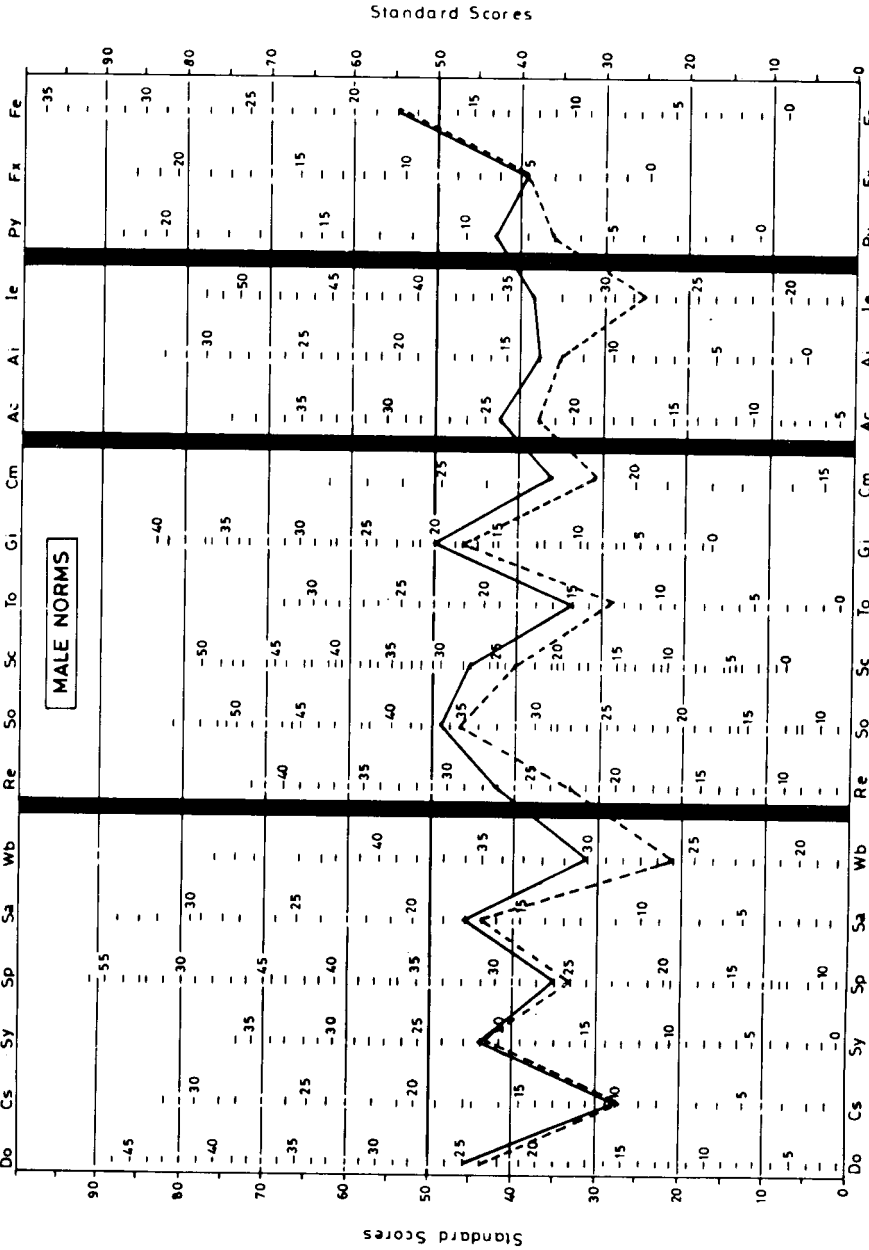
Figure 1: Mean profile on Urdu and English versions of CPI, on American Norms, Solid line; 37 students who took Urdu CPI. Broken line; 33 students who took English CPI.

Table 5

Matrix of CPI (English) Scale Intercorrelation for American & Pakistani data

	Do	Cs	Sy	Sp	Sa	Wb	Re	So	Sc	To	Gi	Cm	Ac	Ai	Ie	Py	Fx	Fe
Do		454	640	405	540	270	355	110	030	305	260	125	430	145	425	285	-110	-090
Cs	488		615	535	485	380	345	145	150	505	340	070	425	410	565	370	190	-655
Sy	637	629		590	585	385	285	155	-070	360	280	180	435	145	530	240	-080	-165
Sp	557	541	491		525	315	030	040	-150	310	-045	055	140	250	465	230	285	-280
Sa	730	394	583	529		110	065	-040	-270	125	-070	125	130	100	315	090	-065	-095
Wb	235	597	437	386	074		460	440	615	650	565	260	620	370	620	410	025	-070
Re	406	319	170	189	131	434		450	480	545	440	280	575	390	505	290	-060	195
So	164	380	028	028	-069	567	525		485	335	390	300	500	140	310	125	-185	140
Sc	-178	271	-028	-220	-400	587	527	597		525	750	110	610	345	370	390	-070	150
To	327	644	451	338	126	649	452	526	507		500	155	575	620	665	465	200	-035
Gi	177	496	158	054	-064	669	602	564	815	579		-035	625	280	360	385	-085	020
Cm	378	350	271	175	116	229	465	280	150	360	135		255	-055	-225	-055	-225	-110
Ac	207	455	408	028	051	719	528	549	680	587	641	375		385	585	380	-130	-100
Ai	117	355	174	102	071	613	547	535	521	527	549	275	620		560	430	440	-075
Ie	355	609	283	376	119	657	488	489	528	554	669	220	478	501		430	135	-075
Py	531	495	438	443	209	556	579	442	361	480	479	329	524	256	545		200	-080
Fx	102	042	066	445	225	050	-142	-217	-421	-015	-223	015	-228	102	-093	-122		-070
Fc	246	-112	-351	-248	-209	-045	131	174	281	-010	073	085	091	-038	-044	100	-180	

Scale intercorrelation for the English version on American data are in the upper diagonal and on Pakistani data are in the lower diagonal. American data is taken from CPI Handbook (Megargee, 1977,pp.106—107) as mid values between male and female indices, to have a combined matrix for the sexes..

Table 6

CPI Scale Intercorrelation Matrix for English and Urdu Version

	Do	Cs	Sy	Sp	Sa	Wb	Re	So	Sc	To	Gi	Cm	Ac	Ai	Ie	Py	Fx	Fc	
Do		488	637	557	730	235	406	164	-178	327	177	378	207	117	355	531	102	246	Do
Cs	508		629	541	394	597	319	380	271	644	496	350	455	355	609	495	042	-112	Cs
Sy	563	629		491	583	437	170	028	-028	451	158	271	408	174	283	438	066	-351	Sy
Sp	498	541	491		529	386	189	231	-220	338	054	175	028	102	374	443	445	-248	Sp
Sa	513	595	136	249		074	131	-069	-400	126	-064	116	051	071	119	209	225	-209	Sa
Wb	053	065	161	179	136		434	567	587	649	669	229	719	613	657	556	050	-045	Wb
Re	302	459	001	-206	457	130		525	527	452	602	465	528	547	488	579	-142	131	Re
So	-211	-136	-169	-063	-092	423	150		597	526	564	280	549	535	489	442	-217	174	So
Sc	-074	347	-303	-177	055	540	429	286		507	815	150	680	521	528	361	-421	281	Sc
To	339	486	089	257	247	499	454	262	574		579	360	587	549	544	480	-015	-010	To
Gi	237	252	209	213	222	543	189	146	590	342		135	641	549	669	479	-223	073	Gi
Cm	127	267	027	191	394	286	358	141	131	244	103		375	275	220	329	015	-179	Cm
Ac	449	223	389	244	092	550	288	352	451	580	664	040		620	478	524	-228	091	Ac
Ai	294	652	-046	240	468	208	493	012	377	729	166	432	165		501	256	102	-038	Ai
Ie	073	-323	334	142	-269	417	001	233	109	254	090	247	273	-075		545	-093	-044	Ie
Py	382	803	-169	210	639	-053	381	-269	324	387	136	155	055	557	-503		-122	100	Py
Fx	180	451	-027	318	380	156	157	045	169	218	169	373	034	285	001	244		-180	Fx
Fe	-305	-452	075	-072	-364	-080	-235	221	-214	-250	092	-179	049	-293	098	-608	134		Fe

Scale intercorrelations for the English CPI are given in the upper diagonal, for Urdu CPI in the lower diagonal on Pakistani data ($N=33$ & 37, respectively).

group appeared otherwise (fig. 1) on CPI, for the same reasons, as the CPI is a test for normal characteristics.

The similarity of Urdu and English Pakistani profiles between themselves and their identical deviance from the American normative data suggests that differences between the two cultural groups were stronger than linguistic differences between the two versions. This showed that difference between the responses of Pakistani and American subjects exist by different nature of social living which affect their personal outlook, interest and values resulting consequently in differential appreciation of the test contents. Cross (1978) also undertook a similar study on CPI and found varied constellation of traits and configuration in cross-national profiles as a result of cultural differences in the comparative groups. The Urdu CPI, therefore, needs to be renormed in Pakistan.

CONCLUSION

Though this was a preliminary study having certain limitations pertaining to the small size of sample and a research design of rather limited scope, it did evidence that the Urdu CPI holds reasonably satisfactroy indices of psychometric qualities by Pakistani data also. For the valid use of the test and specially as a translated test to be used in the target population, it has to be, however, revalidated anew in its own right, using essentially the same set of items and procedures but redefining the constructs of CPI scales in the context of Pakistani culture and interpreting the data by local criterion and norms. In fact, the Pakistani profile does justify an examination and possible redesigning of at least some of the scales (To, Wb, Cs, etc). Simple work, renorming of the test, as is sometimes considered to suffice in cross-national research would be a practice of very limited use and of questionable validity. Whereas individual differences and possible sex differences on the test will be in order, for more meaningful use of the test, it must also be seen whether education, residence (Urban-Rural), socio-economic level and other demographic variables affect CPI scores in our conditions and if so, these correlates should be studied through using controls for the factors to find how much they account for the test results and in what direction.

REFERENCES

Ahmad, I. (1985). *Development of Urdu version of California Psychological Inventory in Pakistan.* Islamabad: National Institute of Psychology.

Bond, H. Michael & Yang Kuo-Shu (1980). Ethnic affirmation by Chinese bilinguals. *Journal of Cross-cultural Psychology, 4,* 411 – 425.

Bouchard, T. J. (1969). Personality, problem solving procedure and performance in small groups. *Journal of Applied Psychology Monorgraphs, 53,* No. 1 Part 2, 1 – 29.

Brislin, R. W., Lonner, W. J. & Thorndike, R. M. (1973). *Cross-cultural research methods.* New York: Wiley.

Butcher, J. N. & Pancheri, P. (1976). *A handbook of cross-national MMPI research.* Minneapolis: University of Minnessota Press.

Cross, D. T., Barclay, A. & Burger, G. K. (1978). Differential effects of ethnic membership, sex and occupation on California Psychological Inventory.. *Journal of Personality Assessment, 42,* 597 – 603.

Farley, F. H., & Cohen, A. (1980). Common items and reliability in personality measurement. *Journal of Research in Personality, 14,* 207 – 211.

Gough, H. G. (1956). *California Psychological Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Gough, H. G. (1957). *Manual for the California Psychological Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Gough, H. G., Chun, K. & Chung, Y. E. (1968). Validation of the CPI femininity scale in Korea. *Psychological Reports, 22,* 155 – 160.

Gynther, M. D. (1972). White norms and Black MMPIs: A prescription for discrimination? *Psychological Bulletin, 77,* 386 – 403.

Levin, J. and Karni, E. S. (1971). A comparative study of the CPI femininity scale: Validation in Israel. *Journal of Cross-cultural Psychology, 2,* 387 – 391.

Megargee, I. E. (1972). *The California Psychological Inventory handbook.* San Francisco: Jossey-Bass.

Mirza, L. (1975). *MMPI Urdu version.* Lahore: Fountain House.

Mirza, L. (1983). Multiple administration of the MMPI with schizophrenics. *Pakistan Journal of Psychology, 14,* 23−36.

Mitchell, J. V. Jr, & Pierce-Jones, J. (1960). A factor-analysis of the California Psychological Inventory. *Journal of Consulting Psychology, 24* 453−456.

Nichols, R. C., & Schnell, R. R. (1963). Factor scales for the California Psychological Inventory. *Journal of Consulting Psychology, 27,* 228−235.

Nishiyama, T. (1975). Validation of CPI femininity scale in Japan. *Journal of Cross-cultural Psychology, 6,* 482−489.

Pitariu, H. (1981). Validation of CPI femininity scale in Romania. *Journal of Cross-cultural Psychology, 12,* 111−117.

Todd, T. (1983). Cross-cultural concerns in standardized testing: Some suggestions about the use of MMPI in Pakistan. *Pakistan Journal of Psychology, 14,* 27−36.