



2-1-2003

# Measuring Personality and Values Across Cultures: Imported Versus Indigenous Measures

Fanny M. Cheung

*The Chinese University of Hong Kong, fmcheung@cuhk.edu.hk*

Shu Fai Cheung

*The Chinese University of Hong Kong, sfcheung@umac.mo*

---

## Recommended Citation

Cheung, F. M., & Cheung, S. (2003). Measuring Personality and Values Across Cultures: Imported Versus Indigenous Measures. *Online Readings in Psychology and Culture*, 4(4). <https://doi.org/10.9707/2307-0919.1042>

---

# Measuring Personality and Values Across Cultures: Imported Versus Indigenous Measures

## Abstract

Cross-cultural studies of personality have shown cultural similarities and differences in the manifestation of personality traits. In interpreting cultural differences in personality traits, we should consider not only the experiences of people in different cultures, but also the measures adopted and the cultural orientations of researchers themselves. In this chapter, we discuss the considerations in adapting an instrument from one culture to another culture. We illustrate the alternative approach of developing an indigenous personality measure in the Chinese culture.

### Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](https://creativecommons.org/licenses/by-nc-nd/3.0/).

## Introduction

Personality assessment constitutes a major function in applied psychology in many Asian countries (Cheung, Leong, & Ben-Porath, 2003). As in the West, for example, clinical psychologists use assessment instruments to aid diagnostic and treatment decisions. The use of scientific methods of assessment enhances the status of the professionals especially in countries where clinical psychologists are a relatively young profession (Cheung, 1996). In Asian countries, the field of personality assessment is largely an "imported" discipline, following the Western tradition and paradigm in psychology. As a result, most of the common personality tests currently used in Asia are tests translated from English. For example, Asian language versions for the Minnesota Multiphasic Personality Inventory-2 (MMPI) include Chinese, Japanese, Korean, Thai, Vietnamese, and Hindi; for the Eysenck Personality Questionnaire (EPQ) - Chinese, Japanese, Korean, and Thai; The State-Trait Anxiety Inventory (STAI) – Chinese, Japanese, Korean, Hindi, Bengali, Kannada and Tamil; and the NEO-PI-R – Chinese, Japanese, Korean, and Filipino. The strategy of applying foreign instruments and constructs in the local culture, assuming cross-cultural validity and relevance, is called the *imposed etic* strategy (Berry, 1989; Church & Lonner, 1998). However, there are important cross-cultural considerations in using translated personality tests.

### Cross-Cultural Considerations in the Use of Translated Personality Tests

Much has been written about issues, problems, and procedures associated with the translation and use of psychological tests in cultures in which the tests were not developed (e.g., Hambleton, 2001; Tanzer & Sim, 1999; van de Vijver & Hambleton, 1996). Focusing primarily on personality tests, inventories, and scales, we discuss eight major issues that need to be addressed when using such devices in another culture:

1. Insuring the adequacy of the translation and adaptation, and appropriateness for the target culture.
2. Equivalence of the translated and the original instrument(s)
3. Local research on reliability and validity of the instrument(s)
4. Standardization of the translated instrument(s) in the target culture.
5. Implications of using the original or local norms
6. Correctly interpreting cross-cultural differences in test scores
7. Culture-relative ethical standards and copyright issues in test use
8. Gaps in culturally relevant constructs

## **Adequacy of Translation and Adaptation**

The strategy for the translation and adaptation of the MMPI-2 in international use (Butcher, 1996) is presented here as an example of how to ensure the adequacy in translation and adaptation.

### **Requirements expected of the test translator**

The test translator should be experienced in both the source and the target languages. Preferably, the translator should be a "true" bilingual who associates two sets of linguistic signs "with two different sets of meanings" (Butcher, 1996, p.29). Individuals with "bicultural" experience, or those who learned the language in the culture of origin, are higher in the degree of true bilingualism.

In translating complex or obscure expressions, translators usually have to choose between complex and rarely used phrases but that are equivalent to the original terms, or phrases that are more natural and commonly used in the target language but less similar to the original terms. Using the more natural phrases can ensure that a wider range of respondents in the target culture can understand the item.

### **Care in maintaining the original meaning**

To maintain content validity, it is necessary to ensure the translated items and the original items are linguistically equivalent. That is, assurance has to be made that the items have the same meaning in the target and the original languages. To achieve this, it is preferable to avoid using abstract terms and use concrete terms, as the equivalence of meaning is more difficult to establish in abstract terms.

### ***Ways to adapt or modify items that could not be translated meaningfully***

When it is difficult to translate a particular item meaningfully into the target language, a literal translation with explanation in parentheses is preferable. Despite the increase in length and complexity, this can enhance the linguistic equivalence between the original item and the translated one. When an item has no equivalent in the target culture, for example, an idiom that has no equivalent in the target language, then it is acceptable to replace the item with another one that is culturally appropriate. However, field tests (discussed below) need to be conducted specifically for this kind of item to establish the equivalence in psychological meaning.

### **Use of a field test and a bilingual retest method to identify problematic items**

Even after several phases of translation, the researchers and translators judge that the translated items are highly equivalent to the original items in psychological and linguistic meanings, the translated items still need to be tried out in a field test, to identify any problematic items. Interview with a small sample on their responses to the items can provide in-depth insight into possible misinterpretation of the translated items, and any

cultural differences in the interpretation. In addition to the field test, both the original and the translated versions can be administered to a group of bilinguals using a test-retest design. The translation/bilingual equivalence can be examined in terms of the correlation of the translated and original measures (Cheung 1985; 1996).

### **Three Levels of Equivalence of the Translated and the Original Instruments**

One cannot assume that a translated instrument is equivalent to the original instrument. We need to demonstrate that they are equivalent. Butcher (1996) presented three levels of equivalence between the original and the translated instruments or scales. Two scales are said to be functionally (or structurally) equivalent if they measure the same constructs in the original and the target cultures, even though the item contents of the two scales may be different. We can evaluate functional equivalence by examining the original and the translated instruments' inter-item or inter-scale correlations, and their patterns of correlation with external variables.

The second level is metric equivalence. In general, metric equivalence refers to the similarities between the original and the translated instruments in psychometric properties, such as item difficulty level, item-scale correlations, and the pattern of loadings in factor analyses.

The third level is scalar (or full score) equivalence--the extent to which the scale scores indicate the same degree, intensity or magnitude of the characteristic being measured in both cultures. Although this level of equivalence is implicitly assumed when we compare the scale mean differences between two cultures, this level of equivalence is the most difficult to establish (Berry, Poortinga, Segall, & Dasen, 2002).

### **Local Research on Reliability and Validity**

Before applying the translated instrument as if it is the original instrument, we need to conduct local research on the translated instrument. An adapted instrument is like a newly developed instrument. We need to establish its reliability in the local culture, including its internal consistency, test-retest reliability, and factor structure. Problems in internal consistency may reflect inadequacies in translation, genuine cross-cultural differences in the manifestation of a characteristic, or the lack of cultural relevance of the imported construct. Other psychometric properties of the translated measure also should be compared to those of the original measure, such as item difficulty and endorsement rate. Likewise, the validity of the translated instrument has to be established through a program of local research. As with original instruments, four aspects of validity need to be studied. How do the scale scores converge with other related instruments (convergent validity)? Do the scale items cover the construct being measured adequately (content validity)? Do the scales predict specified outcomes or discriminate between the normative sample and criterion groups (criterion validity)? What is the nomological net covered in the scale and does the construct measured by the scale cover the same nomological net or meaning cross-culturally (construct validity)?

### **Standardization of the Translated Instrument**

The scores of an instrument are meaningful only when they are referenced against a relevant comparison group. In using the original norms of an imported instrument, we are assuming the normative sample is relevant and comparable to the target respondents. However, we often find cross-cultural differences in test scores. Given these differences, we may misinterpret the meaning of the scores by assuming the equivalence of the norms. This is particularly problematic when diagnostic or selection decisions using cut-off points derived from these norms. Standardizing the translated instrument in the local culture will provide norms that are more accurate for the local population. To ensure the comparability with the original instrument, the translated instrument should be standardized in a normative sample that is both representative of the local population and comparable to the normative sample of the original culture (Butcher, 1996).

### **Use of Original vs. Local Norms**

When should we use the original norms and when should we develop local norms? There are pros and cons for using the original norms with translated measures. Original norms constitute an integral part of the assessment instrument. Interpretation of scale scores is based on research using these norms. The original norms also provide the common yardstick for cross-cultural comparison of scores. This allows for the accumulation of cross-cultural research findings on the instrument.

However, adopting the original norms assumes that the original standardization sample is relevant and equivalent to the local respondents, which may not be the case. For example, Chinese respondents tend to obtain higher T scores on a number of the clinical scales on the MMPI-2 when the American norms are used (Cheung, 1996). The higher scores may not indicate higher prevalence of psychopathology among the Chinese people; instead, the mean differences in scores may reflect cultural differences in desirability of certain behaviors included in the item contents (Cheung, 1995). The Chinese values of modesty and restraint, for instance, may affect the elevation of scores on depression-related scales.

Using local norms, on the other hand, can ensure that the interpretation of scale scores is relevant to the local culture. Conducting standardization studies is a large-scale project that requires resources. Practitioners in the local culture have less access to research resources whereas academic researchers may not be willing to devote their resources to applied research with little theoretical value and publication opportunities. It is only when there is a large demand for as well as substantial support for the use of the instrument before local standardization studies are feasible. Otherwise, test users continue to borrow from imported measures and rely on the original norms in order to have usable assessment tools within their constraints.

### **Cross-Cultural Differences in Test Scores**

If the original norms are used, we have to take note of the cross-cultural differences in the means and distributions of scale scores in the test interpretation. For example, the mean scores among normal Chinese samples on several MMPI-2 clinical scales, including 2 (D), 7 (Pt) and 8 (Sc), are higher than those obtained by their American counterparts. Do the elevated scores on these scales mean greater psychopathology among the Chinese normal samples? Or are there other cultural factors that affect their responses to some of the items on these scales? If that is the case, there may be a risk of overestimating psychopathology if the test scores are interpreted directly according to the original norms. The more important question is whether the scales discriminate between normal and clinical respondents in the local culture as they do in the original culture.

### **Ethical Standards and Copyright Issues in Test Use**

Translating and adapting Western measures is a common practice in psychological assessment in Asian countries. Despite the emphasis on the scientific standards of psychological assessment, there has been relatively little guidance on the ethical standards of test use and test adaptation for international psychologists. Many tests are translated or adapted without the original authors' or publishers' permission, and copyright compliance is not always observed. The Association of Test Publishers (ATP; Hambleton and Patsula, 1999) and the International Test Commission (ITC; Bartrum, 2000) have recently published guidelines on test use and test adaptation, which suggest specific steps for test adaptation, development, administration, and documentation when translating tests from another language. There are also stipulations for the qualification of test users and access to test materials.

### **Gaps in Culturally-Relevant Constructs**

Despite the widely adopted and improved practice of translating personality tests, two theoretical gaps need to be addressed. First, translated tests are almost always based on imported Western measures. The personality constructs are imposed-etic concepts applied to the local culture under the assumption that they are cross-culturally relevant. However, whether the imposed etic personality constructs are universally applicable is an empirical question that needs to be investigated rather than assumed (Cheung & Leung, 1998). Second, indigenous and culturally relevant constructs, i.e. the emic concepts, may be missing from these imported measures. Are there personality characteristics that are important to the understanding of mental health and personality in the local culture that have been left out in Western measures? These two gaps highlight the limitations of using only translated instruments. Some international researchers have endorsed the indigenous approach, which takes into account the "sociocultural realities" of the local culture (Berry et al., 2002, p.460). The emic constructs complement the etic constructs in providing a comprehensive understanding of personality in non-Western cultures. We will illustrate the

combined emic-etic approach in developing an indigenous personality measure for the Chinese people.

### **Development of an Indigenous Personality Measure: The Chinese (Cross-Cultural) Personality Assessment Inventory**

The objective of developing an indigenous personality measure is to construct a comprehensive personality inventory suited to the local needs, while retaining the psychometric standards of established assessment measures. The Chinese Personality Assessment Inventory (CPAI) was designed using a combined emic-etic approach (Cheung, Leung, Fan, Song, Zhang, & Zhang, 1996). It was developed as a collaborative project involving psychologists in Hong Kong and Mainland China. The CPAI includes both universal personality constructs also found in Western personality theories (e.g., Leadership, Optimism vs. Pessimism, Emotionality) and indigenous personality constructs derived from the local context (e.g., Family Orientation, Harmony, Face, Thrift vs. Extravagance, *Renqing*, i.e., Relationship-Orientation, and Somatization). In deriving the personality constructs used in the Chinese culture, the researchers reviewed the psychological and popular literature, and conducted empirical surveys of person descriptions. Based on preliminary studies with large samples, the researcher selected scale items that met specified psychometric criteria. The first version of the CPAI was standardized in 1993 using representative samples of adults from different regions of Mainland China, including Hong Kong (N = 2,444). Four personality factors and two clinical factors were extracted from the CPAI. The four personality factors are *Dependability*, *Social Potency*, *Individualism*, and *Interpersonal Relatedness*. When the CPAI was jointly factor analyzed with the NEO-PI-R, it was found that the *Interpersonal Relatedness* factor was distinct from the Big Five of the Five Factor Model (Cheung et al., 2001). The *Interpersonal Relatedness* factor measures various aspects of interdependence in interpersonal relationship. On the other hand, the domain of Openness was absent from the CPAI.

A new set of openness scales was developed for the CPAI using a similar approach as in the development of the original scales. The second version of the CPAI, CPAI-2, was re-standardized in 2001 using the same sampling methods to obtain representative norms in Mainland China and Hong Kong. The normative sample consists of 1911 adults with an age range of 18 to 70 years.

The CPAI-2 consists of 28 personality scales, 12 clinical scales (including one that is double listed as a personality scale), and 3 validity indexes, with a total of 541 items. There are about 10 items on each personality scale and 20 items on each clinical scale. The items are self-descriptions of behavior to be answered in a true-false format. Even with the addition of six new openness-related scales, four personality factors and two clinical factors were extracted from the CPAI-2 using Principal Components Analysis. The four personality factors resembled those from the original CPAI: *Social Potency*, *Dependability*, *Accommodation*, and *Interpersonal Relatedness*. The two clinical factors,

*Emotional Problems* and *Behavioral Problems*, were also similar to the original factor structure. Table 1 lists the CPAI-2 scales loaded on the personality and clinical factors.

**Table 1**  
The Scales and Factors of the CPAI-2

<b>Personality Factors</b>	
<b>Social Potency</b>	<b>Dependability</b>
Novelty ( <i>NOV</i> )	Responsibility ( <i>RES</i> )
Diversity ( <i>DIV</i> )	Emotionality ( <i>EMO</i> )
Divergent Thinking ( <i>DIT</i> )	Inferiority vs Self-Acceptance ( <i>I-S</i> )
Leadership ( <i>LEA</i> )	Practical Mindedness ( <i>PRA</i> )
Logical vs Affective Orientation ( <i>L-A</i> )	Optimism vs Pessimism ( <i>O-P</i> )
Aesthetics ( <i>AES</i> )	Meticulousness ( <i>MET</i> )
Extraversion vs Introversion ( <i>E-I</i> )	Face ( <i>FAC</i> )
Enterprise ( <i>ENT</i> )	Internal vs External Locus of Control ( <i>I-E</i> )
	Family Orientation ( <i>FAM</i> )
<b>Accommodation</b>	<b>Interpersonal Relatedness</b>
Defensiveness (Ah-Q Mentality) ( <i>DEF</i> )	Traditionalism vs. Modernity ( <i>T-M</i> )
Graciousness vs Meanness ( <i>G-M</i> )	Ren Qing (Relationship Orientation) ( <i>REN</i> )
Interpersonal Tolerance ( <i>INT</i> )	Social Sensitivity ( <i>SOC</i> )
Self vs. Social Orientation ( <i>S-S</i> )	Discipline ( <i>DIS</i> )
Veraciousness vs Slickness ( <i>V-S</i> )	Harmony ( <i>HAR</i> )
	Thrift vs Extravagance ( <i>T-E</i> )
<b>Clinical Factors</b>	
<b>Emotional Problem</b>	<b>Behavioral Problem</b>
Inferiority vs Self-Acceptance ( <i>I-S</i> )	Pathological Dependence ( <i>PAT</i> )
Anxiety ( <i>ANX</i> )	Hypomania ( <i>HYP</i> )
Depression ( <i>DEP</i> )	Antisocial Behavior ( <i>ANT</i> )
Physical Symptoms ( <i>PHY</i> )	Need for Attention ( <i>NEE</i> )
Somatization ( <i>SOM</i> )	Distortion of Reality ( <i>DIR</i> )
Sexual Maladjustment ( <i>SEM</i> )	Paranoia ( <i>PAR</i> )

Note: Abbreviations are presented in italics in the brackets.

The distinctiveness of the *Interpersonal Relatedness* factor from the Big Five is of particular cross-cultural interest. Joint factor analysis between the CPAI-2 and the NEO-FFI again extracted an independent Interpersonal Relatedness factor, even after the addition of six openness-related scales. The unique Interpersonal Relatedness factor has demonstrated incremental validity beyond the etic Western personality factors. The IR

factor scales added predictive value beyond those contributed by the Big Five dimensions in predicting a variety of Chinese social behavior, including filial piety, trust, persuasion tactics, and group communication styles (Cheung, Leung, et al., 2001). The IR factor also significantly explained additional variance in the Somatization scale of the CPAI in the standardization sample, beyond that explained by the Dependability factor.

The usefulness of Interpersonal Relatedness factor suggests that if only imported instruments had been used in with the Chinese people, we would not have identified this salient interpersonal construct in the Chinese culture. Recent studies using an English version of the CPAI with Caucasian and Asian American students (Cheung, et al., 2001; Cheung, Cheung, et al., 2003) further showed that the same four personality factors of the CPAI, including the Interpersonal Relatedness factor, could be extracted from non-Chinese samples. In addition to the English version, the CPAI also has been translated into Korean and Japanese. Studies in these other Asian countries will examine the relevance of the personality dimensions derived in the Chinese cultural context in other Confucian-related cultures. Preliminary results showed that a high factor congruence was obtained on the IR factor from samples of Korean, Japanese, Caucasian American, and Asian American students, using Procrustes Factor Analysis of the CPAI-2 with the Chinese normative sample as the target. The indigenously identified personality dimension of Interpersonal Relatedness not only enhances our understanding of personality in collectivistic cultures, but also suggests that Western theories and measures of personality could be adapted to reflect the neglected interdependence nature of Western cultures.

Our original attempt to develop a comprehensive indigenous personality measure for the Chinese people has led us to reflect on the broader issue of what are etic and what are emic personality traits. By translating the CPAI-2 into other languages, it has become an "imported" measure in the other cultures. We have re-named the CPAI-2 as the "Cross-Cultural" instead of "Chinese" Personality Assessment Inventory. The same cross-cultural considerations that we discussed above apply.

## References

- Bartrum, D. (2000). *International Guidelines for Test Use*. Belgium: International Test Commission. Retrieved from <http://www.intestcom.org>.
- Berry, J. W. (1989). Imposed etics-emics-derived etics: The operationalization of a compelling idea. *International Journal of Psychology, 24*, 721-735.
- Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications* (2nd Ed.). Cambridge: University Press.
- Butcher, J. N. (1996). Translation and adaptation of the MMPI-2 for international use. In J. N. Butcher (Ed), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 26-43). Minneapolis, MN: University of Minnesota Press.
- Cheung, F. M. (1985). Cross-cultural considerations for the translation of the Chinese MMPI in Hong Kong. In J. N. Butcher & C. D. Spelberger (Eds.) *Advances in personality assessment: Vol. 4*. (pp. 131-158). Hillsdale, NJ: Lawrence Erlbaum.

- Cheung, F. M. (1995). *Administration manual of the Minnesota Multiphasic Personality Inventory (MMPI) Chinese edition*. Hong Kong: The Chinese University Press.
- Cheung, F. M. (1996). The assessment of psychopathology in Chinese societies. In M. H. Bond (Ed.). *Handbook of Chinese psychology* (pp. 393-411). Hong Kong: Oxford University Press.
- Cheung, F. M., Cheung, S. F., Leung, K., Ward, C., & Leong, F. (2003). The English version of the Chinese Personality Assessment Inventory. *Journal of Cross-Cultural Psychology, 34*, 433-452. doi:10.1177/0022022103034004004
- Cheung, F. M., Leong, F. T. L., Ben-Porath, Y. (2003). Psychological assessment in Asia: Introduction to the special section. *Psychological Assessment, 15*, 243-247. doi:10.1037/1040-3590.15.3.243
- Cheung, F. M., & Leung, K. (1998). Indigenous personality measures: Chinese examples. *Journal of Cross-Cultural Psychology, 29*, 233-248.
- Cheung, F. M., Leung, K., Fan, R., Song, W. Z., Zhang, J. X. & Zhang, J. P. (1996). Development of the Chinese Personality Assessment Inventory (CPAI). *Journal of Cross-Cultural Psychology, 27*, 181-199.
- Cheung, F. M., Leung, K., Zhang, J. X., Sun, H. F., Gan, Y. G., Song, W. Z., Xie, Dong. (2001). Indigenous Chinese personality constructs: Is the Five-factor Model complete? *Journal of Cross-Cultural Psychology, 32*, 407-433.
- Church, A. T., & Lonner, W. J. (1998). The cross-cultural perspective in the study of personality. *Journal of Cross-Cultural Psychology, 29*, 32-62.
- Hambleton, R. K. (2001). The next generation of the ITC test translation and application guidelines. *European Journal of Psychological Assessment, 17*, 164-172.
- Hambleton, R. K., & Pastsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology, 1*, 1-12.
- Tanzer, N. K & Sim, C. Q. E. (1999). Adapting instruments for use in Multiple languages and cultures: A review of the ITC guidelines for test adaptations. *European Journal of Psychological Assessment, 15*, 258-269.
- Van de Vijver, F. & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1*, 89-99.

### **About the Authors**

Fanny M. Cheung received her Ph.D. in Psychology from the University of Minnesota. She is currently Professor of Psychology and Chairperson of the Psychology Department at The Chinese University of Hong Kong. She is a Past President of the Hong Kong Psychological Society, and a Past President of the Division of Clinical and Community Psychology of the International Association of Applied Psychology. She is the translator of the Chinese versions of the MMPI, MMPI-2 and MMPI-A, and is the Principal Author of the Chinese (Cross-Cultural) Personality Assessment Inventory. Her research interests include cross-cultural personality assessment, Chinese psychopathology, and psychology of gender.

Shu Fai Cheung received his Ph.D. in Psychology from the Chinese University of Hong Kong. He is currently working as Assistant Professor at the University of Macau. His research interests include personality assessment, meta-analysis, structural equation modeling, and the relationship between attitude and behavior.

### **Questions for Discussion**

1. What are the considerations in adapting a personality assessment tool developed in another culture?
2. In what situations is it justified to translate a foreign personality assessment tool and administer it in the local community? In what situations is it not justified?
3. If you have used a translated personality measure, evaluate to what extent that translated measure is equivalent to the original measure.
4. When you administer different language versions of a personality assessment measure in two different cultures, how can you judge whether the scores are comparable? If you find a large mean difference between the two samples, how would you determine whether it is a genuine difference on the dimension measured and not a methodological artifact?
5. What are the advantages and disadvantages of the indigenous approach in personality assessment?
6. Whether it is true or just an oversimplification, it is commonly believed that the Western culture is individualistic in nature. However, an Interpersonal Relatedness factor is identified in the CPAI, which includes personality dimensions not covered by most Western instruments. Do you think this factor is unique to Chinese and/or other collectivistic cultures? Are there Western personality theories that suggest that this interpersonal factor is also applicable to the Western culture?
7. As an interesting exercise in translating items for use in another culture, select 10 items or questions that were part of a recent exam in another class (psychology, history, sociology, or some other academic area). Discuss the various things you would have to take into account if these items were to be used in another culture.