

Reliability and Validity of the Myers-Briggs Type Indicator® Form M when Translated into Traditional and Simplified Chinese Characters

Carl J. Beuke

David G. Freeman

Shanping Wang

Australian Psychologists Press

Abstract

The MBTI® instrument assesses four aspects of normal personality, believed by the indicator's authors to be universal across cultures. Previous versions of the MBTI instrument have been translated into several languages. This study examined the reliability and validity of the most recently revised version of the MBTI instrument, Self-Scorable Form M, when translated into Traditional and Simplified Chinese characters. For each translation, about 100 participants for whom Chinese was their first language, and who were naïve to type theory, completed the instrument. They then estimated their type based on a standardised explanation. Psychometric evaluation of the translated inventories demonstrated that most items functioned acceptably, and that the overall reliability of the scales was also acceptable. The level of correspondence between type as reported by the inventories, and as independently estimated by the participants, was supportive of the validity of the instruments. Overall, the results are consistent with the cultural universality of the aspects of personality measured by the MBTI instrument.

The MBTI instrument is the most widely used personality inventory in the world (Myers, McCaulley, Quenk, & Hammer, 1998). The inventory assesses four aspects of normal personality, believed by the instrument authors to be universal across cultures (Myers et al., 1998). However, some authors have argued that personality structure is culturally variable, and so personality inventories cannot be expected to work in cultures other than those they were constructed in (see Church, 2001, for a review of different perspectives). Translations of psychological instruments can also be affected by different cultural standards of socially desirable attitudes and behaviour, and by subtle differences in the meaning of translated words (Hofstede & Hofstede, 2005). Despite these issues, previous versions of the MBTI instrument have been successfully translated into several languages (Bents & Wierschke, 1996; Sim & Kim, 1993; Nordvik, 1994; Stalikas & Fitopoulos, 1998).

The aspects of personality measured by the MBTI instrument are:

- Extraversion – Introversion (EI)
- Sensing – Intuition (SN)
- Thinking – Feeling (TF)
- Judging – Perceiving (JP)

The instrument authors argue that the MBTI instrument measures dichotomous preferences, rather than continuous traits. The various combinations of the four scales result in 16 possible

personality *types*, each represented by a four letter code indicating a preference on each scale (for example, ISFJ). The theory and use of the inventory are thoroughly described in Myers et al. (1998). Independent evaluations of Form M of the MBTI instrument have concluded that the instrument's reliability and validity, as conventionally measured, are good (Fleenor & Mastrangelo, 2001). However, the authors' assumptions that preferences are dichotomous, and that the interaction between preferences is an important determinant of personality functioning, remain controversial (Arnau et al., 2003; Fleenor & Mastrangelo, 2001).

Previous versions of the MBTI instrument have been translated into Chinese, and the translated instruments have demonstrated acceptable levels of reliability and validity (Osterlind, Miao, Sheng, & Chia, 2004). However, Form M of the instrument, which is psychometrically improved over earlier versions (Fleenor & Mastrangelo, 2001), has not previously been translated into Chinese. The current study evaluates the performance of Self-Scorable Form M of the MBTI instrument, as translated into both the Traditional and Simplified Chinese Character sets. The Traditional Character set is the main form of writing used in Hong Kong and Taiwan, while the Simplified Character set is the main form of writing used in mainland China (Simplified Chinese Character, 2005). Performance is evaluated using item and overall reliability

® Myers Briggs Type Indicator®, MBTI®, and Introduction to Type are registered trademarks of the Myers-Briggs Type Indicator Trust in the US and other countries. The authors thank Ms Peggy Miu of the University of Hong Kong, and her team of colleagues, for providing the original translations. Carlson Leung and Shanping Wang of Asianic Psychologists Press provided invaluable feedback on the original translation. We also acknowledge the kind assistance of the R & D unit of CPP Inc., who provided 15 experimental items that were assessed during the 1996 revision of the MBTI®.

bility analysis. Also, validity is evaluated by examining the correspondence between type as estimated by the instruments, and type as independently estimated by the participants.

Osterlind et al. (2004) make special note of a particular issue with the translation of the MBTI instrument into Chinese. Items from the inventory can be divided into two categories – phrase items and word pairs. Phrase items ask participants to decide which of two sentences best describe them. Word pair items ask participants to decide which of two words appeal most to them. Osterlind et al. (2004) found that word pair items were less likely to succeed after translation than phrase items, in the sense that they were less likely to load onto the expected factor. They speculated that literal translation is likely to introduce ambiguity to word pair items, because there is less context available to suggest the correct interpretation. The current study addressed this issue by allowing the translators to make non-literal translations of word-pair items that preserved their psychological meaning.

Method

Participants. The participants who completed the Traditional Character set MBTI instrument were 100 native speakers of Chinese, based in Hong Kong. Of the participants, 2 did not specify gender, 62 were female, and 36 were male. Participants were managers from a wide range of operational areas, of which 40 took part in company-sponsored management development programmes, while 60 participated voluntarily in a free MBTI program.

The participants who completed the Simplified Character set MBTI instrument were 102 native speakers of Chinese, based in Shanghai. Of the participants, 1 did not specify gender, 65 were female, and 36 were male. Participants included graduates seeking career advice and junior and middle managers from a range of departments and companies, who took the MBTI instrument as part of management development programmes.

No participant had completed the MBTI instrument on any previous occasion.

Materials. The MBTI instrument was translated from English to Chinese by highly experienced MBTI administrators from the University of Hong Kong, who were fluent in both English and Chinese. In addition to the standard Form M Self-Scorable, an additional question sheet was attached which contained demographic questions and 15 experimental questions. The experimental questions were items from experimental versions of the MBTI instrument that had not been selected for the final version of Form M, but that were believed to have the po-

tential to work well in another culture. The translators were very familiar with the theory underlying the MBTI instrument. Both translations were then reviewed and revised by independent experts, who were fluent in both English and Chinese and had conducted MBTI accreditation programs. The slide series “Presenting Type in Organisations” (Australian Psychologists Press, 2004) and the booklet “Introduction to Type®” (Myers, Kirby, & Myers, 1998) were translated in a similar way. “Presenting Type in Organisations” provides a basic illustration of the four dichotomies. “Introduction to Type” provides an explanation of type theory and a thorough description of each of the 16 types.

Procedure

Participants were administered the MBTI instrument in small to medium groups. The administrator first checked whether any participants had previously completed the MBTI instrument. The remainder of the administration procedure was closely based on that recommended in the instrument manual (Myers et al., 1998). Participants were first given a brief introduction to the purpose of taking the inventory (to promote understanding of self and others), and were given mind-setting instructions (see Myers et al., 1998) but were not given any explanation of type theory. The participants then answered the inventory. Next, the administrator presented an explanation of the four type dichotomies, based around “Presenting Type in Organisations”. The Shanghai participants were given the explanation, and shown the slides, in Chinese, and the Hong Kong participants were given the explanation in English. The participants then self-estimated their type based on the explanation. They next self-scored their MBTI inventories. They then compared their self-estimated type to the type reported by the inventory. They then used the full one-page descriptions of each type, presented in the translated “Introduction to Type”, to select the type that they felt best described them (their “best-fit type”).

Results

Distribution of Type. Table 1 shows the distribution of best-fit types among the participants who completed the translated inventory. One of the participants who took the Simplified translation did not produce a complete best-fit type, and so has been excluded from all relevant analyses. It can be seen that for both translated inventories, the participants obtained a wide variety of best-fit types, with each type represented at least

Table 1. Distribution of best-fit types

| | | | | | |
|------|------|------|------|-------|---------|
| ISTJ | ISFJ | INFJ | INTJ | E | 47 (64) |
| 19 | 12 | 4 | 3 | I | 53 (37) |
| (12) | (9) | (3) | (2) | | |
| ISTP | ISFP | INFP | INTP | S | 67 (59) |
| 3 | 4 | 2 | 6 | N | 33 (42) |
| (1) | (2) | (3) | (5) | | |
| ESTP | ESFP | ENFP | ENTP | T | 59 (61) |
| 3 | 6 | 2 | 4 | F | 41 (40) |
| (3) | (2) | (10) | (2) | | |
| ESTJ | ESFJ | ENFJ | ENTJ | J | 70 (73) |
| 14 | 6 | 5 | 7 | P | 30 (28) |
| (25) | (5) | (6) | (11) | Total | 100 |
| | | | | | (101) |

Note: Numbers refer to raw number of participants. Figures not in brackets refer to the Traditional Chinese Character Translation. Figures in brackets refer to the Simplified Chinese Character Translation.

once. Because the sample is not representative of the general population, no further analysis or comparison with the general US population will be presented in this paper.

Sex differences in best-fit type

Because of the sample sizes used, it would not be meaningful to produce separate type tables for males and females. However, Table 2 compares the proportion of males and females with each preference. Because the samples are not representative of the general population, no further analysis or comparison with the general US population will be presented in this paper. Reliability of the Scales.

Item Properties. The recommended procedure for developing and evaluating a translated version of the MBTI is to follow the general procedure used to develop the English version of the instrument (Myers et al., 1998). Because the MBTI instrument is designed to measure dichotomous preferences, rather than continuous traits, the measure that has traditionally been used as the most important indicator of

item quality is the *prediction ratio* (Myers et al., 1998). The prediction ratio is specific to each answer to an item, and measures the likelihood that a person giving that answer will in fact be of the keyed preference (as estimated by the total score on that preference scale). The formulae used to calculate the prediction ratio for J-P item *I*, first for the J response and then for the P response, are:

$$PR \text{ for item } J_i = \frac{\text{Percent of J participants giving } J_i \text{ response}}{(\% \text{ of J participants giving } J_i \text{ response}) + (\% \text{ of P participants giving } J_i \text{ response})}$$

$$PR \text{ for item } P_i = \frac{\text{Percent of P participants giving } P_i \text{ response}}{(\% \text{ of P participants giving } P_i \text{ response}) + (\% \text{ of J participants giving } P_i \text{ response})}$$

Traditionally, items have only been retained if their prediction ratios are 0.63 or greater for at least one keyed response. To ensure that the overall quality of items in the final translated instruments is similar to that of the English MBTI Form M instrument, two other criteria were applied. The item-total correlation of the item needed to be at least .30, and the correlation of the item with the keyed preference scale needed to be at least .10 greater than the correlation of the item with any of the other three preference scales. Data on the characteristics of the items is shown in Table 3. The items are grouped by scale and identified by the item position in which they appear in the inventories (item order is identical in the English and two translated versions). For the Traditional Character translation, the results show that most items performed well. Of the items translated from Form M, 83 of 93 (89%) functioned acceptably. In all other cases, it is possible to replace the faulty item with an acceptable replacement from the experimental items. For the Simplified Character translation, the results also show that most items performed well. Of the items, 80 of 93 (86%) functioned acceptably. It is not possible

Table 2. Comparison of Best-Fit Preferences of Males and Females

| Preference | Traditional Character | | Simplified Character | |
|-----------------------------|-----------------------|---------------------|----------------------|---------------------|
| | Males (n = 36) | Females (n = 62) | Males (n = 65) | Females (n = 36) |
| Extraversion – Introversion | 36% E | 53% E | 50% E | 71% E |
| Sensing – Intuition | 86% S | 56% S | 67% S | 52% S |
| Thinking – Feeling | 72% T | 50% T | 72% T | 55% T |
| Judging – Perceiving | 81% J | 63 % J | 78% J | 71% J |

Table 3. Item Properties

| Item Code | Traditional Chinese Character Translation | | | | | | Simplified Chinese Character Translation | | | | | |
|-----------|---|----------------|-----------------|----------|----------|----------|--|----------------|-----------------|----------|----------|---------|
| | % INFP | Keyed Answer r | Highest other r | Pr(ESTJ) | Pr(INFP) | Exclude? | % INFP | Keyed Answer r | Highest other r | Pr(ESTJ) | Pr(INFP) | Exclude |
| EI04 | 52 | 82 | 24 | 0.87 | 0.87 | No | 46 | 70 | 6 | 0.81 | 0.93 | No |
| EI08 | 36 | 48 | 8 | 0.73 | 0.62 | No | 25 | 51 | 15 | 0.75 | 0.60 | No |
| EI12 | 36 | 54 | 21 | 0.79 | 0.65 | No | 26 | 56 | 12 | 0.82 | 0.65 | No |
| EI14 | 48 | 59 | 18 | 0.82 | 0.77 | No | 48 | 57 | 12 | 0.74 | 0.83 | No |
| EI18 | 48 | 55 | 9 | 0.75 | 0.72 | No | 42 | 58 | 12 | 0.77 | 0.79 | No |
| EI19 | 33 | 57 | 8 | 0.87 | 0.67 | No | 19 | 60 | 8 | 0.91 | 0.63 | No |
| EI22 | 51 | 59 | 15 | 0.76 | 0.76 | No | 41 | 54 | 19 | 0.75 | 0.74 | No |
| EI23 | 36 | 68 | 18 | 0.97 | 0.74 | No | 24 | 73 | 9 | 0.93 | 0.71 | No |
| EI26 | 48 | 56 | 15 | 0.75 | 0.72 | No | 43 | 63 | 9 | 0.78 | 0.81 | No |
| EI27 | 38 | 42 | 12 | 0.69 | 0.61 | No | 22 | 38 | 10 | 0.71 | 0.57 | No |
| EI34 | 33 | 68 | 25 | 1.00 | 0.73 | No | 38 | 61 | 19 | 0.78 | 0.75 | No |
| EI35 | 44 | 68 | 20 | 0.90 | 0.79 | No | 41 | 67 | 15 | 0.82 | 0.85 | No |
| EI42 | 36 | 61 | 11 | 0.91 | 0.71 | No | 32 | 55 | 16 | 0.80 | 0.70 | No |
| EI48 | 46 | 51 | 26 | 0.76 | 0.71 | No | 30 | 52 | 12 | 0.81 | 0.69 | No |
| EI54 | 32 | 63 | 24 | 0.93 | 0.69 | No | 24 | 65 | 8 | 0.93 | 0.71 | No |
| EI60 | 59 | 60 | 16 | 0.74 | 0.82 | No | 52 | 68 | 6 | 0.74 | 0.88 | No |
| EI62 | 60 | 30 | 4 | 0.59 | 0.63 | Yes | 64 | 35 | 16 | 0.57 | 0.64 | No |
| EI66 | 76 | 54 | 14 | 0.62 | 0.85 | No | 52 | 58 | 16 | 0.71 | 0.81 | No |
| EI67 | 70 | 51 | 12 | 0.63 | 0.79 | No | 71 | 59 | 14 | 0.64 | 1.00 | No |
| EI72 | 82 | 60 | 8 | 0.62 | 1.00 | No | 75 | 47 | 15 | 0.59 | 0.85 | No |
| EI77 | 53 | 72 | 20 | 0.83 | 0.85 | No | 44 | 51 | 13 | 0.70 | 0.70 | No |
| SN03 | 30 | 31 | 10 | 0.67 | 0.59 | No | 25 | 17 | 20 | 0.58 | 0.53 | Yes |
| SN05 | 37 | 48 | 15 | 0.75 | 0.70 | No | 50 | 50 | 24 | 0.67 | 0.70 | No |
| SN13 | 50 | 53 | 29 | 0.67 | 0.72 | No | 27 | 52 | 34 | 0.83 | 0.66 | No |
| SN15 | 40 | 52 | 24 | 0.71 | 0.69 | No | 46 | 57 | 26 | 0.74 | 0.76 | No |
| SN24 | 52 | 46 | 16 | 0.65 | 0.71 | No | 58 | 46 | 20 | 0.63 | 0.70 | No |
| SN29 | 16 | 43 | 37 | 0.82 | 0.58 | Yes | 25 | 56 | 39 | 0.84 | 0.64 | No |
| SN32 | 30 | 63 | 25 | 0.85 | 0.73 | No | 25 | 52 | 35 | 0.84 | 0.65 | No |
| SN37 | 31 | 20 | 12 | 0.66 | 0.58 | Yes | 29 | 16 | 23 | 0.62 | 0.55 | Yes |
| SN40 | 59 | 42 | 31 | 0.64 | 0.75 | No | 52 | 28 | 14 | 0.55 | 0.56 | Yes |
| SN44 | 36 | 59 | 24 | 0.80 | 0.73 | No | 39 | 52 | 37 | 0.73 | 0.68 | No |
| SN47 | 15 | 42 | 29 | 0.85 | 0.59 | No | 9 | 24 | 17 | 0.76 | 0.53 | Yes |
| SN50 | 39 | 53 | 28 | 0.74 | 0.72 | No | 48 | 55 | 41 | 0.74 | 0.78 | No |
| SN53 | 40 | 65 | 28 | 0.81 | 0.82 | No | 53 | 45 | 17 | 0.65 | 0.70 | No |
| SN55 | 22 | 48 | 27 | 0.81 | 0.62 | No | 13 | 28 | 24 | 0.65 | 0.52 | Yes |
| SN58 | 21 | 45 | 32 | 0.83 | 0.61 | No | 32 | 46 | 29 | 0.81 | 0.69 | No |
| SN61 | 44 | 32 | 23 | 0.67 | 0.67 | Yes | 57 | 30 | 13 | 0.56 | 0.58 | Yes |
| SN63 | 46 | 65 | 27 | 0.78 | 0.86 | No | 25 | 33 | 21 | 0.65 | 0.56 | No |
| SN73 | 52 | 29 | 22 | 0.62 | 0.65 | Yes | 56 | 34 | 20 | 0.63 | 0.68 | No |
| SN74 | 52 | 43 | 20 | 0.65 | 0.71 | No | 61 | 24 | 7 | 0.54 | 0.56 | Yes |
| SN79 | 26 | 71 | 23 | 0.92 | 0.74 | No | 18 | 55 | 32 | 0.93 | 0.62 | No |
| SN82 | 48 | 63 | 24 | 0.76 | 0.86 | No | 69 | 52 | 16 | 0.63 | 0.86 | No |
| SN83 | 59 | 47 | 29 | 0.67 | 0.82 | No | 44 | 43 | 24 | 0.69 | 0.68 | No |
| SN86 | 51 | 66 | 18 | 0.74 | 0.88 | No | 69 | 61 | 4 | 0.62 | 0.81 | No |
| SN87 | 73 | 51 | 27 | 0.61 | 0.93 | No | 87 | 46 | 20 | 0.55 | 0.88 | No |

| | | | | | | | | | | | | |
|------|----|----|----|------|------|-----|----|----|----|------|------|-----|
| SN90 | 12 | 42 | 31 | 0.90 | 0.57 | No | 18 | 52 | 22 | 0.93 | 0.62 | No |
| SN93 | 50 | 68 | 17 | 0.74 | 0.85 | No | 65 | 67 | 22 | 0.67 | 0.91 | No |
| TF06 | 39 | 44 | 30 | 0.71 | 0.68 | No | 40 | 51 | 28 | 0.72 | 0.70 | No |
| TF16 | 50 | 58 | 12 | 0.68 | 0.74 | No | 51 | 52 | 32 | 0.70 | 0.80 | No |
| TF30 | 43 | 42 | 24 | 0.65 | 0.64 | No | 38 | 44 | 20 | 0.68 | 0.64 | No |
| TF31 | 43 | 63 | 17 | 0.78 | 0.83 | No | 45 | 53 | 33 | 0.67 | 0.68 | No |
| TF38 | 18 | 51 | 19 | 0.81 | 0.59 | No | 32 | 34 | 29 | 0.65 | 0.58 | Yes |
| TF39 | 28 | 58 | 22 | 0.82 | 0.68 | No | 28 | 55 | 16 | 0.81 | 0.68 | No |
| TF45 | 39 | 55 | 10 | 0.77 | 0.76 | No | 24 | 45 | 20 | 0.81 | 0.64 | No |
| TF46 | 70 | 29 | 11 | 0.59 | 0.76 | Yes | 66 | 40 | 11 | 0.60 | 0.73 | No |
| TF51 | 63 | 65 | 13 | 0.67 | 0.94 | No | 60 | 52 | 24 | 0.66 | 0.85 | No |
| TF52 | 47 | 56 | 37 | 0.74 | 0.82 | No | 68 | 16 | 48 | 0.52 | 0.55 | Yes |
| TF56 | 40 | 72 | 17 | 0.82 | 0.84 | No | 39 | 62 | 25 | 0.82 | 0.84 | No |
| TF57 | 16 | 55 | 21 | 0.90 | 0.62 | No | 12 | 47 | 23 | 0.87 | 0.57 | No |
| TF64 | 25 | 49 | 11 | 0.85 | 0.67 | No | 20 | 44 | 29 | 0.84 | 0.62 | No |
| TF69 | 17 | 41 | 15 | 0.84 | 0.60 | No | 15 | 51 | 20 | 0.93 | 0.62 | No |
| TF75 | 14 | 28 | 19 | 0.74 | 0.55 | Yes | 24 | 30 | 15 | 0.69 | 0.57 | No |
| TF78 | 55 | 41 | 11 | 0.66 | 0.75 | No | 46 | 35 | 18 | 0.68 | 0.70 | No |
| TF80 | 41 | 70 | 13 | 0.79 | 0.81 | No | 44 | 61 | 24 | 0.71 | 0.74 | No |
| TF81 | 40 | 61 | 16 | 0.78 | 0.78 | No | 28 | 44 | 10 | 0.78 | 0.66 | No |
| TF84 | 49 | 55 | 25 | 0.72 | 0.78 | No | 46 | 59 | 34 | 0.75 | 0.82 | No |
| TF85 | 15 | 45 | 13 | 0.85 | 0.59 | No | 18 | 50 | 27 | 0.85 | 0.61 | No |
| TF88 | 25 | 64 | 5 | 0.89 | 0.71 | No | 27 | 62 | 19 | 0.87 | 0.72 | No |
| TF89 | 63 | 45 | 12 | 0.61 | 0.75 | No | 57 | 54 | 14 | 0.69 | 0.86 | No |
| TF91 | 79 | 29 | 7 | 0.52 | 0.60 | Yes | 80 | 10 | 21 | 0.50 | 0.52 | Yes |
| TF92 | 21 | 53 | 16 | 0.81 | 0.61 | No | 12 | 42 | 21 | 0.87 | 0.57 | No |
| JP01 | 21 | 57 | 27 | 0.79 | 0.63 | No | 15 | 48 | 18 | 0.88 | 0.63 | No |
| JP02 | 36 | 67 | 30 | 0.79 | 0.82 | No | 37 | 52 | 25 | 0.77 | 0.80 | No |
| JP07 | 20 | 70 | 29 | 0.91 | 0.73 | No | 27 | 66 | 36 | 0.87 | 0.83 | No |
| JP09 | 25 | 50 | 31 | 0.78 | 0.66 | No | 33 | 59 | 21 | 0.80 | 0.81 | No |
| JP10 | 21 | 45 | 14 | 0.84 | 0.67 | No | 18 | 43 | 36 | 0.78 | 0.60 | Yes |
| JP11 | 45 | 58 | 24 | 0.73 | 0.84 | No | 54 | 55 | 26 | 0.69 | 0.93 | No |
| JP17 | 54 | 42 | 20 | 0.67 | 0.81 | No | 59 | 43 | 17 | 0.63 | 0.78 | No |
| JP20 | 33 | 52 | 28 | 0.78 | 0.75 | No | 30 | 64 | 19 | 0.83 | 0.82 | No |
| JP21 | 16 | 45 | 16 | 0.67 | 0.54 | No | 21 | 44 | 21 | 0.80 | 0.64 | No |
| JP25 | 38 | 29 | 17 | 0.64 | 0.61 | Yes | 42 | 43 | 18 | 0.68 | 0.71 | No |
| JP28 | 14 | 66 | 33 | 0.95 | 0.67 | No | 9 | 57 | 28 | 1.00 | 0.63 | No |
| JP33 | 20 | 22 | 41 | 0.73 | 0.58 | Yes | 26 | 51 | 27 | 0.77 | 0.67 | No |
| JP36 | 22 | 65 | 23 | 0.85 | 0.70 | No | 24 | 62 | 28 | 0.81 | 0.68 | No |
| JP41 | 41 | 64 | 27 | 0.76 | 0.85 | No | 47 | 57 | 19 | 0.70 | 0.82 | No |
| JP43 | 58 | 52 | 21 | 0.66 | 0.86 | No | 38 | 37 | 20 | 0.69 | 0.69 | No |
| JP49 | 33 | 66 | 34 | 0.83 | 0.83 | No | 35 | 52 | 37 | 0.72 | 0.70 | No |
| JP59 | 18 | 34 | 14 | 0.73 | 0.57 | No | 6 | 6 | 9 | 0.65 | 0.51 | Yes |
| JP65 | 32 | 62 | 36 | 0.80 | 0.76 | No | 33 | 67 | 26 | 0.84 | 0.90 | No |
| JP68 | 41 | 58 | 19 | 0.72 | 0.76 | No | 23 | 53 | 16 | 0.80 | 0.65 | No |
| JP70 | 25 | 65 | 9 | 0.86 | 0.74 | No | 26 | 56 | 25 | 0.82 | 0.73 | No |
| JP71 | 20 | 46 | 10 | 0.77 | 0.61 | No | 15 | 20 | 9 | 0.65 | 0.53 | Yes |
| JP76 | 24 | 61 | 30 | 0.85 | 0.72 | No | 31 | 62 | 23 | 0.78 | 0.75 | No |

Note for Table 3: INFP% is the percentage of participants who answered the item in the conventionally keyed positive direction (I, N, F, or P). Highest other scale r is the largest absolute magnitude of the correlation of the item with a non-keyed preference scale. All correlations are presented with decimals omitted. PR(INFP) is the prediction ratio of the item when answered in the keyed-positive direction, while PR(ESTJ) is the prediction ratio of the item when answered in the non-keyed-positive direction.

to replace every faulty item with an entirely acceptable replacement from the experimental items; a revised instrument produced from the current item pool would need to contain three SN items and one JP item that satisfy the prediction ratio criterion, but do not satisfy the criterion that the correlation between the item and the keyed preference scale is at least .10 greater than the correlation of the item with any of the other three preference scales.

There was little overlap between items that were unsatisfactory when translated into Traditional Characters and those that were unsatisfactory when translated into Simplified Characters. Only three original items were unsatisfactory in both translations.

There was little difference in the overall performance of phrase items and word pair items. For the Traditional translation, 41 of 46 phrase items (89%) and 42 of 47 word pair items (89%) were satisfactory. For the Simplified translation, 40 of 46 phrase items (87%) and 40 of 47 word pair items (85%) were satisfactory.

All statistics in other section of this paper were calculated using the original item set, including items that were not satisfactory when translated.

Overall Reliability. The internal consistency of the translated scales is shown in Table 4. For the Traditional translation, the internal consistency of the scales is fairly close to that of the English version of the scale (Myers et al, 1998). For the Simplified translation, this is also the case for the E-I and J-P scales, while the S-N and T-F scales are only slightly lower.

Validity of the Dichotomies. Correspondence Between Self-Estimated, Reported, and Best-Fit Type.

To determine whether participants' types as reported by the instrument were consistent with their independently self-estimated types, a comparison of reported and self-estimated types was conducted. To determine whether participants' initial self-estimate of their type was generally consistent with their final conclusion about their type, after having been presented with the results of the MBTI instrument and a full description of their type, a comparison of self-estimated and best-fit type was conducted. To determine whether participants' types as reported by the instrument were generally consistent with their final conclusion about their type, after having read a full description of their type, a comparison of reported and best-fit type was conducted.

Seven of the Traditional participants felt unable to self-estimate on one of the preferences. One of the Simplified participants felt unable to produce a best-fit type on one of the preferences. For simplicity of analysis, and to produce the most conservative estimate of the validity of the translated instruments, these cases have been counted as mismatches between estimates of the participants' preferences on that dichotomy.

Tables 5 and 6 show that in all cases correspondence between different estimates of type was significantly above chance. In almost every case, correspondence between reported and best-fit type was the highest, followed by the correspondence between self-estimated and best-fit type, and then between self-estimated and reported type.

Table 4. Internal Consistency of the MBTI preference scales (coefficient alpha)

| Scale | Traditional Chinese Character Translation (coefficient alpha) | Simplified Chinese Character Translation (coefficient alpha) | Number of Items |
|-----------------------------|---|--|-----------------|
| Extraversion – Introversion | .90 | .90 | 21 |
| Sensing – Intuition | .88 | .83 | 26 |
| Thinking – Feeling | .88 | .83 | 24 |
| Judging – Perceiving | .88 | .86 | 22 |

Table 5. Percentage agreement between different estimates of preferences (compared with 50% rate of agreement by chance)

| Preference | Traditional Chinese | | | Simplified Chinese | | |
|---------------------------|---------------------|-------|-------|--------------------|-------|-------|
| | SE-RT | SE-BF | RT-BF | SE-RT | SE-BF | RT-BF |
| Extraversion-Introversion | 85 | 89 | 92 | 89 | 91 | 96 |
| Sensing-Intuition | 78 | 84 | 94 | 82 | 90 | 89 |
| Thinking-Feeling | 73 | 82 | 91 | 77 | 83 | 88 |
| Judging-Perceiving | 77 | 83 | 93 | 77 | 85 | 90 |

Note: SE = self-estimated type; RT = type as reported by the inventory; BF = best fit type. All percentage agreements are significantly greater than chance ($p < .0001$).

Table 6. Percentage of participants obtaining each degree of agreement between different estimates of type

| Number of Preference Matches | Traditional Chinese | | | Simplified Chinese | | | Chance |
|------------------------------|---------------------|-------|-------|--------------------|-------|-------|--------|
| | SE-RT | SE-BF | RT-BF | SE-RT | SE-BF | RT-BF | |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6.25 |
| 1 | 6 | 2 | 3 | 1 | 0 | 1 | 25 |
| 2 | 15 | 11 | 1 | 15 | 11 | 4 | 37.5 |
| 3 | 39 | 34 | 19 | 41 | 28 | 25 | 25 |
| 4 | 40 | 53 | 77 | 43 | 61 | 70 | 6.25 |

Note: SE = self-estimated type; RT = type as reported by the inventory; BF = best fit type. In all cases the number of 4 preference matches is significantly greater than chance ($p < .0001$).

Table 7. Correlations between continuous preference scores of the translated MBTI Form M instruments

| | <i>EI</i> | <i>SN</i> | <i>TF</i> | <i>JP</i> |
|-----------|-----------|-----------|-----------|-----------|
| <i>EI</i> | – | -.09 | -.02 | -.01 |
| <i>SN</i> | -.19 | – | .27* | .37* |
| <i>TF</i> | .00 | .16 | – | .30* |
| <i>JP</i> | .00 | .39* | .11 | – |

Note: Correlations below the diagonal refer to correlations within the Traditional Chinese translation of the MBTI instrument. Correlations above the diagonal refer to correlations within the Simplified Chinese translation. $p < .01$.

Correlations Between the Preference Scales. Table 7 shows the interscale correlations of the two translated instruments. These correlations are based on raw score points, with (by convention) I, N, F, and P being scored in the positive

direction. For the Traditional Chinese character translation, the only significant correlation is between SN and JP. This pattern is similar to results obtained with the English version of Form M (Myers et al., 1998). For the Simplified Chinese character translation, there are significant correlations between SN and JP, SN and TF, and TF and JP.

Accuracy of Self-Scoring. The accuracy with which every participant had self-scored their profile was checked by comparing reported type as calculated by the participant, using the supplied scoring procedure, and reported type as calculated by computer from the participants' raw answers. For the traditional character translation, 42 of 100 participants made an error in self-scoring the questionnaire, while for the simplified translation, 38 of 102 participants made a similar error. Most errors were in counting or adding. In most cases this error was only of one or two raw score points. Only four of the Traditional participants, and five of the Simplified participants, miscalculated a reported

preference. In all but two cases, the “correct” reported preference was only one or two raw score points away from being reported in the other direction.

Discussion

The distributions of best-fit type within the two samples showed that when the translated MBTI instruments are used with a Chinese-speaking population, they produce a variety of best-fit types, rather than most participants obtaining the same best-fit type. These findings are consistent with the cross-cultural validity of the translated MBTI instruments. Previous research has suggested that perhaps the distribution of type in China is somewhat different from that in the US (Broer & McCarley, 1999). However, so far, no study (including the present one) has examined the distribution of type in a representative sample of the general Chinese population or the Chinese management population. Future research that addresses this need is desirable.

Item analysis showed that most of the translated items worked acceptably. In most cases where an item failed to perform acceptably, an alternative experimental item is available for use in a revised translation. There was little overlap between the items that failed to work in each translation. To the extent that Hong Kong and China share similar cultures, this suggests that most poor items resulted from specific difficulties in translation, rather than from differences between Western and Eastern culture. In contrast to the findings of Osterlind et al. (1994), there was no evidence that word pair items performed any differently from phrase items. The relative success of the word pair items in the present study may be due to the translation processes used, improvements made to the Form M revision of the MBTI instrument, or to the method of analysis employed (item analysis versus factor analysis).

The internal consistencies of both translations are reasonably close to that of the English version of the instrument, and correspond to a reasonably high degree of reliability (Murphy & Davidshofer, 1998). These findings are supportive of the reliability of the translated inventories. Future research could investigate the test-retest reliability of the inventories, and directly compare them to the English version using bilingual participants.

The correspondence between type as estimated by the inventories and type as independently self-estimated by the participants was in all cases well above chance. This provides evidence for the validity of the translated inventories – the descriptions of the participants by

the inventories were largely consistent with their own self-descriptions. Unsurprisingly, the correspondence between self-estimated type and best-fit type, and between reported type and best-fit type, was also high. This is perhaps to be expected, because best-fit type is decided by the participants with reference to their self-estimated type and their reported type. More significantly, reported type was a generally a better predictor of best-fit type than was self-estimated type. This suggests that, after reading the full-length descriptions of each type, participants felt that the MBTI inventory better described them than did their earlier self-estimates. This finding is consistent with the usefulness of the translated instruments as tools for self-exploration. Although it could be argued that participants were responding to social pressure to agree with the inventory’s results rather than with their earlier self-estimates, every effort was made by the administrators to emphasize that participants should choose the best-fit type that best described them. The levels of agreement between reported, self-estimated, and best-fit types were generally comparable to those reported for the English-language studies reported in Myers et al. (1998). Overall, these results are supportive of the validity of the translated instruments.

The pattern of correlations between the scales of the Traditional translation were similar to those between the scales of the English version of the instrument. The scales of the Simplified translation appeared to be more intercorrelated than the scales of the English version. Future research could examine whether this is a genuine difference, or a property of the specific translation or of the non-representative sample used.

One limitation of the present study was that the sample sizes were not sufficient to conduct meaningful factor or item response theory (IRT) analyses of the results. Future research could investigate the factor structures of the translated inventories. Also, research could use IRT to determine which items best discriminate between preferences, in a manner similar to Myers et al. (1998).

The analysis of the accuracy of self-scoring of type showed that errors in calculating raw-score scale totals were common, but that this rarely affected overall preference score. This shows that the translated inventories can be self-scored with confidence, although the introduction of computer scoring would provide a slight degree of extra accuracy.

In conclusion, overall, the translated instruments performed reasonably similarly to their English language counterpart. With the

replacement of items that encountered difficulty in translation, it may be possible to produce instruments with psychometric properties even closer to that of the original. The results are also

consistent with the existence in China and Hong Kong of the aspects of personality measured by the MBTI instrument, and by extension are consistent with their cross-cultural universality.

References

- Arnau, R. C., Green, B. A., Rosen, D.H., Gleaves, D. H., & Melancon, J. G. (2003). Are Jungian preferences really categorical?: an empirical investigation using taxometric analysis. *Personality & Individual Differences, 34*, 233 – 251.
- Bents, R., & Wierschke, A. (1996). Test-retest reliability of the Myers-Briggs Typenindikator. *Journal of Psychological Type, 36*, 42 – 46.
- Broer, E., & McCarley, N. G. (1999). Using and validating the Myers-Briggs Type Indicator in Mainland China. *Journal of Psychological Type, 51*, 5 – 21.
- Church, A. T. (2001). Personality measurement in cross-cultural perspective. *Journal of Personality, 69*, 979 - 1006.
- Fleenor, J. W. & Mastrangelo, P. M. (2001). Test review of the Myers-Briggs Type Indicator Form M. From B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* [Electronic version]. Retrieved September 12, 2003, from the Buros Institute's Test Reviews Online website: <http://www.unl.edu/buros>
- Freeman, D. (2004). *Presenting Type in Organizations* [Computer CD-Rom]. Melbourne, Australia: Australian Psychologists Press.
- Hofstede, G., & Hofstede, G. J. (2005). *Cultures and organisations: Software of the mind* (2nd ed). NY: McGraw-Hill.
- Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological Testing: Principles and Applications* (4th ed). Upper Saddle River, NJ: Prentice-Hall.
- Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (3rd ed). Palo Alto: Consulting Psychologists Press.
- Myers, I. B., Kirby, L. K., & Myers, K. D. (1998). *Introduction to Type®* (6th ed). Palo Alto: Consulting Psychologists Press.
- Nordvik, H. (1994). Type, vocation, and self-report personality variables: A validity study of a Norwegian translation of the MBTI, Form G. *Journal of Psychological Type, 29*, 32 – 37.
- Osterlind, S. J., Miao, D., Sheng, Y., & Chia, R. C. (2004). Adapting item format for cultural effects in translated tests: Cultural effects on construct validity of the Chinese versions of the MBTI. *International Journal of Testing, 4*, 61-73.
- Sim, H., & Kim, J. (1993). The development and validation of the Korean version of the MBTI. *Journal of Psychological Type, 26*, 18 – 27.
- Stalikas, A., & Fitopoulos, L. (1998). The translation and standardization of the MBTI into the Greek language. *Journal of Psychological Type, 46*, 13 – 21.
- Simplified Chinese Character*. (2005, August 23). Retrieved August 25, 2005, from Wikipedia website: http://en.wikipedia.org/wiki/Simplified_Chinese