The Factor Structure of Chinese Personality Terms

Xinyue Zhou,[1] Gerard Saucier,[2] Dingguo Gao,[1] and Jing Liu[1]

[1]Sun Yat-Sen University

[2]University of Oregon

ABSTRACT From the Contemporary Chinese Dictionary, 3,159 personality descriptors were selected and then ranked by the frequency of use. Among those, the top 413 terms with the highest frequency were administered to two independent large samples in China for self-ratings and peer ratings to explore the emic Chinese personality structure as well as to test the universality of other models. One- and two-factor structures found in previous studies of other languages were well replicated. Previous structures with more than two factors were not well replicated, but six- and seven-factor models were at least as well supported as the Big Five. Emic analysis indicated that a seven-factor structure was the most informative structure relatively salient across subsamples of self-ratings and peer ratings, across original and ipsatized data, and across differences in variable selections. These factors can be called Extraversion, Conscientiousness/Diligence, Unselfishness, Negative Valence, Emotional Volatility, Intellect/Positive Valence, and Dependency/Fragility.

This study explores the factor structure of Chinese personality descriptors, testing the universality of models derived from studies of other languages. It examines the between-language replicability of personality structure, thus serving a dual function. On a practical level it indicates with what degree of validity the model can be applied to Chinese, the native language of approximately one fifth of the world's population. On a theoretical level, it tests the Big Five

This research was supported in part by grants from the Ministry of Education of China (No. 06JC840001), 985-2 Research Program (No. 2006-90015-3272210), Youth Grant, and 100 Talents Program of Sun Yat-Sen University.

Correspondence concerning this article should be addressed to Xinyue Zhou, Department of Psychology, Sun Yat-Sen University, Guangzhou 510275, China. E-mail: zhouxyue@mail.sysu.edu.cn.

Journal of Personality 77:2, April 2009
© 2009, Copyright the Authors
Journal compilation © 2009, Wiley Periodicals, Inc.
DOI: 10.1111/j.1467-6494.2008.00551.x

model's universality across languages in comparison with alternative competitor models. At the same time, it develops an "emic," within-language scientific taxonomy of personality descriptors in the Chinese language.

Lexical Studies and the Big Five Model

Goldberg (1981) described a "lexical hypothesis," which proposes that the most important phenotypic attributes tend to be encoded as single words in the natural language. Based on such an assumption, one can conduct "lexical studies," in which a dictionary is used to obtain a comprehensive list of important personality descriptors, enabling investigation of the personality factor structure within a certain language (Saucier & Goldberg, 1996). Saucier and Goldberg (2001) provided a detailed review describing and comparing lexical studies. The lexical approach has led to several useful alternative models of personality structure.

Based mainly on lexical studies using English adjectives (Goldberg, 1990), the Big Five factor structure now represents the most popular lexically derived personality structure. These five factors are commonly labeled as Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect. The Big Five model formed a template for a highly influential personality inventory (Costa & McCrae, 1992) and has become the most widely tested personality factor model in lexical studies (Saucier & Goldberg, 2001).

Several studies have successfully replicated the Big Five model, though not necessarily precisely. On the basis of studies in English, Dutch, and German, Hofstee, Kiers, De Raad, Goldberg, and Ostendorf (1997) concluded that the Big Five replicated well. In addition to these three languages, the Big Five model was also replicated with slight differences in other languages: Polish (Szarota, 1996), Czech (Hřebíčková, Ostendorf, & Angleitner, 1995), Croatian (Mlacic & Ostendorf, 2005), and Turkish (Somer & Goldberg, 1999).

But out of 15 languages studied to date, the Big Five was found readily in only these 7. Indeed, the results of some other studies cannot be easily mapped onto the Big Five space. In independent studies in Italian (Di Blas & Forzi, 1998, 1999) and Hungarian (Szirmak & De Raad, 1994), for example, the Agreeableness factor split into two distinct factors, and the five-factor solution contained no factor interpretable as Intellect. Similar divisions of the Agree-

ableness factor were found in studies of Korean (Hahn, Lee, & Ashton, 1999) and French (Boies, Lee, Ashton, Pascal, & Nicol, 2001), leading to conceptions of a six-factor model.

The Cross-Language Six Model

A six-factor structure has emerged from self-ratings of familiar personality-descriptive adjectives in eight studies of seven different languages (Ashton, Lee, Perugini, et al., 2004). Ashton, Lee, Perugini, et al. (2004) delineated the features of the six-factor structural pattern: Conscientiousness and Intellect (Openness) factors tend to be only slightly different from their Big Five counterparts. Extraversion becomes a factor emphasizing sociability, and Emotional Stability is replaced by a factor emphasizing self-assurance versus emotional vulnerability. Agreeableness becomes a factor emphasizing even temper versus hostility. Another factor (the one least related to the Big Five) has honesty and humility at one pole and deceit and egotism at the other. This adjective-based lexical structure, represented in Ashton, Lee, Perugini, et al. (2004), is here labeled the Cross-Language Six (CL6).

Seven-Factor Models

Some lexical studies have utilized a more inclusive variable selection, including terms that are highly evaluative, that refers to the consistent effects a person has on others, or even those for physical attributes. Most such terms appear to fit within the usual definitions of personality as manifest patterns discernible in an individual's behavior, thinking, or emotions. By omitting such terms, investigators may miss important factors. By being inclusive, one can study the structure of both narrow and inclusive variable selections and usefully compare these.

Inclusive-selection lexical studies have most often reported seven-factor structures. In a study that sampled descriptors from an English dictionary, Tellegen and Waller (1987; summarized by Waller & Zavala, 1993) found a "Big Seven" factor structure; there were two additional factors—Negative Valence (e.g., awful, terrible) and Positive Valence (e.g., outstanding, impressive)—besides variations on the Big Five. Moreover, instead of Intellect there was a Conventionality factor. Some but not all aspects of the Tellegen and Waller structure were then replicated in Spanish (Benet-Martìnez & Waller, 1997) and Hebrew (Almagor, Tellegen, & Waller, 1995). Other

studies (Goldberg & Somer, 2000; Saucier, 1997) found a variant structure with Intellect and Attractiveness factors instead of Positive Valence and Conventionality.

Based on convergent results from the Filipino (Church, Reyes, Katigbak, & Grimm, 1997) and Hebrew (Almagor et al., 1995) lexical studies, Saucier (2003a) laid out an alternative seven-factor structure. In this Multi-Language Seven (ML7), Conscientiousness and Intellect factors are fairly similar to those in the Big Five, and a Negative Valence factor similar to that factor in the Big Seven, but the other four factors are Gregariousness, Self-Assurance, Even Temper, and Concern for Others (vs. Egotism). Essentially, an omnibus Big Five Agreeableness factor is replaced by three (still broad and mutually independent) factors of Even Temper, Negative Valence, and Concern for Others.

How does the replicability of this seven-factor model compare with that of the Big Five? Based on a wider selection of variables, the ML7 has been examined in fewer studies, making comparison more difficult. In studies of English language–type nouns (Saucier, 2003b), the ML7 replicated about as well as the Big Five, and in studies of Greek (Saucier, Georgiades, Tsaousis, & Goldberg, 2005), about as well as either the Big Five or the CL6. However, in neither of these studies did any previous structures of five to seven factors replicate impressively.

Models With Fewer Factors

But the structures that lexical studies indicate are the most recurrent across variable selections—not to mention languages—are those with only one or two factors. Several studies (Boies et al., 2001; Di Blas & Forzi, 1999; Goldberg & Somer, 2000; Saucier, 1997, 2003b; Saucier et al., 2005) have reported that when a single factor was extracted, it is one that contrasts desirable attributes with undesirable attributes, which can be labeled Evaluation. Indeed, judging the meanings of diverse objects in a wide array of cultural settings, Osgood, May, and Miron (1975) found that a global evaluation factor (good vs. bad) was the single largest factor.

Numerous lexical studies (e.g., Boies et al., 2001; Caprara, Barbaranelli, & Zimbardo, 1997; Di Blas & Forzi, 1999; Goldberg & Somer, 2000; Hřebíčková, Ostendorf, Osecká, & Čermák, 1999; Saucier, 1997, 2003b; Saucier et al., 2005; cf. Shweder, 1972; White, 1980) show a consistent pattern in the two-factor solution: a

Dynamism factor, with desirable qualities related to individual ascendancy, and a Social Propriety factor emphasizing morality, solidarity, and virtue. Like the one-factor structure, this structure has been robust across adjective and noun domains, whereas more differentiated structures of type-nouns have differed from those based on adjectives (Saucier, 2003b).

Two dimensions underlying psychopathology—internalizing and externalizing disorder tendencies (Krueger, Chentsova-Dutton, Markon, Goldberg, & Ormel, 2003; Krueger & Tackett, 2003)—offer an alternative conception of what might arise in two-factor structures. Internalizing disorders include depression, anxiety, and phobias; externalizing disorders include psychopathic and substance use disorders. Externalizing disorders probably involve low Morality/Social Propriety and internalizing disorders low Dynamism.

A Big Three model—including broad versions of Extraversion, Agreeableness, and Conscientiousness—also seems to be more consistent than the Big Five. The Big Three was found in three-factor solutions in English (Peabody, 1987; Saucier, 1997), Turkish (Goldberg & Somer, 2000; Somer & Goldberg, 1999), German (Ostendorf, 1990), Italian (Di Blas & Forzi, 1999), Hungarian (Szirmak & De Raad, 1994), and Korean (Hahn et al., 1999).

Indications Regarding the Emic Personality Structure of Chinese

K. Yang and Bond (1990) conducted a study, although not a standard lexical study, of Chinese descriptors in Taiwan. Descriptors were selected from 557 entries that had been previously pooled from a variety of printed media. Then participants were asked to describe one of the six target persons with the selected 150 descriptors. Factors that emerged were Social Orientation–Self–Centeredness, Self–Control–Impulsiveness, Competence–Impotence, Expressiveness–Conservatism, and Optimism–Neuroticism.

The first attempt to examine the personality structure in Chinese language in mainland China was made by Wang, Fang, and Zuo (1995). They extracted 6,156 terms describing individual differences from a dictionary. They also gathered 1,638 terms that were used to describe people from their daily lives from 352 college students. Five scholars classified these terms into these categories: stable personality trait, social evaluation, expression and emotionality, and body movement. The first category was further divided into three subcategories:

terms describing individual traits that are not related to other people, terms describing traits that are displayed in interpersonal situations, and terms related to the outside world. From 507 terms for traits displayed in interpersonal situations, researchers extracted 165 high-frequency terms and had 300 students rate themselves on a 5-point scale, and two factors were extracted (dominance/arrogance and alienation/submission). In other work (e.g., Cui & Wang, 2003), on the basis of unpublished analyses of ratings using Wang et al.'s (1995) terms as well as additions, this team produced a measure for Chinese personality description with seven factors, including Extraversion, Kindness, Talents (likely related to Conscientiousness), Emotionality, Human Relations (related to warmth), Ways of Life (related to ambition), and Behavior Styles (contrasting social propriety with dynamism); the first four factors appear comparable to factors found in various lexical studies, more so than the last three. Wang, Cui, and Zhou (2005) argued that the Big Five is less applicable to the Chinese context than is their own model.

Another study of a lexical nature examined the structure of ancient Chinese personality terms (B. Yang, 2005). The researcher extracted 1,241 personality terms from an ancient history book (Shi-Ji) and further classified these terms into 98 broad personality traits. Then 118 scholars rated 102 historical figures in that book on each of the 98 traits. Four factors emerged from the ratings: Ren (love, caring), Zhi (intelligence), Yong (courage), and Yin (being a hermit). It is worth noting that the first three of these factors seem to correspond to three principal virtues touted by the ancient Chinese philosopher Kong-zi (Confucius, 1983).

These previous studies of Chinese personality descriptors have led to diverging results, perhaps because of wide variance in method. The studies differed in how lexical items were identified (from a dictionary, from free descriptions, from an ancient history book, or from printed media) and in who was rated (the self, selected target persons, or figures from history). The work of Wang, Fang, and Zuo (1995) comes closest to the methodology of a standard lexical study and, unlike the other studies, did not consistently eliminate synonyms.[1] However, this study went well beyond the dictionary and

1. As a rule, elimination of synonyms is not carried out in standard lexical studies because this violates the lexical rationale, by which clusters of related terms are taken to indicate especially important aspects of personality description, indicating "many words for that construct."

culled out a very large number of descriptors considered evaluative. Finally, previous studies of Chinese personality descriptors have not included a fair and rigorous comparative test of etic structures from other languages.

The term-selection and term-elimination procedures in these previous studies of Chinese personality terms did not follow standard lexical procedures used in other languages, making it difficult to compare Chinese results with those from elsewhere. And thus far, there is no agreed-upon scientific taxonomy of personality descriptors in contemporary Chinese. Given the importance of the Chinese language and the sizable proportion of the world's population that uses this language, this reveals an important gap in knowledge about personality structure.

Toward an Optimal Methodology for Lexical Studies

The present studies of Chinese descriptors attempt to establish an emic classification—a structure of personality attributes derived directly from the Chinese lexicon. They also compare the replicability of Big Five structures with alternative competitor structures, addressing a larger issue: Which structure is most universal or most invariant across languages and cultural settings?

Both peer ratings and self-ratings were included. Self-ratings and peer ratings each represent valuable perspectives on personality (Funder, 2005). In the early stages of the development of studies of the natural language of personality, both self- and peer ratings were common. For example, Norman (1963) used both, as did the earliest lexical studies in English (Goldberg, 1990, 1992; Saucier, 1997; Saucier & Goldberg, 1996), Dutch (De Raad, Hendriks, & Hofstee, 1992) and German (Ostendorf, 1990). However, self-report is generally more convenient for and familiar to personality psychologists, and after these early studies investigators tended to revert to the practice of relying on self-report only. This is unfortunate: In some non-Western settings self-ratings are a more awkward format than peer ratings. Talk about the characteristics and tendencies of others (including the phenomena of gossip and reputation; Bailey, 1971; Haviland, 1977) is likely to be universal across cultures, whereas making representations to others about one's own traits may be more common in societal contexts (e.g., urban settings) where one frequently encounters strangers. Substantial and system-

atic differences between structures of self-ratings and of peer ratings have not yet been observed. However, a thorough lexical study will not assume the lack of such differences but rather will examine the replicability of results across the two types of data. In this study, both self-rating data and peer-rating data were analyzed, with original data and with data that had been ipsatized (row-standardized so that each case has the same mean and variance). Ipsatization beneficially removes individual differences in use of the response scale (i.e., differences in mean and variance) from the data, and it is more likely to lead to bipolar factors. However, there is controversy over whether ipsatization is always appropriate.[2] If one examines only ipsatized-data results, one will not understand the result of ipsatizing. Conversely, if one examines only original-data results, information regarding the effect of response-scale usage on structure will be unavailable. So we examined results with both types of data.

In this study, we collected person descriptors in the Chinese language from a dictionary and retained the most frequently used terms. Then, we derived an emic factor structure from these variables based on self-ratings and peer ratings and compared this structure to the Big Five and other candidate models for personality structure.

METHOD

Extraction of Person-Descriptive Terms From the Dictionary

For the first steps, we employed the same modified methodology outlined by Angleitner, Ostendorf, and John (1990) and Saucier et al. (2005). Three independent native speakers—psychology-student research assistants from different areas of China, each speaking a different dialect—examined the most extensively used dictionary in mainland China, the newest edition of the Contemporary Chinese Dictionary (Chinese Academy of Social Sciences, Institute of Linguistics, 2005), including 65,000 entries. This dictionary includes not only common Mandarin words, but also widely recognized words from other dialects, as well as ancient words still in use and some professional/technical terms. We extracted all the

2. Ipsatization can produce artifactual bipolarities (Dunlap & Cornwell, 1994; ten Berge, 1999). Common natural-language person descriptors do not necessarily have balanced keying; for example, there are far more words in English for negative affectivity than for its absence, so individual differences in response mean are confounded with reported negative affectivity. Ipsatization is a partialing procedure that can partial out not just style but also substance.

person-descriptive entries from this dictionary. Adjectives and nouns were both included; indeed Chinese descriptors (of personality at least) tend to be able to function as either adjectives or nouns. Moreover, our approach actually did not exclude all verbs, because as a rule Chinese adjectives easily function as stative verbs (e.g., the term for "shy with strangers" also means "to be shy with strangers"). A single Chinese word may function as an adjective, noun, and verb, depending on the context. We note that the written Chinese language referenced in this dictionary, due to the policy in effect for over 2,200 years that standardizes the written script, generalizes across many dialects (e.g., Mandarin, Cantonese).

Before they started the search, the three judges were instructed that the term selected should fit into a sentence such as "How [adjective] am I?" They were also instructed to exclude terms falling into any of four categories: (a) nondistinctive and applicable to all individuals (e.g., human, born); (b) referring to geographical origin (e. g., Shanghainese), to nationality (e.g., Chinese), or to professional- or job-related identities (e.g., student, teacher); (c) referring to only a part of the person (e.g., shining eyes, long legs) or appearance; or (d) having personality implications that are both metaphorical and tenuous (e.g., mouse, rose). All these instructions were adapted from Saucier et al. (2005) and were given in Chinese.

The first judge selected 2,330 terms, the second 2,853, and the third 2,920. Reliability for these selections (generalized kappa) was .85. Descriptors selected by any one of the judges were included in the initial list, which included 3,159 terms. Next, we culled out unfamiliar terms.

Ratings of Frequency of Use

Major frequency dictionaries for contemporary (simplified character) Chinese were compiled and published in the 1980s, are now relatively out of date, fail to include many commonly used personality descriptors, and, even where they do, fail to specifically reference the frequency of use in describing a person. Accordingly, as in many previous lexical studies, we gathered frequency ratings. A new set of 10 judges rated the terms on the list developed earlier. These judges were volunteers recruited from an introductory psychology class and hailed from six different provinces in China, with four being from rural areas and six from cities. These judges were asked to rate the degree to which they believed the term is frequently used for the description of a person. Ratings were on a 5-point scale: 1 (*this word is never used for the description of a person*), 2 (*this word is rarely used for the description of a person*), 3 (*this word is sometimes used for the description of a person*), 4 (*this word is often used for the description of a person*), and 5 (*this word is extremely often used for the description of a person*).

Interrater correlations for the frequency judgments ranged from .53 to .78, with an average r (after Z transformation) of .61, and coefficient alpha of .93.

These terms were rank-ordered on the basis of their mean frequency rating. We retained the 413 terms that received an average rating of 4.5 or more from the 10 judges.

Questionnaires and Participants

Self-reports were collected from 451 university students from Fudan University in Shanghai. They were asked to rate themselves on the 413 descriptors on the list produced earlier. The age of the participants ranged from 17 to 24, and 49% were women, 51% men. Ratings were on a 5-point scale: 1 (*this word is very inaccurate for the description of me*), 2 (*this word is somewhat inaccurate for the description of me*), 3 (*this word is neither accurate nor inaccurate for the description of me*), 4 (*this word is somewhat accurate for the description of me*), and 5 (*this word is very accurate for the description of me*).

Peer reports were provided by 300 students from Fudan University (Shanghai) as well as 200 students from Sun Yat-Sen University in Guangzhou (formerly Canton). Participants rated peers on the 413 descriptors on the list produced earlier. The age of the participants ranged from 17 to 23 (58% female, 42% male). Ratings were on the same 5-point scale as for self-ratings but adapted for peer ratings. In order to more representatively sample the population of well-acquainted peers, who are not all the very best-liked people, half of the participants were randomly assigned to rate someone they know well and like, and the other half of the participants were asked to rate someone they know well but who is not the person they like best (i.e., not necessarily disliked, but not the best-liked person). There is no indication from previous studies (e.g., Goldberg, 1992; Saucier, 2003b) that such relatively small variations in the liking of targets have a notable effect on structure, nor did we find strong evidence of such effects in these data.

Participants filled out questionnaires in classes, in exchange for partial course credit.

Variable Selections

Because inclusive variable selections contain a wider range of variables, they provide more information than narrow, conventional selections of personality variables. Accordingly, we give prime emphasis to the inclusive full set of 413 variables. However, in order to examine the effects of a narrower variable selection on results, we utilized the same four exclusion

categories as Wasti, Lee, Ashton, and Somer (2008) in a study of six-factor structures in Turkish.

Following the method detailed in that study, 10 raters were instructed to exclude the terms that (a) primarily evaluate someone favorably, without any specific description, or (b) primarily are used as terms of insult or negative evaluation rather than as objective descriptions or (c) primarily describe the effect that someone's characteristics have on others, rather than the actual characteristics that someone has, or (d) primarily describe someone's physical appearance or physical characteristics. For these exclude-versus-include judgments, interrater r ranged from .48 to .75 (mean after Z transformation .63), with coefficient alpha of .92. A term was excluded if at least 7 raters out of 10 agreed that a term fell into one of these four categories.[3] By application of this method, 51 terms were excluded to form a narrower selection of only 362 descriptors. Results from this narrower selection were compared with those from the full selection.

Emic-Structure Factor Analyses

One primary aim was to derive the best emic structure for Chinese descriptors. This aim is exploratory and does not involve imposition of a priori structural hypotheses. Factor analyses were principal components with varimax rotation, the same rotation method used in previous lexical studies. As in most previous lexical studies, we sought the structure that had the largest number of independent sources of variance (i.e., orthogonal factors) that were easily interpretable. Another criterion for a good emic structure was that the structure appeared in relatively similar form in both of our samples, whether self- or peer ratings.

Etic-Structure Analyses

From previous lexical studies in other languages, there already exist several structural models for personality attributes, including those structures of one, two, three, five, six, and seven factors described earlier. Another purpose of this study is to test the cross-cultural generalizability of these structures by using the prime set of Chinese descriptors created

3. A "6/10" (i.e., majority) criterion would have led to removal of numerous terms (e.g., for hypocrisy, arrogance, diligence, and talkativeness) that are prototypical for lexical factors endorsed by proponents of narrow variable selections. With an even stricter "8/10" criterion, only 24 terms would have been removed, likely meaning that results would differ very little from those with the full set of 413 terms. Thus, the "7/10" criterion seemed the optimal compromise.

within our study. Based on the hierarchical pattern of structures identified in previous studies, our hypotheses posited the following: (1) a factor contrasting desirable and undesirable attributes if just one factor is extracted; (2) two factors related to Dynamism and Morality/Social Propriety when two factors are extracted and rotated or, alternatively; (3) two factors related to internalizing and externalizing disorder tendencies; (4) three factors interpretable as Extraversion, Agreeableness, and Conscientiousness when three factors are extracted and rotated; (5) a Big Five structure when five factors are extracted and rotated, particularly when the narrower variable selection is used; (6) a Cross-Language Six structure when six factors are extracted and rotated, particularly when the narrower variable selection is used; (7) a Multi-Language Seven structure when seven factors are extracted and rotated.

To represent the imposed-etic structures in the data, prior to any analyses the second author selected markers for the Big One, the Big Two, the Internalizing and Externalizing factor duo, the Big Three, the Big Five, the Cross-Language Six, and the Multi-Language Seven from the terms (in English translation) in our variable list.[4] Although these marker scales are relatively short (4–10 items each), they enabled a relatively fair comparison among a wide variety of structures. Except for CL6 Emotionality, the marker scales had coefficient alpha reliability of .58 to .77 in self data and .45 to .77 in peer data; the coefficients for Emotionality (six terms) were only .39 (self) and .28 (peer). Ratings on the markers for each scale were summed into a single score, representing a particular factor within a structure. These scale scores were then correlated with factor scores derived from the corresponding emic-structure solution of one, two, three, five, six, or seven factors. These correlations allowed us to compare the degree of replication of the various imposed-etic structures in Chinese data. We note that moderate (or even low) internal consistency for a brief marker scale tapping a very broad construct may not be a major problem. Although the square root of reliability (above .50 for all scales described above) does set a ceiling on correlations with external variables, this is contingent on certain assumptions (such as scale unidimensionality) that may be inapplicable in the situation of these data. Relevant indications can be gathered from comparing the performance of shorter and longer scales tapping the same construct but differing in internal consistency.

Fortunately, longer marker scales were available for representing certain structures. For the Big Five, additional items (not found among the 413 terms) necessary to create larger marker scales (from Goldberg's

4. The list of marker terms is available at www.uoregon.edu/~gsaucier/gsau5.htm.

[1992] 100 unipolar markers, 20 markers per factor) were added to the end of the materials; these scales had coefficient alpha values of .78 to .86 (self) and .75 to .87 (peer). For the Cross-Language Six, we were able to utilize the tables provided by Ashton, Lee, Perugini, et al. (2004) to compose longer marker scales (using 123 terms, also about 20 terms per factor) in total from the total variable set; these scales had coefficient alpha of .79 to .89 (self) and .73 to .89 (peer), with Emotionality now represented by a scale with workable reliability (.83 self, .80 peer). Unfortunately, there was no strong reference source to guide us in creating longer marker scales for other structures but, given their importance in past research, examining the comparative replication of the Big Five and the Cross-Language Six was our highest priority.

If etic structures like the Big Five are truly inferior for application to Chinese populations (as per Wang et al., 2005), our results should replicate previous emic structures better than they replicate imposed-etic structures. We located, within the set of Chinese terms administered to our samples, marker terms for factors from two emic structures from previous quasi-lexical studies of contemporary Chinese. These included 64% of the marker terms specified by Cui and Wang (2003), 79 terms including 3 to 20 for each of seven emic (E7) factors; these marker scales had coefficient alpha of .59 to .85 in peer data and .58 to .87 in self data. We also located 79% of the terms specified by Yang and Bond (1990)—65 terms including 9 to 16 for each of five emic (E5) factors; these marker scales had coefficient alpha of .71 to .77 in peer and .70 to .81 in self data. Based on widely understood assumptions of cultural relativism, and the viewpoint of Wang et al. (2005), we would expect to replicate these previous Chinese-emic structures better than any imposed etic. Based on the assumption that there are cross-cultural universals but that variations in scientific procedures affect their identification, we would expect the reverse (as our procedures resembled those in studies that discovered these imposed-etic models, more than those in previous Chinese emic structures).

RESULTS

First, we deleted cases with more than half of the responses missing, resulting in removal of 27 self-rating protocols and 10 peer-rating protocols. Then, based on the distribution of participant response means, we deleted seven self-rating and five peer-rating protocols that had a mean at least 2.5 *SD* away from the average response mean. After these reductions, we had 417 self-rating and 485 peer-rating protocols, for a combined sample size of 902.

The ratio of participants to variables exceeded 1.0 but was not high. Such a modest ratio is not unusual in lexical studies, and is unproblematic. Factor-solution stability depends largely on the sampling error of the correlation coefficient, which decreases with the square root of sample size, so that sample sizes over 300 generally have adequate loading-pattern stability (cf. Ashton, Lee, & Goldberg, 2004). A huge variable-to-factor ratio (here, at least 50:1) also contributes to such stability (Guadagnoli & Velicer, 1988; MacCallum, Widaman, Zhang, & Hong, 1999).

Emic-Structure Analyses

We conducted exploratory factor analyses on the full set of 413 descriptors, separately for self- and peer rating samples. The ipsatized-data analyses yielded factors that were more prone to be bipolar and were somewhat more consistently interpretable, so we gave priority to factors derived from these analyses. However, we also describe below the variations that occur when relying on original (nonipsatized) data instead. We further describe below the variations that occur when relying on the narrower variable selection, although these turned out to be generally only small variations on the results obtained with the inclusive variable selection.

For ipsatized self-ratings, the first eight eigenvalues accounted in order for 16.2%, 6.3%, 4.3%, 2.8%, 2.0%, 1.9%, 1.7%, and 1.4% of variance. These percentages for ipsatized peer ratings were 11.6%, 6.0%, 3.8%, 2.8%, 2.0%, 1.9%, 1.5%, and 1.3%. With a sample size of 417 (as in our self-ratings), correlations under .10 are not statistically significant ($p < .05$), and none of the 413 terms had all their loadings under .10 when as many as seven factors were extracted and rotated.

Figure 1 depicts the pattern of factor emergence as successively more factors are extracted and rotated in the ipsatized self-ratings (for all 413 terms). Figure 2 depicts the corresponding pattern for ipsatized peer ratings.

In either figure, the one-factor structure contains a dimension contrasting kindness and genuineness (and other favorable attributes) with hypocrisy and selfishness (and other unfavorable attributes). One-factor structures based on original data were similar, as were those based on the narrower selection of 362 terms.

In either figure, the two-factor structure contains one factor contrasting kindness, innocence, and honesty with arrogance and other

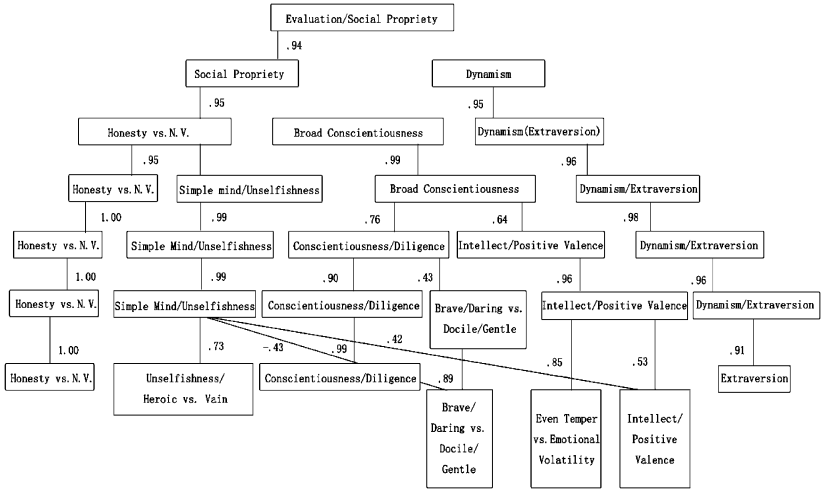

Figure 1
Pattern of factor emergence for ipsatized self-ratings.

attributes indicating social impropriety. The other factor contrasts talkativeness and extraversion with being self-conscious and introverted, with also a prominent infusion of intellect-related content.

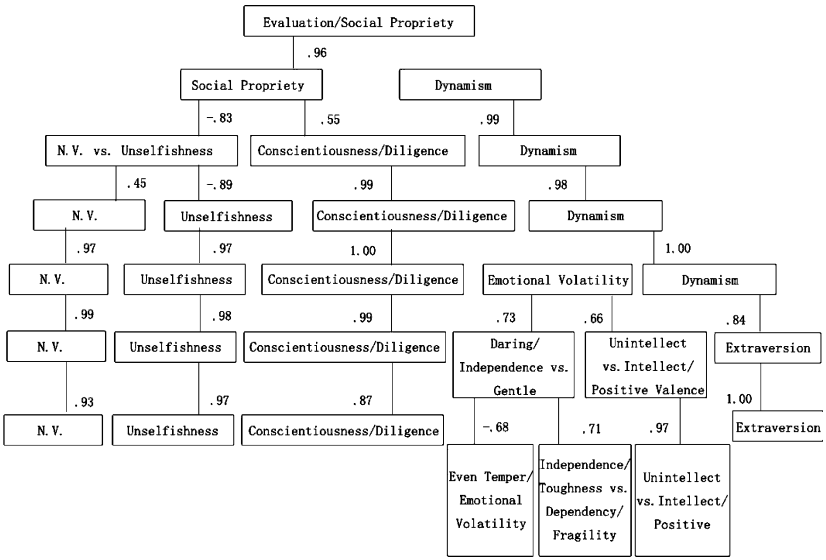

Figure 2
Pattern of factor emergence for ipsatized peer ratings.

Similar Social Propriety and Dynamism factors were found in original data and among the 362 terms in ipsatized data.[5]

In the three-factor structures, the first two factors remained similar but a Conscientiousness/Diligence factor was added, with the most salient terms being diligent, meticulous, and strict versus lazy, inconsistent, and careless. In original data, the third factor could be labeled as Negative Valence—or as ascribed tendencies to Noxious Violativeness of the rights of others (Saucier, 2007); we will call this factor NV from this point; as is true in most other lexical studies (Saucier, 2007), in all variants of the emic Chinese structure, NV content emphasized socially condemned cruelty and amorality.

In the four-factor solutions, in either figure, one sees the Social Propriety factor split into an NV factor and a second "Unselfishness" factor that, while varying somewhat between self- and peer ratings, consistently contrasted magnaminity and undemandingness with self-importance, vanity, overbearingness, and competitiveness. From the standpoint of psychopathology, this split of Social Propriety could represent a split between harmful antisocial tendencies (NV) on the one hand and narcissism (selfishness) on the other. Original data yielded very similar factors.

The five-factor solutions differed substantially according to whether we used peer or self data, although in either case three of the first four factors—Extraversion, Conscientiousness, and Unselfishness—remained relatively constant, and there was a factor with substantial NV content. The fifth self-rating factors in ipsatized data was Intellect/Positive Valence, but the additional peer-rating factor was Emotional Volatility. In original data, there were similar factors, differing similarly between self- and peer ratings. With the set of 362 terms, structures were similar, except that in ipsatized self-ratings the Intellect/Positive Valence factor disappeared in favor of an unusual factor combining (at the same positive pole) masculinity, honesty, and kindness; in ipsatized peer ratings the Emotional Volatility factor was absent, replaced by an unusual factor that was quite difficult to interpret and label. As these descriptions indicate, the emic five-factor structures were unstable, differing quite substantially from one analysis to another.

5. However, in original data with the 362 terms, only, the two-factor solution looked different, having one factor including most favorable terms and another including most unfavorable terms.

The six-factor solutions in ipsatized peer and self-ratings corresponded rather better, though still imperfectly. Five factors were discernible in either type of data: Extraversion, Conscientiousness/Diligence, Intellect/Positive Valence, Unselfishness, and a factor best labeled Dependency/Fragility (possibly related to perceived femininity versus masculinity, especially in the peer ratings). In peer ratings, the sixth factor was NV; in self-ratings, NV was combined with Honesty at opposite poles of the same factor. In original data, there were similar factors, differing similarly between self- and peer ratings. With the set of 362 terms, essentially the same factors as above were observed in peer ratings and in self-ratings, respectively.

In seven-factor solutions, differences between self- and peer-rating structures were rather less pronounced. Both structures had highly similar Extraversion, Conscientiousness, NV, and Intellect/Positive Valence factors, moderately similar Emotional Volatility and Dependency/Fragility factors, and another factor (moderately similar across data sets) referencing Unselfishness (pure-heartedness versus self-importance).[6] "Moderately similar" means there were clearly both data-set-specific and common-across-data-set aspects of factors.

In original data, there were factors similar to the ones found in ipsatized data. With the narrower set of 362 terms, the set of seven factors tended to be quite similar, although the NV factor morphed into Honesty/Genuineness versus Hypocrisy/Duplicity. Terms translated as "unique," "special," and "outstanding," representing Positive Valence, survived the variable reduction and were associated with the Intellect factor regardless of variable selection.

Certainly the peer- and self-derived seven-factor structures were not perfect replicas of one another. For matched pairs of the seven factors, coefficients of factor congruence ranged from good (Unselfishness .96, NV .95, .Extraversion .92, Emotional Volatility .89) to marginal (Conscientiousness .77, Dependency/Fragility .77, Intellect/Positive Valence .68), but indicated definite convergence. These seven-factor structures were relatively more similar across the diverse analyses than was true for solutions of five and six factors. Moreover, factors added in solutions of eight or more factors were

6. These factors were stable across subsamples of best-liked vs. more moderately liked peers. The only exception: Among moderately liked peers, Emotional Volatility content tended to load at the low end of Conscientiousness and was replaced by a difficult-to-interpret factor with heterogeneous content.

low in saturation (factor-loading magnitude), heterogeneous in content, and difficult to interpret. So we adopted the seven-factor structure as the most informative emic structure for Chinese personality descriptors based on these data.

Table 1 presents the terms (in English translation) most associated with each of the seven factors, across self- and peer ratings. Included in the table are only those terms (up to 20 per factor) with the highest average loading, across self- and peer ratings in ipsatized data. All coefficients in Table 1 involve ipsatized data.

Given the relative similarity of the seven-factor structures between the self- and peer-rating samples, we might expect to find the same factors when the peer and self samples were aggregated and analyses conducted on the single "supersample." Indeed, in such an analysis (using ipsatized data and all 413 terms) the factors were Unselfishness, NV, Extraversion, Conscientiousness/Diligence, Emotional Volatility, Intellect/Positive Valence, and Dependency/Fragility, with the same terms as those shown in Table 1 among those having the highest loadings on the factors. Thus, if we had begun by combining our self and peer data, we would have arrived at the same structure depicted in Table 1.

Examining the Replicability of Previous Etic Structures

How closely do the emic factors correspond to factors found in previous studies? We tested our seven hypotheses by computing correlations between marker scales and the most appropriate emic-factor solution with the same number of factors. Table 2 presents tests of our seven hypotheses. Each imposed-etic candidate structure corresponds to a row in the table, where its correlations with the actual emic factors are shown in rank order from highest to lowest. These correlations are those between best-matched pairs of etic marker scales and emic factor scores, such that each marker scale is matched with one factor-score variable in a way that maximizes the average correlations between sets. The table provides replication with peer and self-rating samples using factors from the full set of 413 terms, in both ipsatized and original data. Because previous studies focusing on the Big Five and Cross-Language Six have utilized mainly ipsatized data and a narrower variable selection, the table also provides coefficients for peer- and self-rating samples using factors from the narrower set of 362 terms in ipsatized data.

Table 1
Terms With the Highest Loadings on Each of Seven Emic Chinese Personality Factors

Chinese	English	Peer	Self
Extraversion			
waixiang 外向	Extraverted, Outgoing	.72	.67
huoyue 活跃	Active, Dynamic	.68	.59
kailang 开朗	Open, Frank, Optimistic	.67	.69
huopo 活泼	Lively, Vivacious	.65	.65
reqing 热情	Warm, Enthusiastic	.61	.52
hequn 合群	Sociable, Gregarious	.59	.56
jiantan 健谈	Talkative, Brilliant in conversation	.59	.49
yangguang 阳光	Sunny	.55	.59
kuaile 快乐	Happy, Cheerful	.55	.59
gaoxing 高兴	Happy, Joyful	.53	.55
leguan 乐观	Optimistic	.50	.53
chenmoguayan 沉默寡言	Taciturn, Withdrawn	– .66	– .58
neixiang 内向	Introverted	– .63	– .60
gudan 孤单	Lonely, Alone	– .59	– .60
chenmen 沉闷	Not outgoing, Withdrawn, Dull	– .59	– .59
gudu 孤独	Lonely, Solitary	– .55	– .62
yiyu 抑郁	Depressed	– .54	– .56
youyu 忧郁	Melancholy, Sombre	– .53	– .57
yumen 郁闷	Gloomy, Depressed	– .53	– .53
beiguan 悲观	Pessimistic	– .53	– .53
Conscientiousness/Diligence			
yisibugou 一丝不苟	Unfailing precision, Meticulous	.58	.53
qinfen 勤奋	Diligent, Assiduous	.57	.61
yange 严格	Strict, Rigorous	.53	.52
qinkuai 勤快	Hard-working, Diligent	.53	.44
keku 刻苦	Assiduous, Hardworking	.50	.63
yanjin 严谨	Strict, Rigorous, Meticulous	.50	.42
renzhen 认真	Earnest, Conscientious	.49	.50
nuli 努力	Making an effort, Hard-working	.47	.63

(Continued)

Table 1 (Cont.)

Chinese	English	Peer	Self
qinlao 勤劳	Diligent, Industrious	.44	.54
qieerbushe 锲而不舍	Persevering	.44	.52
jinqu 进取	Enterprising	.41	.50
jianrenbuba 坚忍不拔	Firm, Indomitable	.41	.49
lansan 懒散	Lazy and sluggish	– .62	– .61
lanyangyang 懒洋洋	Languid, Listless	– .61	– .59
santiandayu, liangtianshaiwang 三天打鱼，两天晒网	Lack perseverance	– .61	– .58
landuo 懒惰	Lazy	– .60	– .67
sanman 散漫	Undisciplined, Slack	– .60	– .59
lanchong 懒虫	Lazy bones	– .56	– .58
cuxin 粗心	Careless	– .51	– .46
mahu 马虎	Careless, Sloppy	– .50	– .50
Unselfishness			
chunpu 淳朴	Simple, Honest, Unsophisticated	.64	.24
chunpu 纯朴	Simple, Honest, Unsophisticated	.63	.18
hanhou 憨厚	Simple, Honest, Kind	.59	.38
kuanhou 宽厚	Tolerant, Generous, Kind	.57	.38
dunhou 敦厚	Honest and sincere	.55	.23
kuanrong 宽容	Tolerant, Lenient, Forgiving	.54	.25
kuanhongdaliang 宽宏大量	Magnanimous, Broad-minded	.51	.40
zhonghou 忠厚	Honest and Innocent	.51	.33
houdao 厚道	Honest and Kind, Magnanimous	.50	.33
renhou 仁厚	Benevolent and Generous	.48	.35
jianyiyongwei 见义勇为	Act bravely for a just cause	.41	.42
yingyong 英勇	Heroic, Valiant	.28	.51
zisi 自私	Selfish, Self-centered, Egoistic	– .60	– .40
aoman 傲慢	Arrogant, Haughty, Overbearing	– .56	– .22
yaomianzi 要面子	Be sensitive about one's reputation	– .52	– .42

(Continued)

Table 1 (Cont.)

Chinese	English	Peer	Self
xiaoqi 小气	Stingy, Narrow-minded	– .49	– .40
aimianzi 爱面子	Be sensitive about one's reputation, Vain	– .48	– .37
jichou 记仇	Bearing/harboring/nursing Grudges	– .46	– .36
xurong 虚荣	Vain	– .45	– .44
tanxin 贪心	Greedy, Avaricious	– .45	– .41
Noxious Violativeness (NV; Harmfulness)			
qinshou 禽兽	Beastly, Inhuman persons	.64	.66
xiajian 下贱	Degrading, Low	.62	.70
yindang 淫荡	Lascivious, Lewd	.61	.63
xiongcan 凶残	Fierce and cruel, Bloodthirsty	.60	.66
xiee 邪恶	Evil, Sinister, Wicked, Vicious	.58	.68
langxingoufei 狼心狗肺	Brutal and cold-blooded, Heartless and ungrateful	.58	.67
beijian 卑贱	Low	.58	.65
xialiu 下流	Dirty, Indecent, Lewd, Obscene	.57	.67
huaidan 坏蛋	Scoundrel, Bastard, Rascal	.56	.65
fubai 腐败	Corrupted, Rotten	.56	.61
quede 缺德	Wicked, Unscrupulous, Unprincipled	.54	.74
xionghen 凶狠	Ferocious, Cruel, Fiendish, Cutthroat	.53	.57
jian 贱	Lowly	.51	.62
beilie 卑劣	Despicable, Contemptible,	.48	.75
elie 恶劣	Abominable, Foul, Disgusting	.46	.69
fangdang 放荡	Dissolute, Dissipated, Loose	.46	.61
yinxian 阴险	Insidious, Sinister, Treacherous	.44	.72
kechi 可耻	Shameful, Disgraceful, Ignominious	.43	.75
taoyan 讨厌	Disgusting, Hateful, Repulsive	.39	.69

(Continued)

Table 1 (Cont.)

Chinese	English	Peer	Self
wulai 无赖	Rascal, Scoundrel, Rogue	.39	.69
Gentle/Even Temper			
siwen 斯文	Gentle, Refined	.50	.37
wenjing 文静	Gentle and Quiet	.37	.46
guaiqiao 乖巧	Cute, Lovely	.33	.50
wenrou 温柔	Gentle and soft	.44	.29
guai 乖	Well-behaved, Obedient	.29	.43
tinghua 听话	Docile, behaving as told, Obedient	.29	.43
hanxu 含蓄	Reserved	.35	.36
xini 细腻	Sensitive, Exquisite	.42	.28
anfen 安分	Knowing one's place, Not going beyond one's bounds	.27	.41
chongdong 冲动	Impulsive	– .61	– .31
jixingzi 急性子	Quick-tempered, Short-tempered	– .59	– .43
xingji 性急	Short-Tempered, Impatient, Quick-tempered	– .53	– .38
jizao 急躁	Impatient, Irritable, Hot-tempered	– .53	– .28
baozao 暴躁	Irascible, Irritable, Fiery	– .51	– .28
juejiang 倔强	Stubborn, Unbending, Tenacious	– .42	– .30
jidong 激动	Excited	– .42	– .29
panni 叛逆	Rebellious, Treasonable	– .40	– .39
dadan 大胆	Daring, Bold	– .35	– .50
qiangying 强硬	Tough, Unyielding	– .35	– .43
yonggan 勇敢	Brave, Courageous, Intrepid	– .30	– .42
Intellect/Positive Valence			
dute 独特	Unique, Distinctive	.66	.42
tebie 特别	Special, Unusual, Peculiar	.57	.38
duocaiduoyi 多才多艺	Versatile	.51	.40
tiancai 天才	Genius, Talent	.49	.56
wanmei 完美	Perfect	.48	.53
jiechu 杰出	Outstanding, Prominent	.44	.50
bang 棒	Great, Fantastic	.44	.49
gaogui 高贵	Noble, Morally elevated	.39	.59

(Continued)

Table 1 (Cont.)

Chinese	English	Peer	Self
congming 聪明	Bright, Intelligent, Clever	.38	.42
xinggan 性感	Sexy	.38	.40
ruizhi 睿智	Wise and prescient/far-sighted	.36	.49
chuse 出色	Outstanding, Remarkable, Splendid	.36	.45
jiling 机灵	Smart, Clever	.36	.44
rencai 人才	Talent, Talented person	.35	.45
jizhi 机智	Resourceful, Quick-witted, Tactful	.35	.45
decaijianbei 德才兼备	Having both ability and integrity	.35	.40
putong 普通	Ordinary, Average, Common	– .55	– .55
pingfan 平凡	Ordinary, Common, Mean	– .50	– .49
pingyong 平庸	Mediocre	– .47	– .49
muguangduanqian 目光短浅	Short-sighted	– .37	– .44
Dependency/Fragility			
jiaodidi 娇滴滴	Delicate, Fragile	.52	.23
jiaoqi 娇气	Squeamish, Fragile, Delicate	.50	.31
tianzhen 天真	Naïve, Innocent, Simple-minded, Unsophisticated	.48	.25
haiziqi 孩子气	Childish	.45	.44
youzhi 幼稚	Childish, Naïve	.40	.49
cuiruo 脆弱	Fragile, Weak	.35	.40
diaoman 刁蛮	Headstrong	.34	.31
renxing 任性	Self-willed, Headstrong	.33	.26
ganqingyongshi 感情用事	Sentimental, Be swayed by one's feelings/sentiments, Emotional	.27	.42
nanzihan 男子汉	Manly, True man	– .52	– .21
laolian 老练	Seasoned, Experienced	– .39	– .42
duli 独立	Independent	– .39	– .26
chengshu 成熟	Mature	– .34	– .44
shenchen 深沉	Deep, Impenetrable	– .34	– .33
chenwen 沉稳	Steady, Prudent, Tranquil	– .31	– .59

(Continued)

Table 1 (Cont.)

Chinese	English	Peer	Self
chenzhuo 沉着	Composed, Calm	– .29	– .55
lengjing 冷静	Calm, Sober, Cool, Imperturbable	– .27	– .70
dongshi 懂事	Sensible/intelligent	– .22	– .37
zhending 镇定	Calm and collected, Composed	– .20	– .47
xianshi 现实	Practical, Realistic	– .19	– .44

Note. Peer sample, $N = 485$, self sample, $N = 417$. For each factor, the 20 terms with highest average factor loading across the two samples are shown, in order by their loadings in the (larger) peer sample. Self-rating loadings for the Dependency/Fragility factor and the Negative Valence factor as well as peer-rating loadings for the Dependency/Fragility factor and the Intellect/Positive Valence factor have been reflected.

For evaluating replicability, a key proportion is the signal:noise ratio discernible by comparing the mean convergent correlation with the mean divergent correlation (calculated using absolute values). If replication is strong, this proportion (regardless of the level of internal consistency in the marker scales) will be high; if weak, this proportion will be low.

Table 2 indicates that the Big Two replicates to a greater degree than do other models. For the Big Two, the mean convergent correlations are substantially higher than for all other models, while at the same time the mean divergent correlations are relatively lower. This superior replication was evident in ipsatized as well as original data. The Multi-Language Seven factor structure also had a noticeably higher overall level of replication than the five- or six-factor models, mainly because the divergent correlations were lower.

Early lexical studies in German and English provided replications of the Big Five in which marker scales correlated with matched emic factors above .65 on average (Ostendorf, 1990; Saucier, 1997), and this threshold has been used as a replication threshold in previous lexical studies (Saucier, 2003b; Saucier et al., 2005).

The mean correlations for each row indicate good replication (mean r between matched factors exceeding .65) for the one-factor structure and the Big Two and for the internalizing-externalizing factors in one of the two data sets. If we relied on original-data analyses instead, the conclusions would be the same. With the 362

Table 2
Coefficients Indicating Replication of Etic Factor Structures in Emic Chinese Factors

Etic Structure	Replication *r* Values	Mean Convergent	Mean (and Highest) Divergent	Replication *r* Values	Mean Convergent	Mean (and Highest) Divergent
			All 413 Terms			
		Peer-Ipsatized			Self-Ipsatized	
1	74	74*	—	85	85*	—
2-B2	83 78	81*	16 (19)	85 84	85*	18 (22)
2-IE	67 66	67*	07 (13)	74 63	69*	18 (26)
3	73 73 68	72*	16	80 77 76	77*	15
5	75 73 62 20 08	48	21 (64)	81 71 68 62 44	65*	20 (49)
6	84 72 69 61 34 19	57	17 (61)	81 80 70 58 45 34	61	18 (50)
7	79 79 72 64 56 51 32	62	15 (59)	80 73 72 70 57 37 26	59	18 (75)
E5	75 72 56 25 19	50	25 (60)	68 60 56 43 9	47	23 (55)
E7	89 80 71 61 55 28 23	58	15 (43)	74 64 51 48 45 39 8	47	16 (42)
		Peer-Original			Self-Original	
1	86	86*	—	89	89*	—
2-B2	80 78	79*	23 (37)	83 78	80*	19 (33)
2-IE	74 72	73*	16	76 39	58	24 (32)
3	73 57 48	59	24	64 57 34	53	32
5	68 60 56 47 36	53	25 (59)	70 64 59 49 25	54	26 (54)

(Continued)

Table 2 (Cont.)

Etic Structure	Replication *r* Values	Mean Convergent	Mean (and Highest) Divergent	Replication *r* Values	Mean Convergent	Mean (and Highest) Divergent
6	75 70 55 43 35 32	52	21 (59)	68 63 62 57 55 42	58	20 (67)
7	77 68 64 59 56 53 31	58	18 (63)	73 71 66 66 62 57 22	59	18 (56)
E5	80 72 63 58 27	61	25 (59)	71 71 68 54 29	59	26 (64)
E7	90 70 57 49 48 46 22	55	20 (56)	80 73 72 61 56 39 36	60	19 (51)
			Reduced Set of 362 Terms			
		Peer-Ipsatized			Self-Ipsatized	
1	76	76*	—	84	84*	—
2-B2	81 77	79*	14 (23)	83 81	80*	19 (25)
2-IE	74 64	69*	15 (27)	73 60	66*	21 (26)
3	73 71 66	70*	17 (46)	60 50 45	51	20 (43)
5	76 72 49 25 21	49	20 (66)	76 64 49 45 18	51	24 (65)
6	78 71 70 58 29 18	54	18 (56)	76 73 72 47 43 33	57	19 (49)
7	78 77 71 65 53 38 32	59	16 (58)	71 71 69 60 45 35 33	55	19 (67)
E5	76 74 66 54 35	61	20 (62)	81 72 58 28 22	52	23 (56)
E7	89 81 73 62 51 48 11	59	15 (42)	81 78 71 61 54 40 17	57	19 (47)

Note. B2: Big Two. IE: Internalizing-Externalizing factor duo.
*Mean replication coefficients (mean of correlations between best-matched paris of emic-etic factors) that exceeded the .65 threshold.

terms (ipsatized data), all these one- and two-factor structures were also found to be well replicated.

Beyond the two-factor level, no other structure consistently surpassed the mean .65 threshold. The Big Three passed it in self- and peer-ipsatized data for the full set of terms and for the peer-ipsatized data for 362 terms, but not otherwise. The Big Five actually reached this threshold (.65) in one instance (self-ratings, ipsatized data, full set of terms) but fell far short of it (.48) when the same analysis was repeated in peer ratings (again, ipsatized data, full set of terms), due to inconsistencies among the five-factor structures across variations in data. Overall, although structures with more factors should generally be harder to replicate, among the five- to seven-factor etic structures, the seven-factor structure actually came out slightly ahead of the others. Consistent with the greater stability of the seven-factor structures across analytic variations, the mean replication coefficients for the ML7 were the most consistent in magnitude, ranging from .55 to .62. The Cross-Language Six had mean coefficients ranging from .52 to .61 and the Big Five from .48 to .65.

The correlations of emic factors with the longer marker scales are presented in Table 3. If an etic structure were well replicated, we would see the correlations generating neat pairings of one emic factor and one etic factor scale. The table portrays a more complicated situation. In each comparison two of the emic factors had no correlation higher than .43 with any etic factor scale.

What might be limiting the replicability of etic structures? Some of the Chinese emic factors tended to combine multiple etic factors, whereas others tended to capture variance beyond the etic factors. This problem cannot be substantially solved by rerotation, though it might help to extract additional factors. For the CL6, the problem was least evident in the ipsatized peer data, where etic scales matched up not too badly—.89 for Extraversion, .81 for Honesty, .74 for Conscientiousness, .65 for Agreeableness, .63 for Openness, and .39 for Emotionality—with six of the factors in the seven-factor solution (all 413 terms), the unmatched emic factor being Negative Valence. For the Big Five, the problem was least evident in ipsatized self data, where etic scales matched up reasonably well—.73 for Extraversion, .72 for Agreeableness, .68 for Intellect, .61 for Conscientiousness, and .56 for Emotional Stability—with five of the factors in the seven-factor solution (all 413 terms), the unmatched emic factors being even temper and unselfishness (Negative Valence having matched up more

Table 3
Coefficients Indicating Replication of Five- and Six-Factor Etic Structures in Emic Chinese Factors, Using Longer Imported Scales

Etic Factor Scale	Emic Factors					
	Five Factors From Peer Reports					
	1	**2**	**3**	**4**	**5**	
Extraversion	−.14	.84*	−.05	−.07	−.05	
Agreeableness	.66*	.38	.17	−.25	−.17	
Conscientiousness	.30	.11	.76*	−.15	−.11	
Emotional Stability	.56*	.36	.50*	.05	−.11	
Intellect	−.11	.43	.48	−.25	−.29	
	Five Factors From Self-Reports					
	1	**2**	**3**	**4**	**5**	
Extraversion	−.10	.76*	−.08	.43	.10	
Agreeableness	.73*	.45	.10	−.17	.17	
Conscientiousness	.34	.26	.76*	.01	−.02	
Emotional Stability	.30	.50*	.50*	−.28	.19	
Intellect	.15	.45	.41	.18	−.30	
	Six Factors From Peer Reports					
	1	**2**	**3**	**4**	**5**	**6**
Honesty/Humility	.77*	.17	.16	−.24	−.08	−.02
Emotionality	−.26	−.50*	−.54*	−.01	.20	.34
Extraversion	.01	−.01	.88*	−.06	−.18	−.10
Agreeableness	.72*	.40	−.02	.00	−.22	.29
Conscientiousness	.37	.78*	.05	−.13	−.04	.23
Openness	.00	.38	.30	−.12	−.59*	−.14
	Six Factors From Self-Reports					
	1	**2**	**3**	**4**	**5**	**6**
Honesty/Humility	.76*	.20	.24	.33	.10	.14
Emotionality	−.22	−.58*	−.37	−.15	−.34	−.34
Extraversion	.09	.88*	−.02	.00	−.10	.15
Agreeableness	.57*	.05	.20	.51*	.41	−.18
Conscientiousness	.36	.12	.71*	.16	.39	−.03
Openness	.23	.53*	.16	−.28	.43	.00

Note. Peer sample, $N = 485$, self sample, $N = 417$. All analyses based on ipsatized data, reduced set of terms.
*Correlation having a magnitude of at least .50.

highly with Agreeableness). The replication of the Big Five in peer data and the CL6 in self data was considerably weaker. And none of the matrices in Table 3 indicate a mean replication coefficient of over .60,

although these mean replication levels were a few points higher with the longer (Table 3) than with the shorter (Table 2) marker scales. We note that increasing marker-scale reliability from a moderate to a high level enabled only very modest increases in replication coefficients, indicating relatively little problem with the use of brief marker scales. Overall, these comparisons indicate that neither the Big Five nor the Cross-Language Six is sufficiently comprehensive to capture all the major factors in Chinese personality description, although each etic model captures some of these factors rather well, and the CL6 seems to capture more than the Big Five does.

On the other hand, one- and two-factor structures found in previous lexical studies replicate well in Chinese, in a simple and direct manner. The internalizing-and-externalizing factor duo and the Big Three had a more mixed performance, but were still better replicated across analytic variations than any a priori structure having more than three factors. A table of salient terms in the emic Chinese two-factor solutions (peer and self-ratings, ipsatized data), which provide a reasonable representation of the etic Big Two, are available from the first or second author.

Comparison of the Present Emic With Previous Emic Structures for Chinese

For the five-factor emic structure of Yang and Bond (1995), the most useful comparison is with the etic Big Five. Across the six data variations in Table 2, our emic five-factor structure replicated the Yang and Bond structure with a mean replication correlation of .56, compared to .53 for the Big Five; the only data variation in which the Big Five came out ahead was ipsatized self-ratings. Comparing the seven-factor emic structure of Wang et al. (Cui & Wang, 2003) to the Multi-Language Seven, the mean replication correlations were .56 for the Wang et al. structure and .58 for the ML7; however, the only data variation in which the ML7 had a particularly strong advantage was ipsatized self-ratings. The particular strength of the etic structures in ipsatized self-ratings may occur because these structures were primarily derived in this type of data. In contrast, previous Chinese-emic structures were apparently based on nonipsatized data; some aspects of these structures may be confounded with individual differences in scale usage.

Thus, replicability for previous Chinese-emic structures failed to exceed that for etic structures, despite the generally higher reliability of their emic marker scales. One might conclude that a Chinese emic structure derived by standard lexical-study procedures (as ours was) replicates about equally well (a) structures found in other languages using similar procedures and (b) structures found in Chinese using different procedures. Personality dimensions identified in lexical studies appear to be affected by both the specific language and the procedures used. Our evidence was thus partly consistent with cultural relativism, but partly with the view that there are cross-cultural universals whose discovery depends on application of standard procedures.

DISCUSSION

Lexical studies are important because they address the question of "which personality attributes are important enough to measure" in a relatively objective and empirical way. Moreover, they make possible the derivation of an empirically based structural model for personality attributes that, because they are based on studies in a wide array of languages, is relatively unbiased culturally. And the most useful lexical studies are those that (a) rely on reports by informants as well as self-reports and (b) test the replicability of a wide range of candidate models. The current lexical study of the Chinese language meets both these requirements.

Examining convergent structural tendencies across methodological variations—use of self- or peer-rating samples, original or ipsatized data, and a relatively inclusive or narrow variable selection—we were able to identify in these data the outlines of an emic Chinese structure for personality attributes. The seven factors in this structure are Extraversion, Conscientiousness/Diligence, NV (Negative Valence or Noxious Violativeness), Unselfishness, Emotional Volatility, Intellect/Positive Valence, and Dependency/Fragility. Some features of this structure (Extraversion, Conscientiousness, and perhaps Emotional Volatility) are evident in the Chinese structure promulgated by Wang and colleagues (e.g., Cui & Wang, 2003), but others (including the differentiation of NV and Unselfishness and the appearance of Intellect and Dependency factors) are not. The present structure, however, seems more comparable to those found in previous lexical studies. The Dependency/Fragility factor may be an

interesting culture-specific variant of a more cross-culturally ubiquitous dimension (Resiliency vs. Emotionality); cultures that more highly value interdependency, as East Asian cultures do, have a more tolerant and less dismissive attitude toward dependency than Western cultures have, and dispositions involving dependency may have enhanced meaning and importance within the culture.

Because of relative stability in the two-factor structures, we can also identify an emic structure at the broader two-factor level with Social Propriety and Dynamism factors. However, this two-factor structure corresponds closely with an etic "Big Two" that has now been found in many languages. The emic seven-factor structure we found corresponds loosely to the Multi-Language Seven structure we tested, but replication coefficients for this etic structure fell below the replication-threshold convention (mean *r* of .65 between matched factors) that we continued from prior studies (Ostendorf, 1990; Saucier, 1997, 2003b; Saucier et al., 2005). Replication of the etic five- and six-factor models also fell below that threshold.

Is the .65 threshold arbitrary? This threshold references Ostendorf's (1990) replication of the Big Five in German, as well as Saucier's (1997) replication of the Big Five in an English-language lexical study: Both these had mean replication *r* values for the Big Five above .65. Therefore, for a structure surpassing this threshold it can be claimed that the replication is at least as good as that for the Big Five in the early English and German lexical studies. Perhaps one could claim that a mean *r* in the .55 to .65 range—the range in which fell the overall replication level for all of the etic structures we tested—is a marginally good replication, especially if the replication coefficients for all of the imported etic factors are above .40. However, for structures of more than three factors, exceeding .40 for replications of all matched factors rarely occurred in these analyses. It does now seem that structures of more than three factors are difficult to strongly replicate in emic lexical studies, particularly those in which the language is much different than those in which the structure was originally found or in which the methodology is very much different. However, even these "loose" or "marginally good" replications, when aggregated across many studies and languages, might reveal a structural pattern on which there are substantial emic variations from one language to another. Truly strong and obvious recurrences across lexical studies of very diverse provenance appear to occur only at the one- and two-factor levels (Saucier & Goldberg,

2001, 2006), a finding that underscores the importance of paying due attention to the top levels of the personality-attribute hierarchy. But these most robust top levels do not deserve exclusive attention: Structures with more than two factors provide more information, an advantage that may compensate for their lesser robustness.

Variation in the structure of personality-attribute concepts across languages may have diverse sources. Part of the variation may be study-specific error; that is, it may be due to chance characteristics of the sample and not systematic characteristics of the language community. Part of the variation may be truly culture-specific features, which in turn may be due to effect of key cultural schemas or of biological differences between populations. It seems likely that the more independent sources of variation (i.e., the more factors) allowed, the more chance there is for error or culture-specific features to affect the factors. This predicts increasing "noise" in attempts to replicate imported etic structures in local data, as one increases the number of factors. Such an account can make sense of why one- and two-factor structures are more robust in lexical studies. It does not account, however, for why in the present study a seven-factor structure actually, if anything, exceeded a six-factor structure in replicability, which six-factor structure in turn, if anything, exceeded the replicability for the five-factor structure.

There are several reasons to be cautious, however, in concluding on the evidence from this study that the Multi-Language Seven is the superior model. First, the replication coefficients in this study are below the conventional threshold we set out—a mean r of .65. Second, the ML7 model was derived from scrutiny of the correspondence between results of studies in two languages (Filipino and Hebrew) with some evidence that structures in English and Italian have shown in part some of the same features. This is at most four languages (five if we include Chinese) that have supported the model, which is less than the number that can be adduced to support the Big Five or the Cross-Language Six. Third, if in this study of Chinese descriptors we take the Unselfishness factor as a replica of ML7 Concern for Others versus Egotism and the NV factor as a replica of ML7 Negative Valence, we should acknowledge that Unselfishness and NV might be combined into a single factor, as indeed they were in some self-rating analyses. An Unselfishness versus NV factor corresponds rather well to Honesty combined with Altruism (cf. Lee & Ashton, 2006), and if this combination is made we may have a

structure differing little from the Cross-Language Six (Saucier, 2007). And fourth, more generally, the ML7 structure corresponds rather well to the Cross-Language Six, although with Honesty-factor variance tending to be split into Concern for Others and NV components (Saucier, 2008). Therefore the present support for the ML7 can also be read as support for a variant of the CL6 structure in which Honesty components split into two factors.

In our analyses, structure differed relatively little according to whether we used an inclusive or a narrow variable selection. This may seem surprising given previous empirical indications that more inclusive selections lead to additional factors (Saucier, 1997; Tellegen & Waller, 1987). However, lexical studies with inclusive variable selections have tended to find the Cross-Language Six factors (to a degree greater than they find Big Five factors) with the nature of any additional factors being more a matter of disagreement between studies (Saucier, 2007). Moreover, although Chinese terms have strong doses of evaluation, the language seems to have fewer frequently used terms like the English terms Excellent, Impressive, Awful, and Evil: strongly evaluative but also extremely ambiguous in their descriptive referents. For example, most of the terms showing the highest loadings on NV in Table 1 have definitive descriptive implications regarding specific forms of behavior considered immoral, such as cruelty, licentiousness, and disrespect to family and to those who have helped one. The two terms in that table translated as "outstanding" in fact both strongly imply the descriptive attribute "unusually competent." This helps explain why variable selection had only small effects in this study. It also suggests that trimming highly evaluative terms out of personality studies may in effect amputate psychologically important content domains of morality and competence (cf. Wojciszke, 1994).

Some previous studies with inclusive variable selections have found an attractiveness factor. We did not. We had at most five terms referencing attractiveness, probably too few to form a factor. As an aggregate (summing ratings on these into a single score) they were correlated with the Dynamism factor (about .35 in either sample) and with Intellect/Positive Valence (over .40 in either sample). Given previous studies (e.g., Saucier, 1997), these relations are unsurprising.

A strength of the present study is its use of both self- and peer-rating samples. One caveat, however, is that these data cannot be

used to compare in a systematic way differences between self- and peer reports. Our peer-rating sample was recruited, in part, from a different province in China than was the self-rating sample. Therefore, in comparing these samples one is probably comparing a mixture of two different things: self-ratings versus peer ratings and characteristics of student populations in two parts of China.

Future studies could usefully compare the present emic Chinese structure with structures derived by variant methods, including the seven-factor measurement model of Wang and associates (e.g., Cui & Wang, 2003), the six-factor measurement model developed by Cheung and associates (e.g., Cheung, Kwong, & Zhang, 2003), and the ancient structure of attributes from B. Yang (2005). There is apparent consistency of these studies with the present study in finding dimensions related to unselfishness, extraversion, and competence, although other dimensions seem relatively disparate from one study to another. Future studies of Chinese descriptors should also take into account recent syntheses of lexical-study results (e.g., Ashton, Lee, Perugini, et al., 2004; Saucier, 2007; Saucier & Goldberg, 2001, 2006) that indicate relatively cross-culturally robust dimensions of personality, dimensions which in turn afford greater insight into what in the Chinese structure of attributes is culture specific and what is not.

Conclusions

Studies of the language of personality provide important indications as to what personality tests should measure. And when the language studied is that used by about one fifth of the world's population, the results have some inherent importance. Chinese personality descriptors reveal an "emic" structural pattern that could be an important structural model in its own right. Our studies are a contribution toward discovering this pattern. They indicate that the emic Chinese structural pattern corresponds to that found in other languages most especially at the one- and two-factor levels. But there is also some similarity, albeit one that is comparatively less exact, with recent structural models—the Cross-Language Six and Multi-Language Seven—that go beyond the Big Five. Such findings indicate the importance of testing structural models other than the Big Five if one is to find a personality classification that is optimally usable around the globe, across major boundaries of nation and culture.

REFERENCES

Almagor, M., Tellegen, A., & Waller, N. (1995). The Big Seven model: A cross-cultural replication and further exploration of the basic dimensions of natural language of trait descriptions. *Journal of Personality and Social Psychology*, **69**, 300–307.

Angleitner, A., Ostendorf, F., & John, O. P. (1990). Towards a taxonomy of personality descriptions in German: A psycho-lexical study. *European Journal of Personality*, **4**, 89–118.

Ashton, M. C., Lee, K., & Goldberg, L. R. (2004). A hierarchical analysis of 1,710 English personality-descriptive adjectives. *Journal of Personality and Social Psychology*, **87**, 707–721.

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., et al. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, **86**, 356–366.

Bailey, F. G. (1971). *Gifts and poison: The politics of reputation*. New York: Schocken.

Benet-Martìnez, V., & Waller, N. G. (1997). Further evidence for the cross-cultural generality of the Big Seven factor model: Indigenous and imported Spanish personality constructs. *Journal of Personality*, **65**, 567–598.

Boies, K., Lee, K., Ashton, M. C., Pascal, S., & Nicol, A. A. M. (2001). The structure of the French Personality Lexicon. *European Journal of Personality*, **15**, 277–295.

Caprara, G. V., Barbaranelli, C., & Zimbardo, P. G. (1997). Politicians' uniquely simple personalities. *Nature*, **385**, 493.

Cheung, F. M., Kwong, J. Y. Y., & Zhang, J. (2003). Clinical validation of the Chinese Personality Assessment Inventory (CPAI). *Psychological Assessment*, **15**, 89–100.

Chinese Academy of Social Sciences, Institute of Linguistics. (2005). *The contemporary Chinese dictionary*. Beijing: Commercial Press.

Church, A. T., Reyes, J. A. S., Katigbak, M. S., & Grimm, S. D. (1997). Filipino personality structure and the Big Five model: A lexical approach. *Journal of Personality*, **65**, 477–528.

Confucius (1983). *The analects (Lun yü)* (D. C. Lau, trans.). Hong Kong: Chinese University Press.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Cui, H., & Wang, D. (2003). The re-validation of Chinese personality structure and the results from adjectives evaluation. *Psychology and Behavioral Research*, **1**, 89–95 (in Chinese).

De Raad, B., Hendriks, A. A. J., & Hofstee, W. K. B. (1992). Towards a refined structure of personality traits. *European Journal of Personality*, **6**, 301–319.

Di Blas, L., & Forzi, M. (1998). An alternative taxonomic study of personality-descriptive adjectives in the Italian language. *European Journal of Personality*, **12**, 75–101.

Di Blas, L., & Forzi, M. (1999). Refining a descriptive structure of personality attributes in the Italian language: The abridged Big Three circumplex structure. *Journal of Personality and Social Psychology*, **76**, 451–481.

Dunlap, W. P., & Cornwell, J. M. (1994). Factor analysis of ipsative measure. *Multivariate Behavioral Research*, **29**, 115–126.

Funder, D. C. (2005). *The personality puzzle* (3rd ed.). New York: Norton.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 2, pp. 141–165). Beverly Hills, CA: Sage.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, **59**, 1216–1229.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, **4**, 26–42.

Goldberg, L. R., & Somer, O. (2000). The hierarchical structure of common Turkish person-descriptive adjectives. *European Journal of Personality*, **14**, 497–531.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, **103**, 265–275.

Hahn, D. W., Lee, K., & Ashton, M. C. (1999). A factor analysis of the most frequently used Korean personality trait adjectives. *European Journal of Personality*, **13**, 261–282.

Haviland, J. B. (1977). *Gossip, reputation, and knowledge in Zinacantan*. Chicago: University of Chicago Press.

Hofstee, W. K. B., Kiers, H. A., De Raad, B., Goldberg, L. R., & Ostendorf, F. (1997). A comparison of Big-Five structures of personality traits in Dutch, English, and German. *European Journal of Personality*, **11**, 15–31.

Hřebíčková, M., Ostendorf, F., & Angleitner, A. (1995, July). *Basic dimensions of personality description in the Czech language*. Paper presented at the 7th Meeting of the International Society for the Study of Individual Differences, Warsaw, Poland.

Hřebíčková, M., Ostendorf, F., Osecká, L., & Čermák, I. (1999). Taxonomy and structure of Czech personality-relevant verbs. In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 51–65). Tilburg, The Netherlands: Tilburg University Press.

Krueger, R. F., Chentsova-Dutton, Y. E., Markon, K. E., Goldberg, D., & Ormel, J. (2003). A cross-cultural study of the structure of comorbidity among common psychopathological syndromes in the general health care setting. *Journal of Abnormal Psychology*, **112**, 437–447.

Krueger, R. F., & Tackett, J. L. (2003). Personality and psychopathology: Working toward the bigger picture. *Journal of Personality Disorders*, **17**, 109–128.

Lee, K., & Ashton, M. C. (2006). Further assessment of the HEXACO Personality Inventory: Two new facet scales and an observer report form. *Psychological Assessment*, **18**, 182–191.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, **4**, 84–99.

Mlacic, B., & Ostendorf, F. (2005). Taxonomy and structure of Croatian personality-descriptive adjectives. *European Journal of Personality*, **19**, 117–152.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, **66**, 574–583.

Osgood, C. E., May, W. H., & Miron, M. S. (1975). *Cross-cultural universals in affective meaning*. Urbana: University of Illinois Press.

Ostendorf, F. (1990). *Sprache und Personlichkeitsstruktur: Zur Validitat des Funf-Faktoren-Modells der Personlichkeit [Language and personality structure: Towards the validity of the Five-Factor Model of Personality]*. Regensburg, Germany: Roderer.

Peabody, D. (1987). Selecting representative trait adjectives. *Journal of Personality and Social Psychology*, **52**, 59–71.

Saucier, G. (1997). Effect of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology*, **73**, 1296–1312.

Saucier, G. (2003a). An alternative multi-language structure of personality attributes. *European Journal of Personality*, **17**, 179–205.

Saucier, G. (2003b). Factor structure of English-language personality type-nouns. *Journal of Personality and Social Psychology*, **85**, 695–708.

Saucier, G. (2007). *Personality dimensions recurrent across languages: Support for a Big Six model*. Manuscript submitted for publication.

Saucier, G. (2008). Measures of the personality factors found recurrently in human lexicons. In G. J. Boyle, G. Matthews, & D. Saklofske (Eds.), *The sage handbook of personality theory and assessment: Personality measurement and testing*. (Vol. 2, pp. 29–54) London: Sage.

Saucier, G., Georgiades, S., Tsaousis, I., & Goldberg, L. R. (2005). The factor structure of Greek personality adjectives. *Journal of Personality and Social Psychology*, **88**, 856–875.

Saucier, G., & Goldberg, L. R. (1996). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality*, **10**, 61–77.

Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, **69**, 847–879.

Saucier, G., & Goldberg, L. R. (2006). Personality, character, and temperament: The cross-language structure of traits. *Psychologie Française*, **51**, 265–284.

Shweder, R. A. (1972). *Semantic structure and personality assessment*. Unpublished doctoral dissertation, Harvard University.

Somer, O., & Goldberg, L. R. (1999). The structure of Turkish trait-descriptive adjectives. *Journal of Personality and Social Psychology*, **76**, 431–445.

Szarota, P. (1996). Taxonomy of the Polish personality-descriptive adjectives of the highest frequency of use. *Polish Psychological Bulletin*, **27**, 343–351.

Szirmak, Z., & De Raad, B. (1994). Taxonomy and structure of Hungarian personality traits. *European Journal of Personality*, **8**, 95–118.

Tellegen, A., & Waller, N. G. (1987). *Re-examining basic dimensions of natural language trait descriptors*. Paper presented at the 95th annual convention of the American Psychological Association, New York.

ten Berge, J. M. F. (1999). A legitimate case of component analysis of ipsative measures, and partialling the mean as an alternative to ipsatization. *Multivariate Behavioral Research*, **34**, 89–102.

Waller, N. G., & Zavala, J. D. (1993). Evaluating the Big Five. *Psychological Inquiry*, **4**, 131–134.

Wang, D. F., Cui, H., & Zhou, F. (2005). Measuring the personality of Chinese: QZPS versus NEO-PI-R. *Asian Journal of Social Psychology*, **8**, 97–122.

Wang, D. F., Fang, L., & Zuo, Y. T. (1995). A psycho-lexical study on Chinese personality from natural language. *Psychological Science (China)*, **27**, 400–406 (in Chinese).

Wasti, S. A., Lee, K., Ashton, M. C., & Somer, O. (2008). Six Turkish personality factors and the HEXACO model of personality structure. *Journal of Cross-Cultural Psychology*, **39**, 665–684.

White, G. M. (1980). Conceptual universals in interpersonal language. *American Anthropologist*, **82**, 759–781.

Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology*, **67**, 222–232.

Yang, B. (2005). A factor analysis of the ancient Chinese personality structure. *Psychological Science (China)*, **28**, 668–672 (in Chinese).

Yang, K., & Bond, M. H. (1990). Exploring implicit personality theories with indigenous or imported constructs: The Chinese case. *Journal of Personality and Social Psychology*, **58**, 1087–1095.