
Advances in Personality Psychology

Volume One

Edited by
Sarah E. Hampson



(2000).
East Sussex, England: Psychology Press.

Chapter 1

Cross-language studies of lexical personality factors

Gerard Saucier, Sarah E. Hampson and Lewis R. Goldberg

The rationale for lexical studies rests on the assumption that the most meaningful personality attributes tend to become encoded in language as single-word descriptors. Based on this rationale, a number of studies have been conducted examining the factor structure of adjectival descriptors extracted from dictionaries. Using as an initial working hypothesis an Anglo/German version of the Big Five factor structure, we review lexical studies in English and 12 other languages, and examine their fit with the Big Five. We find that the inclusion of highly evaluative terms and physical descriptors may lead to factors beyond the Big Five. With more conventional variable selections emphasizing disposition descriptors, indigenous structures resembling the Big Five seem to emerge most readily in northern European languages. The Big Five have been only partially reproduced in many other studies, although this might reflect the differences in their procedures for selecting variables. Factor structures with fewer factors (one, two or three) may be more generalizable cross-culturally, perhaps because they are less affected by these methodological variations. At this point, more attention should be given to the influence of the major types of procedural differences on the resulting factor structures. Such variations in research design and analysis may account for much of what otherwise might be prematurely interpreted as cultural differences. Suggestions for improving the lexical-study paradigm are offered.

INTRODUCTION

After many years of dispute, personality psychologists have recently shown some signs of agreement on a framework for classifying and organizing personality traits. A consensus among many in the field, which was evident by the early 1990s (e.g. McCrae and John, 1992), posits that five broad dimensions of personality capture the most important aspects of lexicalized personality traits. Table 1.1 indicates subcomponents associated with each of these 'Big Five' factors from a recent study of English and German personality descriptors (Saucier and Ostendorf, 1999).¹

In an influential critique, Block (1995) cautioned against settling prematurely on the five-factor approach. However, even those most closely allied with the five-factor framework regard it as a working hypothesis, rather than a final solution (e.g. De Raad, 1998; Goldberg and Saucier, 1995).

Table 1.1 Big Five subcomponents found in English and German

Domain	Subcomponent	Example terms in English	
		Positive pole	Negative Pole
Extraversion (I)	Assertiveness	assertive	weak
	Activity-adventurousness	daring	unadventurous
	Unrestraint	talkative	shy
	Sociability	cheerful	unsociable
Agreeableness (II)	Warmth-affection	affectionate	cold
	Generosity	generous	selfish
	Gentleness	agreeable	harsh
	Modesty-humility	modest	egotistical
Conscientiousness (III)	Orderliness	organized	sloppy
	Reliability	responsible	undependable
	Industriousness	ambitious	lazy
	Decisiveness	decisive	inconsistent
Emotional Stability (IV)	Insecurity ^a	relaxed	insecure
	Emotionality ^a	unemotional	excitable
	Irritability ^a	undemanding	irritable
Intellect (V)	Intellect	intelligent	unreflective
	Imagination	creative	unimaginative
	Perceptiveness	perceptive	shortsighted

Source: Saucier and Ostendorf (1999).

Note

^a These subcomponents are labelled by their negative rather than positive pole.

Like any scientific model, its prime function is to stimulate research and either to be proved wrong or, more likely, to be refined in the light of new evidence. Additional dimensions could emerge in studies of different cultures or studies of different forms of personality language (e.g. type-nouns), or different kinds of personality data. Moreover, findings from any of these sources could point to a better structure at a different hierarchical level, one with more factors or fewer factors.

Initial problems in developing a scientifically compelling descriptive classification include (a) defining the universe of personality attributes that will be classified and (b) determining the particular attributes from that universe that are important enough to be represented in the final classification. The lexical approach to personality solves these problems by a major and far-reaching hypothesis: the most important individual differences eventually become encoded as single words in the natural (i.e. non-scientific) language (Allport and Odbert, 1936; Cattell, 1943; Goldberg, 1981; Saucier and Goldberg, 1996b). Therefore, at least for the major modern languages, their dictionaries contain a reasonably comprehensive

set of those individual differences that past speakers of those languages have needed to communicate. This 'lexical hypothesis' has two reasonable corollaries: (a) frequency of use has a rough correspondence with importance, and (b) the number of words referring to a particular personality attribute will be a rough guide to the importance of that attribute for the speakers of the language² (Saucier and Goldberg, 1996b). These two conceptually independent criteria can both be satisfied if one examines the most frequently used subset of relevant terms in the language, and then searches for structure (i.e. factors, clusters) in that subset.

Although the peculiarities of trait structure derived from any particular language could be of considerable interest, cross-language regularities have even more scientific import. In the search for such regularities, the most important personality dimensions will be those that replicate across samples of participants, targets of descriptions, variations in analytic procedures, and across languages (Saucier and Goldberg, 1996b). If a factor structure can be recovered in a wide variety of the world's languages, then this would provide support for the possibility that those dimensions might represent cultural 'universals' of person description (Goldberg, 1981). This chapter reviews the evidence for the reproducibility of structural representations – especially the Big Five – across those languages that have been studied to date. Some of the major studies in English are summarized first to describe how the five-factor framework came to be established. Next, lexical studies in other languages are examined for the reproducibility of the five-factor framework. Because these studies differ in a number of their methodological details, we consider the potential impact of these differences on the findings obtained. Possible alternative structural frameworks are also considered. Finally, we consider the broader significance of this body of research.

THE LEXICAL APPROACH IN ENGLISH

The lexical assumption has been the guiding rationale for personality taxonomers dating back at least to the 19th-century British scientist, Francis Galton (1884), who used a dictionary to identify terms that described personality attributes. Particularly influential were Allport and Odbert (1936), who extracted from an unabridged English dictionary almost 18,000 words that could be used to describe individual differences. Of these, roughly 4500 were designated by Allport and Odbert as likely personality traits, the remainder having been classified as temporary moods or activities, social effects, physical or medical terms, predominantly evaluations, and the like. This classification scheme proved highly influential, leading to the decision by Norman (1967) and Goldberg (1982) to exclude from their initial analyses terms describing one's physical

appearance, the effects one has on others, sheer evaluations, and temporary activities, moods and states.

The detailed history of the initial extraction and later modification of trait lists has been documented repeatedly (e.g. Digman, 1990; Goldberg, 1993, 1995; John, 1990; McCrae and John, 1992; Wiggins and Trapnell, 1997), and thus will not be repeated here. In this summary, we will focus primarily on the English-language studies that permit comparisons with those in other languages.

At the dawn of the computer age, lexical studies in English found impetus from work by Norman (1963, 1967) and Goldberg (1976, 1982). Norman's (1963) use of Roman numerals to identify the Big Five factors has become a hallmark of the lexically based taxonomic tradition; the five factors were ordered roughly by the number of reasonably familiar trait-descriptive adjectives in English available to describe that domain: I (Extraversion) and II (Agreeableness) with the most such terms, followed by III (Conscientiousness), and then IV (Emotional Stability) and V (first labelled Culture, here labelled Intellect) with the least adjectives.

Norman (1967) retraced the footsteps of Allport and Odbert (1936) and extracted person-descriptive terms from a new edition of the same unabridged English dictionary that they had used. Norman's four-person research team refined the resulting set of 18,125 terms by excluding those classified as pure evaluations and as physical and medical terms and categorizing the remainder as (a) stable traits, (b) temporary states and activities or (c) social roles, relationships and effects. Through this process, Norman identified roughly 2800 stable trait terms. Goldberg (1976, 1982) reduced Norman's stable trait pool to 1710 by eliminating the least commonly used terms and those that seemed the least dispositional in nature. Using these same subjective criteria, Goldberg later developed a 540-term set that was employed in a number of studies in which university students described themselves and others (Goldberg, 1990). After additional samples of university students had provided familiarity ratings for these and other common English adjectives, Saucier and Goldberg (1996a) analysed the most familiar subset of 435 adjectives in a combined sample of nearly 900 (including 507 self-descriptions and 392 descriptions of other), and found a clear five-factor structure.

Peabody (1987) developed 53 bipolar scales to represent a large pool of adjectives that included Goldberg's 540 terms. Four college students made judgements of semantic similarity between all of the terms from each pole and each of the bipolar scales. Analyses of these judgements of 'internal structure' revealed the Big Five factors, plus a small 'Values' factor; the first three factors (Assertiveness [Extraversion], Agreeableness, and Conscientiousness) were, however, much larger than the remaining ones. Further analyses of these bipolar scales in self and peer descriptions by college students consistently found the Big Five factors (Peabody and

Goldberg, 1989). But again, the first three factors (i.e. Extraversion [I], Agreeableness [II] and Conscientiousness [III]) were substantially larger than the other two (Emotional Stability [IV] and Intellect [V]).

The English-language studies of Goldberg and Peabody share at least one limitation: potentially relevant terms referring to states, social evaluations, and physical and appearance characteristics were excluded from those item pools. A study by Tellegen and Waller (1987; summarized also by Waller, in press; Waller and Zavala, 1993) cast a wider net. Instead of relying on Allport's or Norman's compilations, an abridged dictionary was divided into roughly sixty 25-page sections, and seven or eight non-contiguous pages from each section were randomly selected. On each selected page, the first adjective that could be fitted into the stems 'tends to be ____' and 'is often ____', or otherwise appeared to be person-descriptive, was extracted.³ A set of 400 terms was selected, including terms describing social effects, pure evaluations and temporary states.

Self-reports using these 400 terms were provided by almost 600 university students. In order to encourage meaningful responses to these terms, many of which would be unfamiliar to most native speakers, participants were presented with portions of the dictionary definitions. Factor analyses (5 to 20 factors) of the 400 variables were conducted, and a seven-factor solution was judged to be the most compelling and meaningful (Waller, in press). The seven factors were judged to correspond to the Big Five, plus orthogonal dimensions labelled Positive Valence (e.g. Important, Outstanding) and Negative Valence (e.g. Evil, Vicious). The last two factors drew on the 'pure evaluation' descriptors excluded in all previous lexical-factor studies. Subsequent studies in Hebrew (Almagor, Tellegen and Waller, 1995) and Spanish (Benet-Martínez and Waller, 1997) used a similar methodology.

These 'Big Seven' studies have several strengths. Solutions with various numbers of factors were examined, the range of variables was very broad, and samples of both variables and subjects were large. However, variable selection was based on dictionary pages, rather than on entries. This procedure approximates to representative variable selection only if the distribution of important descriptors (and their semantic content) has a rectangular (flat) distribution across pages (or parcels of pages) in a dictionary. However, in fact some pages have many familiar person-descriptors (e.g. in English 'in ____'), whereas other pages have absolutely none (e.g. 'x ____'). As a result, page-sampled variable sets may often be unrepresentative of the lexicon, particularly if a language's person-descriptors are disproportionately prone to contain certain prefixes (e.g. good-, self-, well- in English). Additionally, such sampling procedures are likely to net many unfamiliar terms, violating the frequency-importance association that we believe is integral to the lexical approach. For example, in English, Rhadamanthine and Tenebrous are as likely to be sampled as Kind and Depressed.

Recently Saucier (1997) found a structure resembling the Big Five in a set of familiar English adjectives restricted to dispositions and states. When a wider range of terms, including social evaluations and physical attributes, was included, two additional factors emerged: Attractiveness and a factor resembling Big Seven Negative Valence. Whether one obtained the Big Five or the related seven-factor structure depended upon variable selection. In contrast, a three-factor solution was replicable across variable selections. The three factors, Extraversion, Agreeableness and Conscientiousness, were each somewhat broader than those from the Big Five; these are henceforth referred to as the Big Three.

In further studies, Saucier (1998) combined eight American samples (total $N = 3062$) with representative lexical variable selections, and found that structures of one, two and three factors were all more replicable than five-factor structures, with their superior replicability especially evident in peer-rating samples. The factors in the three-factor structure closely resembled the first three of the five-factor structure, suggesting that the most robust version of the Big Five may consist of the three large factors, along with two smaller factors whose replicability is weaker.

In English, then, a Big Five structure has been replicated across different investigators, variable selections and participant samples. These studies have indicated that the form and replicability of this factor structure is affected by at least two major aspects of one's methodology: (a) variable selection, and (b) the kind of targets (self, liked peers, diversely evaluated peers, semantic concepts) that are rated (Peabody and Goldberg, 1989). The extensive lexical studies in English, all conducted in the United States, have pioneered many of the techniques subsequently applied to other languages. In the rest of this chapter, we will compare the findings from these cross-cultural lexical studies.

CROSS-LANGUAGE APPLICATIONS OF THE LEXICAL APPROACH

Cross-language studies of the five-factor framework have used both etic (imported) and emic (indigenous) procedures (Berry, 1969). In the former, an imported (usually Western) framework is tried out in the new culture to see how well it fits (e.g. a Big Five measure is translated into another language), whereas emic approaches allow for the indigenous framework to emerge without imposing constraints (e.g. a representative sample of the language's trait adjectives is analysed). If, for example, the Big Five is reproduced in etic studies, it does not constitute evidence for the cross-cultural ubiquity of the personality factors, but only that people in other cultures are able to employ those dimensions. In most lexical studies, investigators use an emic approach to identify the indigenous factors of

personality description, and then use an etic measure to compare these dimensions with ones found in other languages. The present review is focused only on emic studies, and thus does not include studies primarily designed to develop reliable and valid measures of the English five-factor model in another language (e.g. Benet-Martinez and John, 1998; McCrae and Costa, 1997).

Emic lexical research is summarized in Table 1.2. Not included are ongoing projects for which reports are not yet available; for example, lexical studies are presently being carried out in Portuguese, Quechua and Romanian (De Raad *et al.*, 1998), as well as in Croatian, Greek, Norwegian and Slovak. In Table 1.2, the language under study is provided in the first column; the second column lists the names of the investigators; the third column indicates the source of the terms used in that study; and the fourth column provides the type of data that was obtained (most typically self or peer descriptions, or both). The method(s) used to evaluate factor equivalence are provided in the fifth column, and the final column provides a very rough summary of the findings. The particular Big Five factors that were seemingly identified in each study are indicated by their traditional Roman numerals (I to V).

We begin our review with studies in three subgroups of Indo-European languages: Germanic (e.g. German, Dutch), Slavic (e.g. Czech, Polish, Russian), and Romance (e.g. Italian, Spanish).

GERMANIC LANGUAGES

The German personality taxonomy

Much of psychology has originated in Germany, and this is also true of lexical studies. Klages (1932) articulated the lexical rationale, and Baumgarten (1933) created the first list of (1629) German personality descriptors, both of which influenced the work of Allport and Odbert (1936). Lexical studies of the German language were taken up again in the 1980s when 5101 personality-relevant adjectives (e.g. cynical), 2212 type nouns (e.g. cynic) and 3607 attribute nouns (e.g. cynicism) were extracted from a comprehensive dictionary and supplementary lexical sources. Judges classified the adjectives into 5 broad and 13 subordinate categories, a refinement of the category systems used by Norman (1967); for greater detail, see Angleitner, Ostendorf and John (1990).

Ostendorf (1990) studied those 430 familiar adjectives that a majority of judges had classified as either temperament and character traits or abilities and talents, or their absence. Over 400 adult participants completed self-reports using the 430 adjectives and measures of imported Big Five and circumplex models, and 95 per cent of those individuals were described

Table 1.2 Summary of the lexical studies compared in Chapter 1

Language	Authors	Source of terms	Data collected	Method for determining factor equivalence	Replicated Big Five factors
English	Goldberg (1990)	Unabridged dictionary	Self-ratings Peer ratings	Inspection of factor content	I, II, III, IV, V
	Tellegen and Waller (1987)	Abridged dictionary	Self-ratings	Inspection of factor content	I, II, III, IV, V (plus NV, PV)
	Saucier and Goldberg (1996a)	Familiar terms from Goldberg (1990)	Self-ratings Peer ratings	Inspection of factor content	I, II, III, IV, V
German	Saucier (1997)	Familiar terms (broad set)	Self-ratings Peer ratings	Inspection of factor content	I, II, III, IV, V (plus NV, Att)
	Ostendorf (1990)	Comprehensive dictionary	Self-ratings Peer ratings Big-Five markers	Correlations among factor scores from different types of data	I, II, III, IV, V
Dutch	Brokken (1978)	Unabridged dictionary	Self-ratings Peer ratings	Inspection of factor content	I, II, III, IV, but V less clearly
Czech	De Raad et al. (1992)	Brokken's (1978) set of terms	Self-ratings	Inspection of factor content	I, II, III, IV, but V less clearly
	Hřebíčková et al. (1995)	Standard dictionary	Self-ratings Peer ratings NEO	Correlations with NEO Inspection of factor content	I, II, III, IV, V, with IV less clear
Polish	Szarota and Ostendorf (1997)	Concise dictionary	Self-ratings Peer ratings NEO	Inspection of factor content Correlations with NEO	I, II, III, IV, V, with I and IV less clear
Russian	Shmelyov and Pokhil'ko (1993)	Previous lists Dictionaries	Trait-inference judgements	Congruence with English markers	I, II, III, IV, V, with IV less clear
Italian	Caprara and Perugini (1994)	Abridged dictionary	Self-ratings NEO BFQ	Inspection of factor content Correlations with NEO	I, III, and V, with II and IV rotational variants, and V similar to V in Dutch.
	Di Blas and Forzi (1998)	Unabridged dictionary	Self-ratings Peer ratings	Inspection of factor content Correlations with Big Five markers	I, III, with II and IV each splitting into two factors
Spanish	Benet-Martínez and Waller (1997)	Abridged dictionary	Self-ratings Big Seven measure Big Five measure	Inspection of factor content Correlations between ermic and etic factors	I, II, III (plus PV and NV) content
Hebrew	Almagor et al. (1995)	Abridged dictionary	Self-ratings	Inspection of factor content	II, III, V (plus NV), with I and IV splitting into two factors
Hungarian	Szirmák and De Raad (1994)	Unabridged dictionary	Self-ratings	Inspection of factor content	I, II, III, IV, with V less clear
Turkish	Somer and Goldberg (1999)	3 abridged dictionaries	Self-ratings Peer ratings	Inspection of factor content	I, II, III, IV, V
	Goldberg and Somer (1999)	Most familiar terms from Somer and Goldberg (1999)	Self-ratings	Correlations with Turkish Big Five markers	I, II, III, IV, V
Korean	Hahn, Lee and Ashton (1999)	2 dictionaries Terms from free descriptions	Self-ratings	Correlations with Korean Big Five markers	I and II, with III, IV, and V less clear
Filipino (Tagalog)	Church et al. (1997)	Abridged dictionaries	Self-ratings NEO Big Five markers	Correlations with Big Five markers	I, II, III, V; with IV split into two factors

Note

I = Extraversion, II = Agreeableness, III = Conscientiousness, IV = Emotional stability, V = Intellect, PV = Positive Valence, NV = Negative Valence, Att = Attractiveness, NEO = NEO Personality Inventory, BFQ = Big Five Questionnaire.

by from one to three acquaintances. Analyses of the 430 adjectives generated five highly reliable factors in ipsatized ratings of either self or acquaintances. Correlations of these factors with corresponding factors from the imported measures of the Big Five averaged over 0.70 (see also Ostendorf and Angleitner, 1994, Table 1).

The three-factor solutions were even more replicable; only at the three-factor level were all factor-comparability coefficients for both ipsatized and raw data above 0.90.⁴

Lexical studies in The Netherlands

The first lexical studies in Dutch were carried out by Brokken (1978) and Hofstee (Hofstee, Brokken and Land, 1981). Further data and analyses were provided by De Raad, Hendriks and Hofstee (1992). In the original studies, 8690 person-descriptive adjectives were extracted from a large dictionary and then reduced to 6055 that were judged by at least one of four individuals to be stable traits. At a later stage of the project, those 1203 adjectives that fitted best into the sentence stems 'He/she is ____ by nature' and 'What kind of person is he/she? ____' (determined by the average ratings of university students) were retained. These criteria seem to have served to retain temperament and character terms more abundantly than those describing abilities and talents, with the result that factor V in the Dutch solution has a somewhat less intellectual character than the fifth factor in English and in German.

Using the 1203 adjectives, 200 pairs of university students who knew each other well rated both themselves and their partners (Brokken, 1978). Later, De Raad (1992) collected 200 additional self-ratings from university students on 551 of the 1203 adjectives selected by more stringent use of the same criteria used by Brokken. The ipsatized ratings of these 600 participants (400 self, 200 acquaintance) were pooled in a factor analysis of the 551 adjectives; four, five and six factors were rotated. In each of these solutions the first four factors included Extraversion (I), Agreeableness (II), Conscientiousness (III) and Emotional Stability (IV) from the Big Five. The fifth factor in the five-factor solution contrasted intellectual autonomy and independence with conventionality (e.g. Critical, Rebellious versus Meek, Docile), and was interpreted as a variant of Big Five Intellect. In general, the Dutch studies supported the Big Five (though only weakly for the Intellect factor; see De Raad, 1994).

De Raad and his colleagues have pioneered the application of the lexical approach to personality-type nouns (De Raad and Hoskens, 1990; De Raad and Ostendorf, 1996) and personality-relevant verbs (De Raad *et al.*, 1988). These studies indicate that personality nouns tend to have stronger evaluative connotations than do adjectives. Some of the noun factors (labelled 'Malignity') are reminiscent of Negative Valence from the Big Seven. The

verb studies led to a two-factor representation, with one factor defined by verbs like Care for and Cooperate (versus Curse and Threaten), and the other factor defined by verbs like Decide and Lead (versus Flee and Brood); the factors have some resemblance to Big Five Agreeableness (II) and Extraversion (I), respectively. They probably also relate to factors in two-factor solutions from adjectives in some other languages (e.g. Saucier, 1998).

Studies in these Germanic languages have been supportive of a *stringent* emic Big Five model, according to which the same 'Big Five' factors emerge in five-factor solutions, when based on a representative set of single-word personality descriptors in a language.⁵ The stringent requirements of this model can be distinguished from a *looser* counterpart, in which the appearance of the Big Five factors in any solutions of five or more factors is interpreted as supportive evidence for the Big Five model.

SLAVIC LANGUAGES

Polish

In a study by Szarota (1996; Szarota and Ostendorf, 1997), two judges independently scanned a concise Polish dictionary for person-descriptive adjectives, and the 1811 terms selected by both judges were retained; to this list were added a further 28 terms taken from other lists of Polish personality descriptors. Ten judges then rated these 1839 terms for clarity of meaning and personality relevance, and classified them using the system developed by Angleitner, Ostendorf and John (1990). The 290 adjectives categorized as dispositions were used for self and peer ratings in a sample of 369 high-school students. In both the self and the peer ratings of ipsatized data, five factors similar to the Big Five factors could be identified. Agreeableness (II), Conscientiousness (III) and Intellect (V) were virtually identical to the English and German counterparts. However, the Extraversion (I) factor had no sociability facet, and Emotional Stability (IV) included content related to self-control.

Czech

Hřebíčková, Ostendorf and Angleitner (1995) provided a preliminary report on a lexical analysis of the Czech language. All potentially personality-relevant terms were extracted from a standard Czech dictionary. The 4145 such terms were categorized by six judges into 13 different types of descriptors, based on the classification system developed by Angleitner, Ostendorf and John (1990). Those 366 terms that the majority of judges had classified as dispositions were rated for self-descriptiveness by 397

research participants. Using ipsatized data, factor solutions with five and more factors were examined and interpreted both by inspection of the high-loading terms and by correlations with expert prototypicality ratings with respect to the Big Five. Four of the Big Five factors were clearly identifiable, with Emotional Stability being the only problematic factor (the Czech version of Emotional Stability was also related to Agreeableness and Intellect when correlated with the prototypicality indices for these factors). Generally, however, the Czech study provides substantial support for the Big Five structure.

In a subsequent study, Hřebíčková *et al.* (1999) studied the structure of Czech personality-relevant verbs. The two-factor solution was similar to the Dutch two-factor verb structure described above. The authors reported solutions with up to six factors; notably, in no solution did a factor emphasizing Intellect content appear.

Russian

Shmelyov, Pokhil'ko and Kozlovskaya-Telnova (1988, 1991; Shmelyov and Pokhil'ko, 1993) completed the earliest Russian lexical work without access to reports of the major earlier studies. Consulting two dictionaries and previous lists of trait-descriptive terms, they compiled a set of 2090 personality-relevant terms, including 666 nouns (most of which seemed to refer to undesirable traits) and 1424 adjectives. Studies of the internal structure (semantic similarity) of large subsets of this pool were reported by Shmelyov and Pokhil'ko (1993). The investigators rotated 15 factors, far more than in other studies, and these were compared with factors from translations of adjective clusters and cluster labels from the studies by Goldberg (1990). The four largest Russian factors (Altruism [II], Intellect [V], Energy-gaiety [I], and Conscientiousness [III]) and to a lesser degree the sixth (Emotional Stability [IV]) were interpreted as highly congruent with the Big Five factors from Goldberg (1990).

At present, the findings from these Russian studies are difficult to compare with those from other studies, due to their exclusive use of internal rather than external data, and the lack of evidence about the nature of factors at higher hierarchical levels. Nonetheless, most of the Big Five factors (Emotional Stability less strongly) appeared in these concept-rating analyses. Therefore, we can regard the Russian studies, like the other two in Slavic languages, as both (a) loosely supportive of the Big Five and (b) indicating possibly lower replicability for the Emotional Stability factor. An unpublished study of Croatian, a south Slavic language, also found a recognizable Big Five structure in ipsatized ratings, although there was some rotation in the plane defined by the Agreeableness and Emotional Stability factors (Mlačić, 1999). In contrast, the replication of the Big Five has been more problematic in analyses of Romance languages.

ROMANCE LANGUAGES

Studies in Italy

Two independent taxonomic projects have been conducted in Italy – a 'Roman' project conducted by Caprara and Perugini (1994), and a 'Trieste' project conducted by Di Blas and Forzi (1998, 1999). The procedures used in the Roman project were similar to those used in the Dutch studies, whereas the procedures used by the Trieste investigators were similar to those used in the German project (De Raad, Di Blas and Perugini, 1997).

In the Roman project (Caprara and Perugini, 1994), an abridged dictionary was scanned for person-descriptive adjectives and 8532 were selected. Later, 1337 of these terms were retained by a criterion of utility for describing personality, based on the judgements of four experts. Further utility ratings by 22 lay judges were used to reduce the number of terms to 492, which were administered to 274 research participants for self-ratings. The five-factor solution from analyses of these ipsatized data revealed Big Five Extraversion [I] and Conscientiousness [III] factors. A third factor was labelled Quietness [or Peacefulness] versus Irritability; a fourth factor, labelled Selfishness versus Altruism, included adjectives related to tough-mindedness and emotionality. The investigators concluded that the third and fourth factors were rotational variants of Agreeableness (II) and Emotional Stability (IV). The fifth factor, labelled Conventionality, included descriptors related to conformity and traditionalism, and thus resembled the fifth factor in Dutch.

In the Trieste project (Di Blas and Forzi, 1998), five judges searched through an unabridged Italian dictionary for person-descriptive terms. They extracted 11,010 terms, including 3780 adjectives, 1428 adjectives that can be used as type nouns, 2566 type nouns, and 3326 attribute nouns. The 1586 most personality-relevant, frequently used and clear terms (as determined by ratings from at least 10 secondary school students) of the adjectives and the adjectives that can be used as type nouns were later classified by university students into the 13 categories developed by Angleitner, Ostendorf and John (1990). The most prototypical dispositional terms (as defined by high inter-judge agreement), which included 344 adjectives and 124 adjectives that could be used as type nouns, were then reduced to 314 and 118 of each kind by the elimination of synonyms and highly ambiguous terms.

The set of 314 adjectives was administered to a heterogeneous sample of 427 individuals for self-ratings and another sample of 277 secondary school students for ratings of someone they knew well. The two data sets were ipsatized and analysed both separately and jointly, with three- to nine-factor solutions being examined. Self and other ratings produced comparable three-factor solutions that were easily identifiable as Big Five

factors I, II and III. In the five-factor solutions, Extraversion and Agreeableness each split into a pair of more specific factors (Assertiveness and Sociability; Quietness [or Peacefulness] and Tender-mindedness). Distinct aspects of factor IV appear to have been merged with the interpersonal aspects of factors I and II to produce Assertiveness (versus Fearfulness) and Quietness (versus Irritability), with some other aspects associated with Tender-mindedness. A rather weak version of an Intellect factor only emerged when at least seven factors were rotated.

Subsequently, Di Blas and Forzi (1999) attempted to replicate this structure in a study of ipsatized self-ratings from 369 participants using a set of 369 terms.⁶ To construct this new variable set, the investigators excluded 63 terms with very skewed response distributions from their combined set of 314 adjectives and 118 adjectives that can be used as type nouns (described above). In these new analyses, four factors from the Italian five-factor solution, labelled Sociability (weakly related to Big Five factor I), Placidity/Quietness (a blend of Big Five factors II and IV), Conscientiousness (III), and Self-assurance/Assertiveness (a blend of Big Five factors I, IV and V), included content from the first four of the Big Five, but there were no exact counterparts.

Findings from the two Italian projects differ with regard to the nature of the Extraversion factor and the identification of a factor comparable to Big Five Intellect. The Roman set of descriptors included some non-dispositional terms, notably temporary states and social effects, and thus appears to be broader than the Trieste set. Generally, the Italian studies illustrate an important principle: a language containing terms describing Emotional Stability or Intellect may still not yield these Big Five lexical factors, because the terms may be highly associated with terms from other factors.

Spanish

Dictionary-sampling procedures were employed in a study of Spanish personality-descriptors (Benet-Martínez and Waller, 1997). Every fourth page of a 1666-page dictionary was inspected, using selection criteria like those of Tellegen and Waller (1987) and 299 adjectives were selected. As in the American study, the adjectives were accompanied by a synonym or short definition. Self-ratings were elicited from a sample of 894 university students. Prior to factor analysis, the effects of gender, age and primary language (Spanish versus Catalan) were statistically controlled, and the data ipsatized. Solutions of five to eight rotated factors were examined, and the seven-factor solution was reported to be 'easily interpreted', whereas the others were 'structurally ambiguous'. The seven-factor solution corresponded in some respects to the seven factors reported by Tellegen and Waller (1987) and measured by the Inventory

of Personality Characteristics (IPC; Tellegen, Grove and Waller, 1991). Correlations between the seven indigenous factors and those from the IPC were between 0.60 and 0.79 for 'Pleasantness' (which resembles both Positive emotions and Extraversion factors in other studies), Agreeableness, Conscientiousness (labelled 'temperance'), and Positive Valence. The other three factors were substantially smaller in size. One of them correlated moderately (0.47) with Negative Valence, whereas the other two indigenous factors, Engagement and Openness (Introspectiveness might be an alternative label), had low correlations with the remaining IPC factors.

The five-factor solution included Agreeableness, Conscientiousness, Conventionality and Negative Valence factors. There was also a large factor that counterposed Positive Valence and Positive Emotionality with Negative emotionality; this factor might resemble Assertiveness (from Italian) or the broader, Big Three factor of Extraversion.

In summary, Benet-Martínez and Waller's (1997) seven-factor solution might be seen as a replication of the three largest factors in Peabody and Goldberg (1989), as well as Positive Valence and (weakly) Negative Valence. As in the Italian studies, Emotional Stability subcomponents were spread across several factors, and Intellect was not clearly recovered. Given the use of page-sampling procedures, the potential for replication of either the Big Seven or Big Five models may be limited. Because the Spanish factors include versions of the first three Big Five factors, it would be useful to examine one-, two-, three- and four-factor solutions in these data.

As our review has progressed from Germanic to Slavic and on to Romance languages, the emic replication of the English/German Big Five has waned steadily. Does this trend extend itself into non-Indo-European languages? On both linguistic and cultural grounds, if the Big Five are not culturally robust, then they should be even less evident in non-Indo-European languages.

NON-INDO-EUROPEAN LANGUAGES

Hebrew

Hebrew, like Arabic, is a member of the Semitic language group, and thus could provide an interesting contrast with other lexical studies. Almagor, Tellegen and Waller (1995) carried out such a study using a dictionary page-sampling technique similar to that used in the Spanish study just described. Two research assistants searched every fourth page of a 1600-page Hebrew dictionary and extracted the first personality-descriptive adjective or noun they encountered. This process yielded 326 terms, which were reduced to 252 after removing synonyms. University students ($N = 637$) rated the 252 words for self-descriptiveness.

The investigators focused on a seven-factor rotation, which they compared with the Big Seven model (Tellegen and Waller, 1987) by observing the factor loadings of the Hebrew translations of Big Seven marker traits. The Hebrew factors labelled Agreeability (II), Dependability (III) and Negative Valence seem roughly equivalent to their Big Seven (and thus Big Five) counterparts. However, the terms with the highest loadings on the Positive Valence factor were all related to Intellect (e.g. Sophisticated, Sharp, Knowledgeable). Emotional Stability terms were found on factors labelled both Positive Emotionality/Agentic (e.g. Depressed, Sad) and Negative emotionality (e.g. Bad-tempered, Angry, Nervous), whereas the Positive Emotionality/Communal factor included terms that seem highly related to Sociability (e.g. Likeable, Enthusiastic, Friendly, Talkative).

This Hebrew study was limited by its dictionary page-sampling method, which reduced the representativeness of the terms, as did the decision to reduce the item pool by almost 25 per cent by omitting all synonyms. None the less, some of its findings were similar to those found in a few other languages. At the seven-factor level: (a) Emotional Stability terms were associated with more than one factor; (b) Extraversion split into two factors (labelled Positive Emotionality/Agentic and Positive Emotionality/Communal); (c) Intellect-related terms were mixed with some other highly desirable terms to form a factor labelled Positive Valence; and (d) a Negative Valence factor included all of the highly pejorative terms in the variable set, plus a few of the most desirable ones (e.g. Honest, Sincere, Dependable, Reliable). Thus the so-called Negative Valence factor was moderately bipolar, unlike its counterparts in English and Spanish Big Seven studies.

Hungarian

Szirmák and De Raad (1994) reported the findings from some recent lexical studies in Hungarian, one of the Finno-Ugric languages. Hungarian is of particular interest because it is a non-Indo-European language that is spoken by central Europeans who live in close proximity to numerous European languages and cultures. Over 8000 person descriptors were extracted from dictionaries. Five judges provided ratings of familiarity, personality-relevance, and stability (trait versus state) for nearly 4000 adjectives; the 624 trait-descriptive terms with the highest mean ratings on these three criteria were retained. Self-ratings on these 624 terms were provided by 400 university students, and were ipsatized prior to analysis.

Three-, four-, five- and six-factor solutions were examined (De Raad and Szirmák, 1994). In the four-factor solution, the factors resembled Big Five Extraversion (I), Agreeableness (II), Conscientiousness (III) and

Emotional Stability (IV), the first three of which made up the three-factor solution. In the five-factor solution, Agreeableness divided into two factors, one emphasizing calmness versus irritability, and the other (labelled Integrity) emphasizing egotism-related variables. The six-factor solution added an Intellect factor. Thus the Hungarian studies support the Big Three and Big Four (Big Five minus Intellect) models and also, by a loose rather than stringent criterion, the Big Five model.

Turkish

The Turkish language, a member of the Altaic group, was studied by Somer and Goldberg (1999) and Goldberg and Somer (in press). Five native speakers picked out 2200 person-descriptive adjectives from three modern abridged Turkish dictionaries. Most, but not all, terms describing physical attributes, sheer evaluations, and slang terms were removed, leaving 1300 terms that were judged for familiarity as personality descriptors by 150 university students. Three separate studies were carried out with different subsets of the most familiar of these 1300 terms.

The first two of these Turkish studies were reported by Somer and Goldberg (1999). In their Study 1, they used 474 familiar adjectives grouped into 358 synonym clusters and arranged as 179 pairs of antonym clusters. Ratings of self, liked-peers and disliked-peers were obtained for both poles of the antonym clusters from 232 university students. Two, three, four, five, six and seven factors were rotated. Analyses of the self and liked targets produced an easily interpretable five-factor solution reflecting the Big Five, whereas analyses of the liked plus disliked peers produced a five-factor solution in which the first four of the Big Five factors emerged quite clearly but the fifth factor appeared as a blend of Openness-Imagination with Attractiveness. The three-factor solution in both cases produced broad versions of factors I, II and III, as in most of the languages in which such solutions have been reported.

In a second study, the 358 terms were reduced to a simpler and more user-friendly set of 235, and self and peer ratings were obtained from a much larger sample (945 university students); roughly half of the sample described themselves, and the other half described either liked, disliked or neutrally evaluated peers. Solutions for one to seven factors were obtained for the combined self and liked-peer targets, as well as for the combined liked, neutral and disliked-peer targets. The five-factor solution reproduced the Big Five quite cleanly, but with factor V closer to Openness than Intellect. The Big Five were also identifiable in the five-factor solution for the ratings of the liked, neutral, and disliked-peers. Again, for both sets of ratings, the three-factor solution produced broader versions of factors I, II and III.

To answer any possible charges that the synonym clusters used in the two studies by Somer and Goldberg (1999) might have been prestructured in some way, Goldberg and Somer (in press) administered the 498 most familiar single terms from the initial set of 1300 Turkish person-descriptive adjectives to over 600 university students for self-descriptive ratings. In the total item pool, using ipsatized ratings, Big Five factors were identifiable, along with a broad Attractiveness factor and another factor composed of items of extremely low endorsement rates. When the item pool was restricted to terms that are less pejorative and more clearly related to personality traits, the Big Five factors were recognizable. Again, as in Somer and Goldberg (in press), the fifth factor blended Intellect with Modernism/Progressivism. And Extraversion (I) and Emotional Stability (IV) tended to be more highly associated with each other in the Turkish data than they are in the English/German version of the Big Five.⁷

The Turkish findings demonstrate that the Big Five can be reproduced in an emic study in a non-Indo-European language, at least as long as ipsatized ratings are employed, albeit with some possible cultural particularities for factor V. This factor may be the most susceptible of the five to effects of variable selection and to cultural differences (Somer and Goldberg, in press). Turkey has a predominantly Islamic cultural heritage while also being Westernized, and the traditionalism versus modernism theme in this factor may reflect that context.

Korean

In a study by Hahn, Lee and Ashton (1999), two Korean dictionaries and free descriptions of personality written by university students were searched for personality terms by a single investigator. Physical attributes, evaluations and temporary states were excluded, as well as some synonyms of included terms, resulting in about 1000 terms that were reduced to 785 when four judges assessed them for familiarity and adequacy for describing personality. These 785 were rated for frequency of use by 125 university students, and 406 of high familiarity were selected. Undergraduates ($N=435$) from three Korean universities rated themselves on these 406 terms, and the data were ipsatized. A set of 46 Korean markers of the Big Five were identified and confirmed in factor analyses. Three to seven factor solutions were examined and correlated with the Korean Big Five marker factors to assist in their interpretation.

In the three-factor solution, the familiar Big Three were found. In the four-factor solution, the Korean emic factors were labelled Extraversion (correlating 0.91 with Big Five Extraversion), Agreeableness (correlating 0.87 with Big Five Agreeableness), Conscientiousness/Precision (related

to both Big Five Conscientiousness and Intellect), and Masculinity/Emotional Stability (correlating 0.79 with Big Five Emotional Stability). The Emotional Stability factor included a number of gender-related terms (e.g. Manly, Feminine), along with those reflecting self-efficacy or potency (e.g. Strong, Weak); these features may account for its lower replication coefficient. In the five-factor solution, the Extraversion, Agreeableness and Masculinity/Emotional Stability factors remained much the same, whereas terms reflecting Conscientiousness and Intellect merged to form a Methodical intellect factor, with the other Conscientiousness factor now emphasizing Dependability. The six-factor solution produced a factor labelled Truthfulness that was judged to be similar to the Integrity factor identified in Hungarian by Szirmák and De Raad (1994).

The investigators concluded that the Korean personality factors were quite similar to the Big Five. However, there were notable divergences from the Anglo-Germanic Big Five, at least with respect to the axis locations of several factors.

Filipino

Filipino (Tagalog), a member of the Austronesian language family, has borrowed some of its vocabulary from Spanish. It is the first and only native language of a Pacific Island or tropical region to be examined in a lexical personality study. Church, Katigbak and Reyes (1996) searched a comprehensive Filipino dictionary, extracting 6900 person-descriptive adjectives, which were classified into the categories developed by Angleitner, Ostendorf and John (1990). The 682 most familiar and personality-relevant of these were used for self-ratings by both college and high school students (Church *et al.*, 1997); these ratings were ipsatized. When five factors were rotated, four of them included content resembling that found in the Big Five, but rearranged into factors labelled Gregariousness, Socialization, Perceived competence, and Egotism; a fifth Negative Valence factor was composed entirely of pejorative terms. In the six-factor solution, a Temperamentalism factor split apart from Perceived competence.

In a replication study (Church, Katigbak and Reyes, 1998), 740 Filipino college students provided self-ratings with a revised set of 502 terms, which also included some highly desirable and highly undesirable terms that might be expected to form Positive and Negative evaluation factors. The research participants also rated themselves on translations of Big Five marker scales. In analyses of ipsatized data, seven Filipino factors were labelled Gregariousness (I), Concern for others versus Egotism (II), Conscientiousness (III), Self-assurance (IV), Intellect (V), Negative Valence/Infrequency and Temperamentalness, another type of (negative)

Emotional Stability. Positive Valence markers (Big Seven) blended with terms related to Intellect to form factor V in this representation.

In both of these Filipino studies, at least four of the seven Filipino factors seem to correspond to the English/German Big Five. Is this a replication of that model? John and Srivastava (1999), applying a loose criterion for replication, interpreted it as supportive; however, by a stringent criterion, the Filipino studies do not replicate the Big Five model. With respect to the Big Seven, a Negative Valence dimension replicated similar factors found earlier in English, Spanish and Hebrew. There was no unambiguous Positive Valence factor in the Filipino analyses, and the appearance of an Attractiveness factor (see Saucier, 1997) would be unlikely due to the exclusion of terms referring to physical characteristics and social effects.

In several respects, the Filipino structure was the *least* like that found in English and other northern European (Germanic, Slavic) languages. The Conscientiousness factor was smaller than in previous studies, thus seeming to verify other reports indicating that this factor may be subject to differential emphasis according to the cultural setting (Slotboom *et al.*, 1998; Victor *et al.*, 1998). As was the case in the Romance language and Hebrew studies, Emotional Stability content was distributed across at least two factors, which was also true of Extraversion content. Moreover, unlike the findings from every other language in which three-factor solutions have been examined, the Filipino three-factor solution did not include broad versions of Big Five factors I, II and III (the Big Three). Instead, Agreeableness and Conscientiousness content remained intertwined until a lower level in the hierarchical structure.

Nonetheless, there are some similarities between the Filipino structures and those reported in Italian by Di Blas and Forzi (1998, 1999). In the seven-factor solutions from both languages, Extraversion content bifurcated into Gregariousness (or Sociability) and Self-assurance (or Assertiveness) factors, whereas Agreeableness content bifurcated into factors that might be labelled Peacefulness versus Irritability (or Temperamentalism) and Altruism versus Egotism. Moreover, at no hierarchical level in either language was there a single Emotional Stability factor; instead, this content was distributed across other factors. The structures that were derived from Hungarian, Hebrew and Spanish each showed at least some of the same bifurcations of Extraversion and Agreeableness, suggesting that deviations from the Big Five structure are neither language-specific nor entirely random. Perhaps a useful alternative to the Big Five will spring from these initial findings.

PREVIOUS COMPARISONS AMONG LANGUAGES

De Raad, Perugini and Szirmák (1997) used coefficients of congruence to make pair-wise comparisons of personality factors among five European languages, based on pairs of terms that seemed to be reasonably similar in each of the two languages. Congruence was consistently high (> 0.80) for the first three of the Big Five factors, somewhat lower for factor IV and even lower for factor V. In a study within one language, De Raad, Di Blas and Perugini (1997) demonstrated that structures found in two different Italian studies were similar, despite differences in their procedures for variable selection. Moreover, the two structures were more similar to one another than they were to the imported German factor structure. De Raad *et al.* (1998) compared seven languages (English, Dutch, German, Hungarian, Italian [Roman project], Czech and Polish). Within each language, they identified subsets of terms that had reasonably clear English equivalents, and they then computed congruence coefficients between (a) the factors based on all the terms that had been included in that variable selection and (b) the factors based on the subset of English translations. Across these seven languages, some evidence for the equivalence of the first four of the Big Five factors was consistently found, whereas factor V was not as consistently identified.

Hofstee *et al.* (1997) compared Dutch, English and German Big Five structures. They identified a set of 126 triads of terms that were at least roughly equivalent in all three languages. They then compared the factor structure for these 126 terms across the three languages and found modest cross-language factor congruence. They concluded that the Big Five replicated across these three quite similar languages, although the degree of equivalence was far from perfect.

Each of these comparative studies focused solely on the five-factor structures. Future studies of this type might profitably examine a variety of hierarchical levels. As one example, Saucier and Ostendorf (1999) examined the replicability of specific subcomponents of the Big Five in parallel analyses conducted with English and German lexical data. They found 18 replicating subcomponents, indicating there are some cross-language regularities at a hierarchical level more specific than that of the Big Five, at least in these two Germanic languages.

Most of these cross-language comparison studies have focused on the question 'Does the Big Five replicate in analyses of other languages?' In part, the answer depends on whether one applies a loose or a stringent criterion for replication. Clearly, content from each of the Big Five factors can be found among person-descriptors in all of the languages studied so far. However, with the exception of Turkish, only studies of Germanic and Slavic languages have found a relatively clear Big

Five structure where it would most be expected – namely in five-factor solutions. Upon reflection, it becomes clear that a question like ‘Does the Big Five replicate in analyses of other languages?’ is more complex than it first appears.

WHEN IS A DIFFERENCE A REAL DIFFERENCE?

How does one tell when a difference in factor representations between two languages reflects a true cultural difference, as compared with some methodological or other artefact? This problem is more complex than the related question of ‘factorial invariance’ (e.g. Hofer, 1999; Horn, 1991; Meredith, 1993; Sorbom, 1974), where exactly the same variables are measured in different groups of persons. In cross-language studies, in addition to differences between the languages and the cultures, virtually everything else may differ as well, including:

- 1 the methods for selecting an initial set of terms (e.g. from unabridged versus abridged dictionaries, using one versus more than one dictionary, examining all entries versus sampling entries or pages);
- 2 the size and representativeness of this initial set of terms;
- 3 the procedures used to cull the initial set of terms into one or more smaller subsets (e.g. the number of judges involved, the nature of those judges, the types of judgements used in the selection process, the order of those judgements in the culling process);
- 4 the extent to which the final set of terms includes descriptors of individual differences beyond those normally associated with personality traits (e.g. physical characteristics, temporary states, social effects, pure evaluations);
- 5 the size of the final set of terms;
- 6 the shape and extremity (skew) of the distribution of those terms on an evaluative dimension;
- 7 the nature of the final variables chosen for use (e.g. single terms, antonym pairs forming bipolar scales, synonym clusters, antonymic pairs of synonym clusters);
- 8 the nature of the judgements used to obtain indices of association among the variables (e.g. ‘internal’ versus ‘external’ data);
- 9 the size and nature of the subject sample used to provide those indices;
- 10 in the most common case, with external data, the nature of the targets being described (e.g. self versus others, and with others the extent of familiarity and liking of the subjects with the targets they describe);
- 11 the particular kinds of rating scales used to gather the final judgements;

- 12 the kinds of procedures (if any) used to cull potentially off-task individuals from the participant sample (e.g. the use of indices of semantic consistency);
- 13 the handling of missing data;
- 14 the extent to which the sample has been culled of semantically inconsistent subjects or subjects with excessive missing data;
- 15 the methods (if any) used to handle individual differences in subjects’ use of the response scale (e.g. Z-scoring, acquiescence partialling; see Ten Berge, 1999);⁸
- 16 the methods of data analysis (e.g. multidimensional scaling, cluster analysis, exploratory or confirmatory factor analysis);
- 17 in the most common case, with component or factor analysis, the number of factors extracted and the procedures used for factor rotation (e.g. orthogonal versus oblique rotations); and, finally,
- 18 the procedures used to compare the resulting factors across pairs of languages (e.g. expert or other judgements of factor similarity, correlations between indigenous and imported dimensions, analyses of subsets of translation-similar terms in the two languages).

Given these and other differences among studies, is it any wonder that investigators might disagree about the evidential basis for a particular structural representation? However, not all of these methodological differences are likely to be equally powerful as influences on the resulting structures. What is sorely needed at this juncture are some rigorous studies, within *each* of two or more single languages, examining the differential impact of these potential effects. Such purely methodological studies may not be of widespread interest, but their findings should have profound implications for our comprehension of the meaning of between-language factor differences.

In the interim before such findings are available, it will be difficult to say anything definitive about cultural differences in factor representations. Certainly, the more replicable the factor structure, the less susceptible it will be to these method effects. And, other things being equal, the more terms that are associated with a particular factor, the more replicable should be that factor. This reasoning would suggest that the most hardy factor should be the first unrotated one, which typically differentiates desirable from undesirable personality attributes. It is hard to imagine any language which would not include some descriptors of both types, and thus it ought to be the case that some sort of quasi-evaluative factor, contrasting good personality qualities with bad ones, should be recoverable in any language. Indeed, Osgood, May and Miron (1975) found that a large evaluation factor was widely distributed cross-culturally in judgements about heterogeneous sets of concepts. Such an evaluative

factor could turn out to be the only truly 'universal' aspect of personality description.

Clearly, then, there is value in looking for emic factors at a broader level in the personality-trait hierarchy than the Big Five. Two-factor solutions from several lexical studies suggest a consistent pattern: one factor tends to involve attributes associated with dynamism and individual ascendancy, whereas the other tends to involve attributes associated with socialization, social propriety, solidarity, and community. As related to the Big Five, the first factor is most highly associated with Extraversion (I), Intellect (V) and the fearfulness aspects of factor IV, whereas the second is most highly associated with Agreeableness (II), Conscientiousness (III) and the irritability aspects of factor IV. In addition to its appearance in lexical studies (e.g. Di Blas and Forzi, 1999; Hřebíčková *et al.*, 1999; Saucier, 1997, 1998), this pattern has appeared in other studies of personality structure (Digman, 1997; Paulhus and John, 1998; Shweder, 1972; White, 1980). It may represent projections of individualism and collectivism (Triandis *et al.*, 1988) into within-group variation.⁹ The reproducibility of this two-factor solution in diverse languages needs more attention.

A three-factor structure, which includes broad versions of Extraversion (I), Agreeableness (II), and Conscientiousness (III) can probably be found in many, if not most, major modern languages (e.g. De Raad and Szirmák, 1994; Goldberg and Somer, in press; Hahn, Lee, and Ashton, 1999; Ostendorf, 1990). However, it is important that readers understand that this 'Big-Three' factor structure is *not* equivalent to the dimensions of Psychoticism, Extraversion and Neuroticism proposed by Eysenck (1991) as 'basic super-factors'. In the two models, only Extraversion is the same. Psychoticism has been shown to be a blend of the orthogonal Big Five factors II and III (Goldberg and Rosolack, 1994), whereas Neuroticism (the opposite pole of Big Five Emotional Stability) is not one of the 'Big Three' lexical factors.

At the five-factor level, we can no longer expect that the exact English/German representation of these broad domains will appear with any variable set in all of the world's languages. Tests of the Big Five structure remain useful, but alternative structures should be tested as well. Moreover, there may be factors beyond the Big Five (Saucier and Goldberg, 1998).

REFINING THE LEXICAL-STUDY PARADIGM

Up to this point, we have been interpreting the findings from studies conducted within a standard framework, which we might label the 'lexical-study paradigm'. This paradigm has evolved, as most do, from the

influential template of a few initial studies. As experience has accumulated, the template has enabled researchers to learn a great deal, but some limitations of this template have also become evident. These limitations point to ways in which we might refine future lexical studies, and thereby learn even more.

First, previous studies have been prone to rely on easily obtainable data from university students, who might in fact constitute a single cosmopolitan population across national and linguistic boundaries. It is possible that such samples represent the subgroup within each language community that is most Westernized and most likely to have some facility in English, the language in which the Big Five was first discovered. Non-student, rural or even non-literate populations would provide more diversity. Moreover, previous studies have probably leaned too heavily on self-descriptions. Extensive self-examination may be an unfamiliar and possibly bizarre task in some cultural settings, based as it is on contemporary Western norms of individualism, autonomy and introspection. More attention should be given to how personality descriptors actually are used in a wide range of cultural settings.

Second, researchers should attend to the representativeness of the languages (and cultures) being sampled. Although nearly half of the world's population inhabits tropical regions, only one of the lexical studies reviewed in this chapter (Filipino) examined a language that originated in the tropics. Indeed, European languages have been heavily sampled while other geographic regions and linguistic families have remained untouched. Better sampling of languages could be obtained by pursuing two distinct strategies. In strategy (a) studies should be undertaken in non-European languages that have a very large number of speakers, such as Chinese, Bengali, Hindi and Japanese. Although useful studies of personality descriptors in Chinese (Yang and Bond, 1990) and Japanese (Isaka, 1990) have been conducted, they have not used a method comparable to those in the studies reviewed here. In strategy (b) at least one language should be studied from each of the larger language families. For example, in Africa it would be useful to study one of the Niger-Congo languages and one of the Nilo-Saharan languages. In Asia, the larger unsampled language families include Sino-Tibetan (which includes Chinese), Austro-Asiatic, Daic and Dravidian. Studies of one or more languages native to Australia, New Guinea or the Americas would also contribute to a more representative sampling.

Third, studies have largely tested a prototype Big Five structure developed originally from studies in English. Near-perfect reproduction of this structure has been characteristic mainly of emic structures from language families (i.e. Germanic, Slavic) with their origin in northern Europe. We ought to avoid ethnocentrism, with psychologists in effect imposing a

northern European folk psychology worldwide. It would be worthwhile to test other structures alongside the Big Five, such as those found in Italian or Filipino, and those based on fewer and broader factors. Fuller examination of unrotated factors might be useful; they may be as replicable as the rotated ones (Saucier, 1998).

Fourth, there is some arbitrariness in the common practice of segregating out a portion of the lexicon as referring to stable, dispositional (or 'temperament') traits, studying this portion, and ignoring other person-descriptors. Old assumptions, under which certain characteristics are heritable and others not, have been undermined by modern behaviour genetic findings (e.g. Bouchard *et al.*, 1990). These findings indicate moderate heritability for a very wide range of psychological characteristics, though not necessarily for very specific behaviours (see Chapter 4 for Borkenau *et al.*'s review of studies). Therefore, so-called 'traits' are not so separate from other characteristics as Allport (1937) and others had assumed. Moreover, as Allport (1937, Chapter 2) made clear, there are many ways to define personality, and this is one reason there is little consensus about how to distinguish between personality and other terms. Past emphasis on a narrow band of disposition terms may have been based understandably on low computing power, a belief in the importance of comprehensive selection from the lexicon, and the need to begin the search for structure in some logical place. However, the understanding of personality characteristics can only benefit from being placed in context with other characteristics of persons (e.g. states, social effects, physical appearance).

Fifth, restriction of variable selections to adjectives has been typical, but this procedure complicates extensions to other languages. As Dixon (1977) made clear, languages vary in their reliance on adjectives and some languages lack a substantial adjective class. Although person description can be called an 'adjective function' (Saucier and Goldberg, 1996b, p. 30), some of this function can be carried out with nouns. Moreover, in some languages, verbs may be used to denote person-related states and characteristics. Arguably, the goal of variable selection should be the development in each language community of a set of their most frequently used person-descriptors, regardless of whether these are found in adjective, noun or verb form. Such a set is more likely to be comparable across a wider range of languages than is a simple sampling of adjectives.

ASSUMPTIONS ABOUT PERSONALITY ATTRIBUTES

Researchers have followed Allport (1937) in tending to assume that personality descriptors are used by people mainly to denote individual differences

with respect to stable traits. But these descriptors may also function to comment on and regulate momentary behaviours, as in sentences like 'Stop being such a grouch' or 'You should be more assertive'. Consistent with this view, the findings of Chaplin, John and Goldberg (1988) suggest that most personality descriptors can function to denote either stable traits or temporary states. Therefore, it is more accurate to label these descriptors as referring to, and yielding models of, 'attributes' or 'characteristics', rather than necessarily 'traits'. Some, but not all, attributes ascribed to a person turn out to be stable traits; the trait nature of an attribute cannot be assumed *a priori*.

One purpose of lexical personality research is to understand the semantic categories that members of a culture use to comprehend one another, and thereby be able to predict others' subsequent behaviours. What is the origin of these categories? Some investigators adopt a biogenic assumption: in their view, cross-cultural regularities mean that we all share some genetically based personality hardware. An alternative is a sociogenic assumption: cross-cultural regularities indicate that various human cultures have found it useful to develop the same system for categorizing human behaviour, perhaps because of regularities in the human environment or in human social groups.

Under the biogenic assumption, the personality factors derived from lexical studies are assumed to be indicators of a biologically based personality structure. For example, McCrae and Costa's (1998) five-factor model (FFM), which was initially developed independently of the lexical Big Five (Costa and McCrae, 1980), later added two factors derived from lexical analyses (Costa and McCrae, 1985). At a purely descriptive level, the lexical Big Five and the FFM now look similar, and thus they are frequently confused. However, McCrae and Costa take a much more committed biogenic position than the typical lexical researcher, arguing for a genetic basis for their five dimensions, and citing cross-language studies as supporting this claim (McCrae and Costa, 1997). We remain sceptical of such a straightforward mapping of personality categories on to biological systems. It is quite possible that lexical personality factors represent, in part, a set of culturally derived categories into which biologically derived differences are fitted. Lexical factors cannot be assumed to reflect automatically the main lines of biological influence.

Some critics of the lexical approach have argued against studying everyday language, even as a starting point for the development of scientific theories (e.g. Block, 1995). They argue that the basic constructs in personality psychology should come from expert theories, rather than judgements by naive research participants. Our response is that such arguments assume too extreme a discrepancy between lay and expert concepts. Scientists impose categories on what they observe, just as do lay people.

Lay conceptions affect scientists, and scientific conceptions that become widespread or popular tend to percolate into the language. Therefore, it is inevitable that the scientific language of personality description will overlap substantially with the lay language. Accordingly, lexical studies can make an important contribution, helping to assure that personality models are parsimonious, yet comprehensive in content.

Clearly, the study of different lexicons can lead to a useful and highly generalizable classification system for personality traits, but this classification system should not be reified into an explanatory one. A model of descriptions does not provide a model of causes, and the study of personality lexicons should not be equated with the study of personality.

NOTES

Work on this chapter was supported by Grant MH-49227 from the National Institute of Mental Health, US Public Health Service. The authors are grateful to Michael Ashton, Veronica Benet-Martinez, Jack Block, Roy D'Andrade, Boele de Raad, Mario Forzi, Janice C. Goldberg, Martina Hřebíčková, John A. Johnson, Jeannette M. Mageo, Ivan Mervielde, Boris Mlačić, Dean Peabody, Marco Perugini, Leonard G. Rorer, Richard Shweder, Piotr Szarota, Auke Tellegen, Harry Triandis, Erika Westling and Jerry S. Wiggins for their thoughtful comments and suggestions.

Correspondence may be addressed to any of the three authors: Gerard Saucier, Department of Psychology, University of Oregon, Eugene, OR 97403, USA (e-mail: gsaucier@oregon.uoregon.edu); Sarah E. Hampson, Department of Psychology, University of Surrey, Guildford GU2 5XH, UK (e-mail: s.hampson@surrey.ac.uk); or Lewis R. Goldberg, Oregon Research Institute, 1715 Franklin Boulevard, Eugene, OR 97403-1983, USA (e-mail: lewg@ori.org).

- 1 We use the phrase 'Big Five' to refer to the common elements in the five-factor representations from lexical studies in English (e.g. Saucier and Goldberg, 1996a) and in German (Ostendorf, 1990). This practice should not be understood as an ethnocentric reification of the English/German structure, but only as a short-hand convention.
- 2 This corollary of the lexical hypothesis has important implications for the selection of variables in lexical studies. Because the most important attributes are likely to spawn a large number of synonyms (or quasi-synonyms), it defeats the goal of representative sampling to cull such redundancies from the final set of terms to be studied. Unfortunately, as we will indicate later in this chapter, the findings from a number of studies may have suffered as a result of such synonym deletion. On the other hand, just how many synonyms an attribute (e.g. Kind) might generate is somewhat indeterminate: Lexicons are open enough that new (but low-frequency) terms can be readily invented, especially using metaphors (e.g. Open-hearted) or nominalizations (e.g. a Mother Theresa) (see Henss, 1995).
- 3 The inclusion of the phrase 'tends to be' and the adverb 'often' may have served to bias the selection against attributes that are relatively non-fluctuating, such as those referring to physical attractiveness. If this hypothesis is correct, it might explain some of the differences between the 'Positive

Valence' factor found in this study and the 'Attractiveness' factor found by Saucier (1997).

- 4 In raw-data analyses in German, an Emotional Stability factor did not appear until at least six factors were extracted and rotated; in the raw-data five-factor solutions, Agreeableness was split into two factors, labelled SD (low social desirability) and 'Warmherzigkeit' (Warm-heartedness) (Ostendorf, 1990, p. 178).
- 5 A Norwegian lexical project has been carried out over the past few years, but reports of these findings have not yet been published. Engvik (1999, personal communication) has reported that a clear Big Five factor structure was easily identifiable in analyses of 859 personality-trait adjectives. Indeed, it appears that even the Norwegian factor V was nearly identical to the English/German version of the Intellect factor.
- 6 Di Blas and Forzi (1999) analysed raw data as well as ipsatized data. The raw-data solutions are described, in a footnote, as having more unipolar factors but otherwise similar results.
- 7 When raw data instead of ipsatized ratings were analysed in this study of Turkish descriptors, a Big Five structure could not be identified, due to the fusion of aspects of Extraversion and Intellect into a single factor (probably related to Dynamism, Self-assurance, Assertiveness, and Positive Valence factors found in other lexical studies) beginning at the two-factor level. The raw-data five-factor solution included this Dynamism factor, plus factors that might be labelled Social propriety, Unpleasant affect, Conscientiousness, and Emotionality; the two-factor solution included the first two of these, and the three-factor solution the first three. Thus the raw-data analyses in the second study supported neither the Big Five nor the Big Three, but did find a two-factor structure resembling that in analyses of descriptors from other languages (e.g. Saucier, 1998).
- 8 There has been some controversy with respect to the advisability of ipsatizing. Ten Berge (1999) refuted some of the critiques made of that procedure, but noted that ipsatizing removes individual differences in mean item response, and 'to the extent that the mean reflects non-spurious differences between individuals' (p. 101), any removal of this component of responses leads to a loss of useful information. Ten Berge (1999) noted that when a personality scale is quite unbalanced with respect to positively and negatively keyed items, 'the sum of scores on the items represents a mixture of the trait under consideration and acquiescence, and ipsatizing the data will indeed remove the most important source of variation, an error of strategy' (pp. 93-94).
In other words, ipsatization makes sense under the assumption that all individual differences in response mean and standard deviation are spurious (i.e. due to individual differences in acquiescence and extremeness response tendencies). But this assumption is most tenable when each of the attributes under consideration is represented by an equal number of positively and negatively worded items, a situation that cannot be expected to hold in most variable selections based on natural language descriptors. As Hofstee (1994) argued, removal of acquiescence variance might be accomplished with more precision by partialling the response mean computed across a balanced (sub)set of variables (e.g. a set of antonym pairs).
In order to gain more clarity on the effect of these procedures, it would be useful if investigators reported their findings from both ipsatized and non-ipsatized data (as in Goldberg, 1990; Ostendorf, 1990; Saucier, 1998), as well

as from residual data after removing the response means from a balanced subset of variables. As indicated in this chapter, most lexical analyses have been conducted with ipsatized data. Exceptions are (a) the Big Seven studies in English (Tellegen and Waller, 1987) and Hebrew (Almagor, Tellegen and Waller, 1995) in which only raw data were apparently used, and (b) some studies (Di Blas and Forzi, 1999; Goldberg and Somer, 1999; Ostendorf, 1990; Saucier, 1998) that have analysed both raw and ipsatized data. The few comparisons of the two kinds of data suggest that in analyses of raw data the Big Five is weaker, whereas the Big Three and higher-order solutions (one or two factors) are not weakened substantially.

- 9 It is interesting that the third unrotated factor found in English, German and some other languages, which is orthogonal to these first two, has been labelled Tight-Loose (impulse control versus impulse expression). Triandis and his colleagues (1988) have discussed another cultural syndrome which they also labelled Tight-Loose.

REFERENCES

- Allport, G.W. (1937) *Personality: A Psychological Interpretation*. New York: Holt, Rinehart and Winston.
- Allport, G.W. and Odbert, H.S. (1936) 'Trait names: A psycho-lexical study', *Psychological Monographs* 47(1, Whole No. 211).
- Almagor, M., Tellegen, A. and 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. and John, O.P. (1990) 'Towards a taxonomy of personality descriptors in German: A psycho-lexical study', *European Journal of Personality* 4: 89-118.
- Baumgarten, F. (1933) 'Die Charaktereigenschaften' ['The character traits'], *Beiträge zur Charakter- und Persönlichkeitsforschung* (Whole No. 1). Bern: A. Francke.
- Benet-Martínez, V. and John, O.P. (1998) 'Los Cinco Grandes across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English', *Journal of Personality and Social Psychology* 75: 729-750.
- Benet-Martínez, V. and 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.
- Berry, J.W. (1969) 'On cross-cultural comparability', *International Journal of Psychology* 4: 119-128.
- Block, J. (1995) 'A contrarian view of the Five-Factor approach to personality description', *Psychological Bulletin* 117: 187-215.
- Bouchard, T.J., Lykken, D.T., McGue, M., Segal, N.L. and Tellegen, A. (1990) 'Sources of human psychological differences: The Minnesota study of twins reared apart', *Science* 250: 223-228.
- Brokken, F.B. (1978) *The Language of Personality*. Meppel, The Netherlands: Krips.
- Caprara, G.V. and Perugini, M. (1994) 'Personality described by adjectives: The generalizability of the Big Five to the Italian lexical context', *European Journal of Personality* 8: 357-369.
- Cattell, R. B. (1943) 'The description of personality. II. Basic traits resolved into clusters', *Journal of Abnormal and Social Psychology* 38: 476-507.
- Chaplin, W.F., John, O.P. and Goldberg, L.R. (1988) 'Conceptions of states and traits: Dimensional attributes with ideals as prototypes', *Journal of Personality and Social Psychology* 54: 541-557.
- Church, A.T., Katigbak, M.S. and Reyes, J.A.S. (1996) 'Toward a taxonomy of trait adjectives in Filipino: Comparing personality lexicons across cultures', *European Journal of Personality* 10: 3-24.
- Church, A.T., Katigbak, M.S. and Reyes, J.A.S. (1998) 'Further exploration of Filipino personality structure using the lexical approach: Do the Big Five or Big Seven dimensions emerge?', *European Journal of Personality* 12: 249-269.
- Church, A.T., Reyes, J.A.S., Katigbak, M.S. and Grimm, S.D. (1997) 'Filipino personality structure and the Big Five model: A lexical approach', *Journal of Personality* 65: 477-528.
- Costa, P.T., Jr and McCrae, R.R. (1980) 'Still stable after all these years: Personality as a key to some issues in adulthood and old age', in P. Baltes and O.G. Brim (eds) *Life Span Development and Behavior*, vol. 3, pp. 31-54. New York: Academic Press.
- Costa, P.T. Jr and McCrae, R.R. (1985) *The NEO Personality Inventory Manual*. Odessa, FL: Psychological Assessment Resources.
- De Raad, B. (1992) 'The replicability of the Big Five personality dimensions in three word-classes of the Dutch language', *European Journal of Personality* 6: 15-29.
- De Raad, B. (1994) 'An expedition in search of a fifth universal factor: Key issues in the lexical approach', *European Journal of Personality* 8: 229-250.
- De Raad, B. (1998) 'Five big Big Five issues: Rationale, content, structure, status and cross-cultural assessment', *European Psychologist* 3: 113-124.
- De Raad, B., Di Blas, L. and Perugini, M. (1997) 'Two independent Italian trait taxonomies: Comparisons with Italian and between Italian and Germanic languages', *European Journal of Personality* 11: 167-185.
- De Raad, B., Hendriks, A.A.J. and Hofstee, W.K.B. (1992) 'Towards a refined structure of personality traits', *European Journal of Personality* 6: 301-319.
- De Raad, B. and Hoskens, M. (1990) 'Personality-descriptive nouns', *European Journal of Personality* 4: 131-146.
- De Raad, B., Mulder, E., Kloosterman, K. and Hofstee, W.K.B. (1988) 'Personality-descriptive verbs', *European Journal of Personality* 2: 81-96.
- De Raad, B. and Ostendorf, F. (1996) 'Quantity and quality of trait-descriptive type nouns', *European Journal of Personality* 10: 45-56.
- De Raad, B., Perugini, M. and Szirmák, Z. (1997) 'In pursuit of a cross-lingual reference structure of personality traits: Comparisons among five languages', *European Journal of Personality* 11: 167-185.
- De Raad, B., Perugini, M., Hřebíčková, M. and Szarota, P. (1998) 'Lingua Franca of personality: Taxonomies and structures based on the psycholexical approach', *Journal of Cross Cultural Psychology* 29: 212-232.

- De Raad, B. and Szirmák, Z. (1994) 'The search for the "Big Five" in a non-Indo-European language: The Hungarian trait structure and its relationship to the EPQ and the PTS', *European Review of Applied Psychology* 44: 17-26.
- Di Blas, L. and 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. and 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.
- Digman, J.M. (1990) 'Personality structure: Emergence of the five-factor model', in M.R. Rosenzweig and L.W. Porter (eds) *Annual Review of Psychology*, vol. 41, pp. 417-446. Palo Alto, CA: Annual Reviews.
- Digman, J.M. (1997) 'Higher order factors of the Big Five', *Journal of Personality and Social Psychology* 73: 1246-1256.
- Dixon, R.M.W. (1977) 'Where have all the adjectives gone?', *Studies in Language* 1: 19-80.
- Eysenck, H.J. (1991) 'Dimensions of personality: 16, 5, or 3? Criteria for a taxonomic paradigm', *Personality and Individual Differences* 12: 773-790.
- Galton, F. (1884) 'Measurement of character', *Fortnightly Review* 36: 179-185.
- Goldberg, L.R. (1976) 'Language and personality: Toward a taxonomy of trait-descriptive terms', *Istanbul Studies in Experimental Psychology* 12: 1-23.
- 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. (1982) 'From Ace to Zombie: Some explorations in the language of personality', in C.D. Spielberger and J.N. Butcher (eds) *Advances in Personality Assessment*, vol. 1, pp. 203-234. Hillsdale, NJ: Erlbaum.
- 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. (1993) 'The structure of phenotypic personality traits', *American Psychologist* 48: 26-34.
- Goldberg, L.R. (1995) 'What the hell took so long? Donald Fiske and the Big-Five factor structure', in P.E. Shrout and S.T. Fiske (eds) *Personality Research, Methods, and Theory: A Festschrift Honoring Donald W. Fiske*, pp. 29-43. Hillsdale, NJ: Erlbaum.
- Goldberg, L.R. and Rosolack, T.K. (1994) 'The Big Five factor structure as an integrative framework: An empirical comparison with Eysenck's P-E-N model', in C.F. Halverson, Jr, G.A. Kohnstamm and R.P. Martin (eds) *The Developing Structure of Temperament and Personality from Infancy to Adulthood*, pp. 7-35. New York: Erlbaum.
- Goldberg, L.R. and Saucier, G. (1995) 'So what do you propose we use instead? A reply to Block', *Psychological Bulletin* 117: 221-225.
- Goldberg, L.R. and Somer, O. (in press) 'The hierarchical structure of common Turkish person-descriptive adjectives', *European Journal of Personality*.
- Hahn D.-W., Lee, K. and Ashton, M.C. (1999) 'A factor analysis of the most frequently used Korean personality trait adjectives', *European Journal of Personality* 13: 261-282.
- Henss, R. (1995) 'From Aal to Zyniker: Personality descriptive type nouns in the German language', *European Journal of Personality* 9: 135-145.
- Hofer, S.M. (1999) 'Assessing personality structure using factorial invariance procedures', in I. Mervielde, I.J. Deary, F. De Fruyt and F. Ostendorf (eds) *Personality Psychology in Europe*, vol. 7, pp. 35-49. Tilburg, Netherlands: Tilburg University Press.
- Hofstee, W.K.B. (1994) 'Personality and factor analysis: Bind or bond', *Zeitschrift für Differentielle und Diagnostische Psychologie* 15: 173-183.
- Hofstee, W.K.B., Brokken, F.B. and Land, H. (1981) 'Constructie van een standaardpersoonlijke-eigenschappenlijst (SPEL)' ['Construction of a standard list of personality descriptive adjectives'], *Nederlands Tijdschrift voor de Psychologie* 34: 443-452.
- Hofstee, W.K.B., Kiers, H.A.L., De Raad, B., Goldberg, L.R. and Ostendorf, F. (1997) 'A comparison of Big-Five structures of personality traits in Dutch, English and German', *European Journal of Personality* 11: 15-31.
- Horn, J.L. (1991) 'Discussion of issues in factorial invariance', in L.M. Collins and J.L. Horn (eds) *Best Methods for the Analysis of Change*, pp. 114-125. Washington, DC: American Psychological Association.
- Hřebíčková, M., Ostendorf, F. and Angleitner, A. (1995) '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, July.
- Hřebíčková, M., Ostendorf, F., Osecká, L. and Čermák, I. (1999) 'Taxonomy and structure of Czech personality-relevant verbs', in I. Mervielde, I.J. Deary, F. De Fruyt and F. Ostendorf (eds) *Personality Psychology in Europe*, vol. 7, pp. 51-65. Tilburg, The Netherlands: Tilburg University Press.
- Isaka, H. (1990) 'Factor analysis of trait terms in everyday Japanese language', *Personality and Individual Differences* 11: 115-124.
- John, O.P. (1990) 'The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires', in L.A. Pervin (ed.) *Handbook of Personality: Theory and Research*, pp. 66-100. New York: Guilford.
- John, O.P. and Srivastava, S. (1999) 'The Big Five trait taxonomy: History, measurement and theoretical perspectives', in L.A. Pervin and O.P. John (eds) *Handbook of Personality: Theory and Research*, 2nd edn, pp. 102-138. New York: Guilford.
- Klages, L. (1932) *The Science of Character* (translated 1932). London: George Allen and Unwin. (Original work published 1926).
- McCrae, R.R. and Costa, P.T., Jr (1997) 'Personality trait structure as a human universal', *American Psychologist* 52: 509-516.
- McCrae, R.R. and Costa, P.T., Jr (1998) 'Toward a new generation of personality theories: Theoretical contexts for the Five-Factor Model', in J.S. Wiggins (ed.) *The Five-Factor Model of Personality: Theoretical Perspectives*, pp. 51-87. New York: Guilford.
- McCrae, R.R. and John, O.P. (1992) 'An introduction to the five factor model and its applications', *Journal of Personality* 60: 175-215.

- Meredith, W. (1993) 'Measurement invariance, factor analysis and factorial invariance', *Psychometrika* 58: 525-543.
- Mlačić, B. (1999) 'Hrvatska taksonomija deskriptora osobina ličnosti' ['Croatian taxonomy of personality descriptive terms'], unpublished doctoral dissertation, University of Zagreb, Croatia.
- 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.
- Norman, W.T. (1967) *2800 Personality Trait Descriptors: Normative Operating Characteristics for a University Population*. University of Michigan; Ann Arbor, MI: Department of Psychology.
- Osgood, C.E., May, W. and Miron, M. (1975) *Cross-Cultural Universals of Affective Meaning*. Urbana: University of Illinois Press.
- Ostendorf, F. (1990) *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit [Language and Personality Structure: Toward the Validation of the Five-Factor Model of Personality]*. Regensburg, Germany: S. Roderer Verlag.
- Ostendorf, F. and Angleitner, A. (1994) Enthusiasts contra pessimists, *Psychological Inquiry* 5: 159-162.
- Paulhus, D.L. and John, O.P. (1998) 'Egoistic and moralistic biases in self-perception: The interplay of self-descriptive styles with basic traits and motives', *Journal of Personality* 66: 1025-1060.
- Peabody, D. (1987) 'Selecting representative trait adjectives', *Journal of Personality and Social Psychology* 52: 59-71.
- Peabody, D. and Goldberg, L.R. (1989) 'Some determinants of factor structures from personality-trait descriptors', *Journal of Personality and Social Psychology* 57: 552-567.
- Saucier, G. (1997) 'Effects of variable selection on the factor structure of person descriptors', *Journal of Personality and Social Psychology* 73: 1296-1312.
- Saucier, G. (1998) 'A hierarchy of personality factors in the natural language', paper presented at the 9th European Conference on Personality, Guildford, England, July.
- Saucier, G. and Goldberg, L.R. (1996a) 'Evidence for the Big Five in analyses of familiar English personality adjectives', *European Journal of Personality* 10: 61-77.
- Saucier, G. and Goldberg, L.R. (1996b) 'The language of personality: Lexical perspectives on the five-factor model', in J.S. Wiggins (ed.) *The Five-Factor Model of Personality: Theoretical Perspectives*, pp. 21-50. New York: Guilford.
- Saucier, G. and Goldberg, L.R. (1998) 'What is beyond the Big Five?', *Journal of Personality* 66: 495-524.
- Saucier, G. and Ostendorf, F. (1999) 'Hierarchical components of the Big Five personality factors: A cross-language replication', *Journal of Personality and Social Psychology* 76: 613-627.
- Shmelyov, A.G. and Pokhil'ko, V.I. (1993) 'A taxonomy-oriented study of Russian personality-trait names', *European Journal of Personality* 7: 1-17.
- Shmelyov, A.G., Pokhil'ko, V.I. and Kozlovskaya-Telnova, A.Y. (1988) *Praktikum po experimental'noi psichosemantike: Tezaurus lichnostnikh [Trait Project on Experimental Psychosemantics: Trait-names Thesaurus]*, Moscow: Moscow University Press.
- Shmelyov, A.G., Pokhil'ko, V.I. and Kozlovskaya-Telnova, A.Y. (1991) 'Reprezentativnost' lichnostnikh chert v soznanii nositelya russkogo yazika' ['Representation of personality traits in the consciousness of a Russian-language speaker'], *Psichologicheskii Zhurnal* 2: 27-44.
- Shweder, R.A. (1972) 'Semantic structure and personality assessment', unpublished doctoral dissertation, Harvard University.
- Slotboom, A.-M., Havill, V.L., Pavlopoulos, V. and De Fruyt, F. (1998) 'Developmental changes in personality descriptions of children: A cross-national comparison of parental descriptions of children', in G.A. Kohnstamm, C.F. Halverson, I. Mervielde and V.L. Havill (eds) *Parental Descriptions of Child Personality*, pp. 127-153. Mahwah, NJ: Lawrence Erlbaum.
- Somer, O. and Goldberg, L.R. (1999) 'The structure of Turkish trait-descriptive adjectives', *Journal of Personality and Social Psychology* 76: 431-450.
- Sorbom, D. (1974) 'A general method for studying differences in factor means and factor structures between groups', *British Journal of Mathematical and Statistical Psychology* 28: 229-239.
- Szarota, P. (1996) 'Taxonomy of the Polish personality-descriptive adjectives of the highest frequency of use', *Polish Psychological Bulletin* 27: 342-351.
- Szarota, P. and Ostendorf, F. (1997) 'Taxonomy and structure of Polish personality-relevant adjectives', unpublished manuscript.
- Szirmák, Z. and De Raad, B. (1994) 'Taxonomy and structure of Hungarian personality traits', *European Journal of Personality* 8: 95-118.
- Tellegen, A. and 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, NY, August.
- Tellegen, A., Grove, W.M. and Waller, N.G. (1991) 'Inventory of personality characteristics No. 7 (IPC7)', unpublished manuscript, University of Minnesota, Department of Psychology, Minneapolis, MN.
- 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.
- Triandis, H.C., Bontempo, R., Villareal, M.J., Asai, M. and Lucca, N. (1988) 'Individualism and collectivism: Cross-cultural perspectives on self-in-group relationships', *Journal of Personality and Social Psychology* 54: 323-338.
- Victor, J.B., Dent, H.E., Carter, B., Halverson, C.F. and Havill, V.L. (1998) 'How African-American parents describe their children', in G.A. Kohnstamm, C.F. Halverson, I. Mervielde and V.L. Havill (eds) *Parental Descriptions of Child Personality*, pp. 169-187. Mahwah, NJ: Erlbaum.
- Waller, N.G. (in press) 'Evaluating the structure of personality', in C.R. Cloninger (ed.) *Personality and Psychopathology*. Washington, DC: American Psychiatric Press.
- Waller, N.G. and Zavala, J.D. (1993) 'Evaluating the Big Five', *Psychological Inquiry* 4: 131-134.

- White, G.M. (1980) 'Conceptual universals in interpersonal language', *American Anthropologist* 82: 759-781.
- Wiggins, J.S. and Trapnell, P.D. (1997) 'Personality structure: The return of the Big Five', in R. Hogan, J. Johnson and S. Briggs (eds) *Handbook of Personality Psychology*, pp. 737-764. San Diego, CA: Academic Press.
- Yang, K. and Bond, M.H. (1990) 'Exploring implicit personality theories with indigenous or imported constructs: The Chinese case', *Journal of Personality and Social Psychology* 58: 1987-1095.