Recurrent Personality Dimensions in Inclusive Lexical Studies:

Indications for a Big Six Structure

Gerard Saucier

University of Oregon

Running Head: Personality Structure

Correspondence should be addressed to:

Gerard Saucier
Department of Psychology
1227 University of Oregon
Eugene OR 97403-1227 USA
e-mail: gsaucier@uoregon.edu
tel.: 541-346-4927
fax: 541-346-4911

September 2008

Abstract

Previous evidence for both the Big Five and alternative six-factor model has been drawn from lexical studies with relatively narrow selections of attributes. This study examined factors from previous lexical studies using a wider selection of attributes, in seven languages (Chinese, English, Filipino, Greek, Hebrew, Spanish, and Turkish), and found six recurrent factors, each with common conceptual content across most of the studies. The previous narrow-selection-based six-factor model outperformed the Big Five in capturing the content of the six recurrent wide-band factors. Adjective markers of the six recurrent wide-band factors showed substantial incremental prediction of important criterion variables, over and above the Big Five. Correspondence between wide-band six and narrow-band six factors indicate they are variants of a 'Big Six' model that is more general across variable-selection procedures and may be more general across languages and populations.

At present, the favored scientific structural representation for personality attributes is the Big Five. The Big Five model (and the related Five Factor Model [FFM]; McCrae & Costa, 1985) posits that inter-individual personality variation tends to fall into five independent dimensions, labeled as Extraversion, Agreeableness, Conscientiousness, Emotional Stability (versus Neuroticism as in the FFM), and Intellect (in the Big Five) or Openness (in the FFM). Measures of the Big Five are frequently used in many types of psychological research. Publications that helped establish the Big Five paradigm (e.g., Digman, 1990; Goldberg, 1990; John, 1990) have been at times among the most cited in psychology (Pendlebury, 1996). However, evidence in this report indicates a strong case for an alternative to the Big Five.

Recurrent Structures of Personality Attributes in Lexical Studies

Norman (1963) and Goldberg (1981) proposed that the Big Five model be tested in studies of the natural language -- lexical studies. Lexical studies draw on a strong variable-selection rationale: Because of the social importance of personality attributes, the most crucial attributes tend to become sedimented in human lexicons. A lexical study extracts from a dictionary a relatively representative and comprehensive set of frequently-used natural-language personality descriptors, administers them in self- or acquaintance-report format, and examines the factor structure generated by the descriptors' intercorrelations. Each language is studied separately, but factors arising convergently in diverse studies are important because they indicate psychological patterns that transcend cultural and linguistic boundaries.[1] Lexical studies have now examined over a dozen languages (for a review, see Saucier & Goldberg, 2003).

The earliest lexical studies (De Raad, Hendriks, & Hofstee, 1992; Goldberg, 1990; Ostendorf, 1990) appeared to confirm Big-Five-based expectations: Five factors based on correlations among hundreds of personality adjectives from Dutch or English or German could be interpreted as Big Five variants. But as lexical studies began to appear in non-Germanic languages, anomalous findings began to appear. First, Szirmák and De Raad's (1994) study of Hungarian personality descriptors failed to find the Big Five where expected in a five-factor solution, and

pointed to an alternative six-factor structure. Subsequent studies of Italian (DiBlas & Forzi, 1997) and French (Boies et al., 2001) led to similar six-factor structures. Many lexical studies that confirmed the Big-Five have yielded this six-factor alternative, if an additional factor is extracted.

In a landmark synthesis, Ashton et al. (2004) showed that six-factor solutions from seven languages (not just Hungarian, Italian, and French, but also Dutch, German, Polish, and Korean) contained, content-wise, a fairly consistent set of factors. Ashton et al. provided a useful interpretive summary of the factors. Table 1 documents the degree of recurrence of specific personality concepts on each factor, tabulated directly from tables in Ashton et al. (2004).

How closely are these *six factors* related to the Big Five? *Conscientiousness, Openness,* and *Extraversion* have content closely resembling that in Big Five Conscientiousness, Intellect, and Extraversion (cf., Goldberg, 1992). The *Agreeableness* content resembles that in both Big Five Agreeableness and Emotional Stability. *Emotionality* content resembles that in Neuroticism, although Emotionality has a better defined favorable pole (e.g., Courage, Self-Assurance) and more content related to sentimentality. *Honesty/Humility* content relates moderately to Big Five Agreeableness, but much of its content stands outside the Big Five. With its additional factor, this six-factor model is neither reducible to nor entirely reproducible from the Big Five.

De Raad et al. (2008) compared six-factor structures in 13 lexical studies, including those synthesized by Ashton et al. (2004). The study examined all possible pairings of the 13 structures, comparisons based on that subset of variables in each paired study that were judged to have corresponding meaning. A limitation of this approach is that typically only about 30% of the variables in one study corresponded with those in another, and inferences are not straightforward because they must be drawn based on a large number of pairwise comparisons. Nonetheless, in De Raad et al. (2008) convergence across studies, as indicated by mean coefficients of factor congruence, was discernible though moderate. Across all pairings, the mean congruence between matched factors in a pair of studies (factors from one study rotated to a target matrix established by the structure from the other study) was .75. The same mean congruence of roughly .75 has been

found in similar studies examining the replicability of five-factor structures (De Raad et al., 1998; Hofstee et al., 1997). Thus, six-factor structures appear about as robust as the Big Five, while providing one added source of variance beyond the Big Five.

These factors can be called the Cross-Language Six (e.g., Saucier, 2008), but perhaps more accurately the narrow-band Cross-Language Six (NCL6), because they come out of studies with highly restricted selections of variables. That is, the synthesis of Ashton et al. (2004) omitted a sizeable group of lexical studies: those that used a more inclusive variable-selection strategy.

Evidence from Wide-band Lexical Studies

Lexical studies to date fall into two groups based on inclusiveness of variable selection. Influenced by early work of Allport and Odbert (1936), many lexical studies – including all those cited above – have excluded terms judged to be highly evaluative (e.g., Stupid, Wicked, Outstanding) or to refer to emotional states (e.g., Sad, Angry, Frustrated) or to the characteristic effects an individual has on others (e.g., Charming, Dangerous, Disgusting) (Saucier, 1997; Tellegen, 1993). The rationale for exclusion of these categories: Personality dispositions are indicated only by terms that primarily describe an "enduring pattern of typical behavioral tendency," that is, tendencies in action, thought, and emotions (Ashton & Lee, 2005. p. 8). There is concern, for example, that highly evaluative terms (such as Awful, Wicked, Outstanding, and Impressive) might produce factors reflecting patterns of response to items with unusual extremeness in desirability rather than any meaningful substantive, descriptive content.

Although concepts within such excluded categories may fall outside a prototypical core set of "enduring patterns of typical behavioral tendency," they do fit easily within typical definitions of personality. They are all qualities discernible in individuals' behavior, thinking, and affect. There is no need for researchers to put on blinders, excluding such descriptors from view, when a sober appraisal is yet to be made of what validity they might potentially have. Studies including such descriptors may yield a larger number of predictively useful dimensions (e.g., Simms, 2007).

Consistent with such a rationale, eight lexical studies to date have had more inclusive

variable selections: two studies in English (Saucier, 1997; Tellegen & Waller, 1987), as well as one each in Filipino (Church et al., 1997), Greek (Saucier et al., 2005), Hebrew (Almagor et al., 1995), Spanish (Benet-Martínez & Waller, 1998), Turkish (Goldberg & Somer, 2000), and most recently Chinese (Zhou et al., in press). None of these studies replicated the Big Five, but the replicability of a six-factor model has not been thoroughly examined across these studies.

The earliest of these "wide-band" lexical studies (Tellegen & Waller, 1987) produced what was labeled a Big Seven structure, consisting of four factors closely resembling those from the Big Five (Extraversion [Positive Emotionality], Agreeableness [Agreeability], Conscientiousness [Dependability], and Emotional Stability [Negative Emotionality]), one with a more distant resemblance (Conventionality, related to the low end of Intellect/Imagination), and two represented as new factors drawing on highly evaluative descriptors (Positive Valence and Negative Valence). Attempts to replicate these results, using the same page-sampling methodology as in the original study, followed in independent lexical studies in Spanish (Benet-Martinez & Waller, 1998) and Hebrew (Almagor, Tellegen, & Waller, 1995). Results were partially supportive – the studies did all find factors that could be readily labeled Agreeableness, Conscientiousness, Negative Valence, and Positive Valence.

After examining considerable convergences between seven-factor structures in the Hebrew study and a later Filipino study (Church et al., 1997), Saucier (2003a) proposed to test an alternative "Multi-Language Seven" (ML7) model. This structure tends to resemble the narrow-band Cross-Language Six, except that Agreeableness is re-labeled as Even Temper, Extraversion as Gregariousness, Openness as Intellect, Emotionality as (by its opposite pole) Self-Assurance, and in place of the Honesty factor one finds two factors: Concern for Others and Negative Valence. The ML7 model has been tested in competition with the Big Five and narrow-band Cross-Language Six in several lexical studies (Saucier, 2003b; Saucier et al., 2005; Zhou et al., in press), typically appearing roughly as replicable as either of the other models, so long as variable selection is inclusive. There is no evidence, however, that either the Big Seven or

the ML7 appears under narrower, more restrictive variable-selection conditions.

Against this background, the present studies address two important questions. First, do these eight inclusive-selection lexical studies support structures proposed in previous research? Second, do these studies converge on a consistent structural pattern, and if there is such a pattern, how does it differ from prior structural models?

Study 1

Method

Each lexical study uses a distinct set of variables – an optimally representative set of descriptors from the language under study – as well as a distinct set of subjects. These features make a lexical study well-suited to identifying the structure indigenous to a language and population. But these features also make exact comparison of structures challenging, because typical quantitative indices of factor similarity require either a common set of subjects (for correlations between factor scores) or a common set of variables (for congruence coefficients). One might rely judgments of similarity made by raters (experts or trained subjects), but these have a subjective aspect that is difficult to remove. Partly because of the abstract, demanding nature of tasks where one judges how much groups of terms resemble each other, cross-rater agreements between such judgments could be biased by shared consensual schemas.

Fortunately, much can be accomplished by a straightforward tabulation-and-comparison approach that is highly objective. Ashton et al. (2004) were able to demonstrate a recurrent six-factor structure by listing the terms (translated into English by previous authors) having their highest association with each of six factors in each of eight studies and, comparing lists, drawing conclusions about content held in common. Table 1 identifies the most recurrent terms in those eight studies. The present study extends the general approach of Ashton et al. (2004) to a different set of studies.

Examination of inclusive-variable-selection studies would be expected to yield less convergence than did the narrow variable-selection studies examined by Ashton et al. (2004), for

several reasons. First, Ashton et al. (2004) included only studies already observed to show the six-factor structural pattern, and studies that deviated from this pattern were simply omitted. In the present analysis no studies were *a priori* excluded as too deviant. Second, inclusive-selection studies have varied considerably in their methodology: Three studies included here extracted terms from a dictionary by sampling some but not most dictionary pages and therefore included some relatively unfamiliar terms. Third, a more inclusive selection gives a larger universe of concepts from which to sample, reducing the likelihood of finding common terms across studies and languages. And fourth, the eight studies examined here include four non-European languages, involving more geographic diversity than in Ashton et al. (2004), which examined just one such language. With a wider diversity of cultural settings, one should expect less convergence.

For five of the eight studies, the tables of factor loadings presented in the original research reports were examined and relied upon (Almagor et al., 1995, Table 1; Benet-Martínez & Waller, 1997, Table 1; Church et al., 1997, Appendix B; Goldberg & Somer, 2000, Table 1; Zhou et al., in press, Table 1). The Greek study (Saucier et al., 2005) had reported a six-factor solution; an unpublished table of the seven-factor solution was used. Sample sizes were substantial: 1531 for Filipino, 991 for Greek, 894 for Spanish, 892 for Chinese, 637 for Hebrew, and 631 for Turkish.

The two other studies involved English. The unpublished English lexical study of Tellegen and Waller (1987) had no comparable table, so we relied instead upon the factor loadings of terms selected to reference its structure in an English-language sample (N=565) reported by Benet and Waller (1995, Table 1). Carried out in Minnesota, this study is labeled "English-MN." The English lexical study of Saucier (1997, Table 7; N=700) reported an analysis with an unusually inclusive variable selection, numerous terms for physical attractiveness being incorporated among the 500 terms.[2] Saucier (1997) described analyses using a nearly as inclusive selection of 455 "non-physical" descriptors which leaves out the most overt attractiveness terms (as well as other physical descriptors like Tall, Tiny, and Fat). I relied here on the seven-factor solution from this non-physical selection. Carried out in Oregon, this study is labeled "English-OR."

All studies used self-descriptions. The Chinese data included 485 participants describing a well-acquainted peer as well as 417 describing self, with attention focused on the structure these two subsamples converged upon, defined by average loadings across self and peer structures.

For each factor in each study, the terms with their highest loading on a factor were taken as representatives of that factor.[3] Published tables differed in how many terms were provided for each factor; at least 8 terms were provided for each factor from one study (Tellegen & Waller, 1987; in Benet & Waller, 1995), at least 9 in two studies (Almagor et al., 1995; Saucier et al., 2005), and at least 11 in the five others. On the high end, as many as 18 terms for some factors were provided in all studies except one English study (Tellegen & Waller, 1987; Benet & Waller, 1995) which had a maximum of 11. For exploratory analyses seeking the pattern on which the eight studies converged, it was useful to equalize the contributions of the studies to the final outcome, so a limit of 10 terms per factor was imposed; such a limit beneficially served to focus attention only on the core prototypical content of each factor. For testing expectations based on prior structures, equalizing the contribution of the eight studies was less important, whereas power to detect representatives of previous structures was more important; in this case the limit was set instead at 18 terms per factor.

Analyses Testing Prior Expectations

The comparative replication of four etic structures was tested. The Big Five was represented by the 100 marker terms provided by Goldberg (1992). The narrow-band Cross-Language Six was represented by the 78 terms in Table 1 that appeared on a given factor in at least three studies. The Multi-Language Seven was represented by the 60 marker adjectives provided by Saucier (2003a); because the Filipino and Hebrew results served as a basis for derivation of this model, those two studies are the least informative from a replication standpoint. The Big Seven was represented by the 70 terms used as markers by Benet and Waller (1995); obviously all 70 appeared on the proper factor in that study, so that study is uninformative with regard to replication of the Big Seven.

If each model is correct, what should occur in these seven-factor solutions drawn from eight studies? Expectations are straightforward for the Big Seven and Multi-Language Seven: Many

marker terms should appear from each factor (hereafter called "construct") in the model, and markers for each construct should appear on their own separate factor. Thus the Big Seven or ML7 scales should correspond perfectly to the seven factors in each study. For the Big Five, understood as a stringent model, we would expect in seven-factor solutions the Big Five plus two orthogonal (perhaps meaningless) additional factors, so that the five sets of markers should correspond to five of the seven factors, and not at all to the other two. Applying the Big Five model more loosely, the two additional factors could be subdivisions of single Big Five constructs. Parallel expectations would hold for the narrow-band Cross-Language Six: stringently, that each seven-factor solution would contain these six plus one factor unrelated to the six, or more loosely, that each solution would have five of the six constructs, plus one of the six constructs divided in two.

To evaluate these hypotheses, there are two key indicators. The first involves construct representation: the number of the model's constructs whose markers fail to appear anywhere on the seven factors -- such failure to appear suggests that the construct is not present. The second involves isomorphism: Among the marker terms appearing on any of the seven factors, how many of them appear on the wrong factor, that is, can be scored as a "miss" rather than a "hit"? A 'hit' occurs when a marker term (e.g., Bold for Extraversion) appears on the factor that is the most common attractor of that set of marker terms (e.g., of all of the Big Five Extraversion markers), making that factor interpretable in terms of the construct (e.g., as Extraversion). A 'miss' occurs when a marker term appears on a different factor than the most common attractor – thus, the wrong factor. This second indicator was used explicitly (and the first implicitly) by Goldberg (1992) in evaluating the replication of the Big Five factors as produced by a classic marker set of 100 terms.

A perfectly replicating model should have all of its constructs represented in the model derived from new data – the percentage of missing constructs should be zero. Moreover, in a perfectly replicating model, the marker-terms that operationalize those constructs should appear neatly on separate factors, with no marker-terms appearing on the wrong factor and thus no marker-terms scoreable as a miss rather than a hit. That is, a perfectly replicating model should

have only hits and no misses -- the percentage of terms scored as a miss should be zero. The closer a model approaches zero on these two indicators, the less error and noise obstructs its appearance, and the more it can be considered perfectly replicated. If an etic structure replicates, both of these indicators will be impressive; weakness in either indicator would indicate poor replication.

If a five- (or six-) factor model is the optimal one, how should it behave in a solution with seven factors? Under stringent expectations, any extra factors should be relatively meaningless and unrelated to the first five (or six). Under loose expectations, within a study one of the five (or six) factors might divide itself into multiple subfactors in a seven-factor solution (e.g., an interpretation of two or even three "subfactors" of Extraversion is allowed), though ideally this should occur in a way that generalizes across studies (e.g., Extraversion should then be the one factor that consistently divides into subfactors).

Additional exploratory analyses examined the manner in which the eight studies converge on a common structure. The focus of these analyses was affected by the results of the initial tests of prior expectations, therefore these analyses are described later, after initial results.

Results

Testing Expectations Based on Prior Models

Table 2 documents the occurrence of overlap between the marker-scale items and the terms appearing among those (up to 18) having the highest loadings on each factor. For example, the O factor terms for the narrow-band Cross-Language Six were Original, Creative, Intellectual, Intelligent, Sharp, Clever, Gifted, Ironic, and (reverse-keyed) Conservative and Conventional. Three terms on Filipino factor 7 had, as translated, correspondence to any terms in this set, so these are counted in the table as 3 O under Filipino factor 7, and as such give evidence that this is a Filipino version of the O factor. Looking across all the Filipino factors, one finds marker-scale items from all six of the NCL6 represented (H, E, X, A, C, and O) and so the number of missing constructs is zero. Matching up these NCL6 factors with six from the Filipino structure, the only anomalies – indicated by italics -- are one H item overlapping with factor 3, one H item with factor

7, and 2 X items overlapping with factor 6. Thus with four mismatches out of 21 total overlapping terms on that row in the table, the miss-proportion is 4/21. This can be reduced to 3/21 if the H item on factor 3 is interpreted to indicate that factor 3 is a second, subdivided Honesty factor.

Table 2 details, first, how factors from the eight studies were associated with Big Five marker scales (i.e., the Big Five constructs). Studies varied markedly in how many Big Five constructs were missing, from zero (Chinese, Turkish, English-OR) to two (Spanish, English-MN); Intellect was the construct most typically missing. Miss-rates also varied markedly, four being zero or near zero and thus excellent (Spanish, English-MN, English-OR, Turkish) and four being in the 20 to 50 percent range and thus relatively poor (Chinese, Filipino, Hebrew, Greek). Two studies (Turkish, English-OR) had a good Big Five replication – both indicators close to zero. Across studies and factors, Big Five constructs were missing at an 18% rate, and the miss-rate was 25% – decreasing to 15% in a looser application of the model in which Agreeableness (II) and Emotional Stability (IV) each split into two factors in seven-factor solutions.

Table 2 details, next, how factors from the eight studies were associated with narrow-band Cross-Language Six marker scales (the NCL6 constructs). No more than one construct was ever missing in a study, and in those three studies where a construct was missing, it was always the same construct (Honesty). The studies also varied relatively little in their miss-rate percentages, all falling in a range from 8 to 27 percent. Across studies and factors, the miss-rate was 14% for the NCL6. This decreased to 13% when accepting a looser application of the model – allowing Honesty to split in the seven-factor solution. Across studies and factors, NCL6 constructs were missing at a 6% rate.

Table 2 details, next, how factors from these studies were associated with Multi-Language Seven marker scales (the ML7 constructs). There was much variation between studies in how many ML7 constructs were missing, from zero (Filipino, Hebrew, and Chinese) up to three (Greek, English-MN) or four (Spanish). Miss-rate percentages also varied considerably, from 0 to 50, across studies. Across studies and factors, the miss-rate was 22% and constructs were missing at a

29% rate, if the Filipino and Hebrew studies were excluded. Replication indicators for the ML7 were better than those for the NCL6 in Filipino and Hebrew, but one would expect such given the derivation of the ML7 model from these two studies (Saucier, 2003a). The NCL6, in an advantageous contrast, was not developed based on any of the languages here, yet appears to generalize substantially to them.

Table 2 details, finally, how factors from the eight studies were associated with Big Seven marker scales (the Big Seven constructs). Markers for all seven constructs were identified in only one study (English-MN), the same study in which the model was derived. Excluding that study, overall, constructs were missing at a 31% rate, and the marker-item miss-rate was 20%, both percentages being similar to those for the ML7. A major source of weakness was that, other than the study (English-MN) in which the Big Seven was derived, the Conventionality construct was isomorphic with a naturally appearing factor in only two studies – Spanish and Turkish.

The best-replicated model, then, appeared to be the narrow-band Cross-Language Six (NCL6). This model's constructs were missing only rarely from factors derived in these studies; the missing-construct rate was 6% as compared with 18 to 31% for other models. The marker-term miss-rate for the NCL6 was also the lowest, 14% as compared with 15-25% for other models.

Differences between an observed and a hypothesized (test) proportion can be analyzed with a binomial test. Here, the NCL6's missing-construct proportion (.0625) and marker-item miss-rate (.1364) was compared to that for other models. With critical value (p) at .01, by margins exceeding chance expectations (a) all the other models had higher missing-construct proportions than the NCL6 and (b) the Big Five (except in the looser application) had a higher marker-item miss-rate proportion than the NCL6. These analyses show that the narrow-band Cross-Language Six (NCL6) accounts for the wide-band factors better, by a margin beyond what would be expected by chance.

However, none of the four models had a perfect replication rate. At best, they still imperfectly captured the structural pattern emerging from these inclusive-variable-selection lexical studies. To capture the convergent structural pattern -- to identify a model that would have

had closer to a zero miss-rate and a zero missing-construct rate across these eight studies – a more exploratory analytic strategy was used.

The Structural Pattern on Which the Eight Studies Converge

To identify any convergent structural pattern across the eight studies, attention was devoted to terms appearing as high-loading variables in two or more studies, serving to link a factor in one study to a factor in another. There were exactly 100 such recurrent terms, based on defining "high-loading" as the 10 highest loading terms.

Hierarchical cluster-analysis was used to discern the structural pattern on which these eight studies converge.[4] A data matrix was constructed in which the 100 recurrent terms were cases, and 56 factors (seven for each of the eight studies) were variables. In this 100 X 56 matrix, the appearance of a term on a given factor was represented by 1, its non-appearance on a factor by 0. To the extent that terms cluster together, it indicates that across studies they tend to appear on the same factor. To be clear, it does not directly signify that across *subjects* the terms cluster together, it only implies that they would: Certainly any terms that tend to appear on the same factor across studies should also tend to cluster together in a single, typical sample from those studies.

Figure 1 shows a hierarchical agglomerative clustering of the 100 terms, a within-groups average-linkage procedure with correlation (phi) coefficients as a distance measure. In average-linkage procedures, an entity (case or variable, or a cluster of such) is joined to a cluster when it has a high calculated average of similarities to the entities within that cluster, and that "average linkage" is the highest for any entity at that juncture in the joining (agglomeration) process; the within-groups procedure defines the distance between two clusters as the average distance between all pairs of objects belonging to either cluster. Use of a between-groups average-linkage procedure (defining distance between two clusters as the average distance between all cluster 1 objects and all cluster 2 objects) produced similar clusters. So did use of these procedures with an alternative distance measure (squared Euclidean distance), although this variation produced high-level clusters (at the key 5 to 7 cluster level) that were far more unequal in

size and thus less comparable to factor-analytic structures. The content in each cluster tended to include a distinct unfavorable pole, and in Figure 1 terms associated with the unfavorable pole have a hyphen (-) in front of them. Direction (negative or positive) of association with factors is not encoded in the co-occurrence matrix, but in the eight studies the unfavorable-pole items in Figure 1 always occurred at the opposite pole of a factor from any favorable-pole items in their cluster.

The figure has five major groupings that further divide into six and then seven groupings, reflecting the agglomeration sequence in the clustering. Of the five groupings, they can be identified (from top to bottom, with labels indicated) with Agreeableness (A), Originality/Talent (O), Conscientiousness (C) and Negative Valence (NV) joined together, Extraversion (Ex), and then a grouping representing Resiliency versus Negative Emotionality of an "internalizing" variety (R vs. INE; *externalizing* negative emotionality, such as quick-temperedness, is found on the low end of Agreeableness). Except for the joining of NV to C, this configuration has some affinities to the Big Five. Next, NV and C split to make six clusters, which resemble the narrow-band Cross-Language Six except for the replacement of Honesty by NV. To create seven clusters, Agreeableness (A) splits into two parts: even temper (hostility or its absence) and cooperation or accommodatingness (peaceful, easygoing, and tolerant, rather than stubborn and combative). The next split yields eight clusters, as Originality/Talent splits into Intellect (I) and Positive Valence (PV).

The major clusters just described were then related to specific sets of factors from the eight studies. Cluster-membership assignments for each of the 100 terms for the six-cluster solution in the within-group average linkage analysis (distance measure phi), were tabulated for each factor from each study. An attempt was then made to assign each factor in each study to a cluster, the assignment being made if (a) that cluster was the location for at least two of its terms and (b) more terms were assigned to that cluster than to any other cluster. This procedure enabled 49 of the 56 factors (8 studies, 7 factors each) to be assigned to a cluster, as shown in Table 3.

Table 3 shows several factors to be relatively consistent across studies. Conscientiousness was identified in all but the Filipino study; in Filipino there was a factor labeled Conscientiousness, but it emphasized frugality and piety content to such a degree that its most salient terms did not match up well with Conscientiousness terms from other studies. Negative Valence was identified in all but one study; in Turkish there was a factor labeled Negative Valence, but it simply happened to have terms that, when translated into English (e.g., as unprincipled, impolite, and swindler), did not precisely match those from other studies. Extraversion (Gregariousness/Cheerfulness) was identified in all studies but Spanish; found there instead was an Openness factor that had content (e.g., quaint, strange, mystical, bohemian) relatively unique compared to factors from other studies. An Originality/Talent factor comprising some combination of Intellect and Positive Valence (PV) content appeared in all studies. As noted earlier, the 8-cluster solution separated Intellect and PV, but in only one study (English-OR; Saucier, 1997) did Intellect and PV content constitute two separate factors. In that study the PV content was substantially mixed with terms implying attractiveness (e.g., adorable, exciting, appealing) in a 'Positive Social Stimulus Value' factor.

In five of the eight studies there was clearly a factor referencing Resiliency versus Internalizing Negative Emotionality. The Chinese study's Dependency/Fragility factor tended to resemble resiliency, however it had relatively unique content (involving fragility and childishness) compared to factors from other studies. In Turkish, the resiliency content was spread across several factors. In one English study (MN) no content appeared on any of the factors that could be categorized as resiliency versus internalizing-negative-emotionality.

As Table 3 shows, in five of eight studies, Agreeableness content tended to split into Even Temper and Cooperation/Accommodating subfactors, the former emphasizing patience and inirritability (involving affect, especially anger, and its regulation), the latter emphasizing forgiveness and undemandingness (involving an interpersonal style). Of the other three studies, English-OR had a single factor emphasizing primarily accommodating, Filipino had a single factor emphasizing mainly even temper, and Greek had a single factor clearly representing both aspects.

The Multi-Language Seven also splits Agreeableness into two. However, the pattern in the third row of Table 3 indicates that the ML7 Concern for Others factor would be better replaced by one emphasizing cooperation and accommodatingness, which is the part of Agreeableness that tends most commonly (in studies with inclusive variable selections) to bifurcate from even temper. This bifurcation, however, has not been observed in studies with restrictive variable selections. Therefore, a broader Agreeableness (A) factor that combines even temper and cooperation/accommodating seems more robust across variable-selection strategies.

The last row of Table 3 lists the seven factors (of the 56) that were least readily fit into the pattern above. These factors contain a very heterogeneous range of content (e.g., egotism, frugality/piety, honesty/honorableness, attractiveness, openness). It seems unlikely that any additional cross-language factor can be detected from these lexical studies.

Thus, these eight studies converge on a "wide-band cross-language six" (WCL6) structure. Table 4 lists the English terms most recurrent on each of the six relatively pan-cultural dimensions from Table 3. Terms are included only if they were among the 10 highest loading terms in at least two studies, and always on the same dimension as delineated in Table 3. For Agreeableness terms, indications are provided (via asterisks and daggers) to distinguish terms associated with Even Temper from those associated with Cooperation/Accommodating; two terms (Gentle and Quarrelsome) being associated with both. This table can be considered a counterpart to Table 1, which listed the most recurrent terms on six factors in narrow-band lexical studies. There are fewer terms in Table 4 than in Table 1, partly because of a more stringent threshold for inclusion (only the 10 highest loading terms from each factor in each study), and partly because structures from inclusive-selection studies have converged less strongly than have structures from narrow-selection studies (which may be due to larger variations in methodology, geographical provenance, and culture).

Overall, inclusive-variable-selection studies converge on six factors (or alternatively seven, if we allow the Agreeableness factor to be divided in two). The six factors link and integrate the

results of diverse studies. The next study examines the degree to which these six correspond with the six lexical factors of Ashton et al. (2004) and with other structural models.

Study 2

An obvious and intriguing hypothesis is that the six-factor pattern found in the eight wide-band studies (the WCL6) is no different from the narrow-band Cross-Language Six (NCL6) found in the studies synthesized by Ashton et al. (2004). If these patterns are the same, it implies that inclusiveness versus narrowness of variable selection has little effect at the six-factor level.

The Big Five is an established model, and any rival model should demonstrate important advantages over the Big Five. The six factors of Ashton et al. (2004) may be more replicable across wide-band lexical studies than are the Big Five. A further hypothesis is that, with respect to important criteria, six dimensions will offer better prediction than the Big Five.

Method

The recurrent terms listed in Table 4 were used as markers for dimensions on which inclusive-selection lexical studies tend to converge. Of the 75 Table 4 terms, all but three (Inhuman, Irascible, and Smiling) had been administered to the sample used in this study. Internal consistency and inter-scale correlations for these marker scales are presented at the top of Table 5. Aggregated scores based on these 72 markers were correlated with adjectival scales for previous five- and six-factor models. All materials were administered in self-report format in English.

Participants were members of the Eugene-Springfield community sample (58% female, mean age 51 in 1993; Grucza and Goldberg [2007] provide more details). Personality-descriptive adjectives were administered to the sample at one of five times (in 1993, 1995, 1998, 2001, and 2002), a majority of those used here being administered in 1995. For the predictive validity comparisons, desirable would be a measure comparable in length to our Big Five measure, and administered at entirely the same time as this measure; for this purpose a 47-item subset of the 72 terms, all 47 having been administered in 1995, was used. This subset had 4 to 13 terms for each dimension; to raise the Resiliency scale to four items, the term Afraid was added (in place of

Cowardly and Fearful which had been omitted because not administered in 1995). Internal consistency for the scored aggregates ranged from .76 to .87 for the longer scales, and .77 to .82 for the subset scales.[5] The subset scales corresponded very well with their longer counterparts, with matched correlations all in the .91 to .96 range. When subjected to exploratory factor analysis (principal axes, varimax rotation) the items in either set of scales yielded the anticipated six-factor structure in this American community sample, with less than 20% of the terms having their highest loading on a factor other than that expected given the Table 4 assignments.

Sample size was 520 for the correlational analyses; 520 subjects completed adjective self-ratings at all five occasions and thus had scores for all (Big Five, NCL6, and WCL6) measures. The regression analysis used that set of 440 participants who not only completed the 1995 adjective ratings but also provided data when each of the 12 criterion measures was administered.

Representing the Big Five were the Mini-Markers (Saucier, 1994), a more efficient short form of Goldberg's (1992) 100 unipolar markers, with all items administered in 1995. Also used were adjective scales for the (narrow-band) Cross-Language Six presented by Saucier (2008). Internal consistency coefficients, presented by Saucier (2008) in the current sample ranged from .76 to .83 for the Big Five, and from .65 to .76 for the narrow-band Cross-Language Six.

If the lexical structure derived from wide-band variable selections (i.e., the WCL6) provides better support for the narrow-band Cross-Language Six than for the Big Five, then correlations should indicate that the scored aggregates relate more strongly and directly to the NCL6 than to the Big Five. In multiple regression, the structure whose factors yield higher multiple correlations, in predicting (on average) each of the WCL6 aggregates, can be considered more comprehensive. In hierarchical regression, if the NCL6 trumps the Big Five, the change in R-squared should show that considerable variance in the six-factor aggregates is accounted for only when the narrow-band Cross-Language Six are added to the Big Five.

Variables related to health and psychopathology are important criterion variables for a personality measure (cf., Roberts et al., 2007; Grucza & Goldberg, 2007). Simms (2007; cf.,

Durrett & Trull, 2005) found that a Negative Valence measure added significantly to the Big Five in prediction of six out of 10 personality-disorder diagnostic scales. Criterion variables selected *a priori* for these analyses were the following, with year of administration provided in parentheses: assessment of personal health status (2002), self-reported history of medical problems (2003) and of mental health diagnoses (2003; depression, schizophrenia, bipolar, or anxiety/panic disorder), depressive symptoms (2002; CES-D, Radloff, 1977), phobic symptoms (2006; Fears Questionnaire; Marks & Mathews, 1979), obsessive-compulsive symptoms (1999; Foa et al., 1998), dissociative tendencies (1999; Goldberg, 1999), borderline personality tendencies (2000; Leichsenring, 1999), smoking (2006; lifetime history of smoking, i.e., over 100 cigarettes), compulsive drinking (2006; aggregate of 14 indicators of compulsive drinking tendencies), risk-posing behavior after drinking (aggregate score of how many of six risky behaviors -- driving, getting in an accident, fighting, being arrested, doing something that could hurt oneself, or could hurt another-- were ever engaged in after drinking too much), and history of lawbreaking behaviors (2006; aggregate of 7 items, referencing shoplifting from a store, breaking into a building or vehicle, stealing a vehicle, delinquent gang activity, assaulting someone in authority, attacking someone, being arrested). Analyses involved computation of multiple correlation (R) values for each criterion, one based on the Big Five, the other based on the consensus WCL6 model that emerged from Study 1. These stand-alone R values are the best estimates of the predictiveness of each model on its own. To address the issue of how much each model adds to the other (its predictiveness controlling for the other), hierarchical regression was used to test the size and statistical significance of changes in R^2 values.

Results

The top part of Table 5 presents correlations between the aggregated scores for the WCL6 content clusters. Averaging .24 in magnitude, none exceeded a magnitude of .47. Not presented in the table are correlations between the two subfactor scales for Agreeableness. If the subfactors are referenced with only the Table 4 terms (6 for Even Temper, 8 for Accommodating) associated

distinctly with one or the other subfactor, that correlation was .49. If the two remaining Table 4 terms (Gentle versus Quarrelsome) are added to the Even Temper scale, so that recurrent Agreeableness terms are fully rather than selectively represented, the correlation increased to .58.

The middle of Table 5 also presents correlations with Big Five and narrow-band Cross-Language Six scales. One finds good one-to-one correspondence between NCL6 scales and the WCL6 aggregates derived in Study 1, although this correspondence is weaker for Negative Valence and Resiliency (*r* .55 and .57) than for the other four dimensions (*r* .80 to .88). To provide some comparison, correlations between differing major measures of five-factor Agreeableness and of Openness/Intellect (e.g., Goldberg, 1992, Table 6) have shown similarly weak correspondence.

The NCL6 Agreeableness scale correlated nearly equally with the Even Temper and Accommodating subfactor scales (.72 and .76 respectively); in contrast, Even Temper was more highly correlated with Big Five Emotional Stability than Agreeableness (.63 versus .39), whereas Accommodating had the reverse pattern (.43 versus .68). Although the Big Five had good (.61 to .87) one-to-one correspondence with five of the six aggregates, none of the Big Five uniquely accounted for Negative Valence. For the Big Five, the problem is that both Agreeableness and Negative Valence factors appear, and the Big Five has only one factor in their place. The NCL6 has the advantage that one of its factors (Honesty) uniquely corresponds to NV.

Table 5 also provides multiple correlations for the set of Big Five and the set of NCL6 scales, in predicting each of the WCL6 aggregates. In five of six cases (all but Conscientiousness), multiple correlations were higher for the NCL6 set than for the Big Five set. In hierarchical regression, when the NCL6 scales were added as a second block of predictors with the Big Five scales as the first block, the average change in R^2 was .10; the change in R^2 was significant ($p < .001$) for all of the six WCL6 aggregates. In contrast, when the block-order was reversed (NCL6 first, Big Five second), the Big Five scales led to less change in R^2 -- only .04 on average (though all changes in R^2 were still significant, $p < .002$). Thus, despite their lower reliability, and the inexactness of matching between two pairs of factors (Honesty-Negative Valence and

Emotionality-Resiliency), NCL6 scales still accounted for an additional 6% on average of the variance in each WCL6 aggregate from Study 1.

The analyses reflected in Table 5 indicate an advantage for the narrow-band Cross-Language Six over the Big Five. The NCL6 accounts better than the Big Five for those dimensions that tend to be recurrent in lexical studies with inclusive variable selections.

However, the NCL6 fails to give a perfect account of the six inclusive-selection personality dimensions, derived in Study 1. Clearly, variable selection does have some effects on structure. A particular anomaly is Negative Valence, which had by far the lowest multiple R-squared values of any scored aggregate. This indicates that Negative Valence is the factor least subsumable into either the Big Five or the NCL6 model. However, a NCL6 framework can account for NV as a more strongly evaluative variant of Honesty (consistent with the key -.55 correlation between them in Table 5) that appears when the variable selection is more inclusive. The WCL6 does not equal the NCL6, but these are evidently rather closely related structures.

It is illuminating to compare the NCL6 to the Multi-Language Seven (ML7) model, in terms of accounting for the WCL6 dimensions. Saucier (2003a) provided a set of 60 adjective markers of the ML7 model, with coefficient Alpha in this sample ranging from .70 to .81. WCL6 dimensions corresponded well with six of the ML7 scales, from .72 to. 82. The seventh ML7 scale, Concern for Others, correlated .48 with WCL6 Agreeableness; ML7 Concern for Others correlated .56 with Accommodating, whereas ML7 Even Temper correlated .77 with WCL6 Even Temper. The multiple R values were fairly homogeneous across the WCL6 dimensions, all in a range from .78 (NV) to .89 (Extraversion). The mean multiple R-squared value of .69 slightly exceeded that (.65, Table 5) for the NCL6; the increment in predictiveness was largely due to better correspondence with the NV and Resiliency dimensions, but may also stem from some over-fitting of the ML7 measurement model to this particular sample (the same in which these ML7 markers were developed). Overall, this analysis indicates that the eight lexical studies do tend to converge on the Multi-Language Seven structure, but with Concern for Others replaced by Accommodating.

Table 6 presents multiple correlation coefficients comparing the Big Five (40 item measure) and the six dimensions (47 item measure) found in the present study, in predicting 12 criterion variables. The R coefficients provided are based on each model entered alone—without the scales in the other model. The six factors produced higher multiple R values for all 12 criterion variables. For these 12 criteria, the average R-squared for the Big Five alone was .082 and the average R-squared for the WCL6 aggregate scales alone was .121; that is, the six factor model accounted for an additional .039 of variance, on average, in these criteria. This corresponds to a 47% increase (.039/.082) in the amount of variance accounted for in a criterion variable, when using the WCL6 model alone versus the Big Five model alone.

When WCL6 scales were entered as step 2 in hierarchical regression, with Big Five scales as step 1, the change in R^2 – for which *p* values are provided in Table 6 -- was significant for 10 of 12 criteria (eight with $p < .001$). When the order of entry was reversed (WCL6 in step 1, Big Five in step 2), the change in R^2 was significant for only 5 of 12 criteria (just one with $p < .001$).

Analyses reported in Table 6 involve the combined (WCL6) Agreeableness scale. When the Agreeableness scale was split into two variables, Even Temper and Accommodating, the multiple R values in the WCL6 column were essentially the same, never higher by more than .01, except in one case. With split Agreeableness predictors, the R for the smoking-behavior criterion increased from .13 to .16; in this instance, the zero-order *r* for Even Temper (-.12), unlike that for Accommodating (.02) and for other six-factor scales, was significant, although the overall R still was not. Such a small gain in prediction, purchased with a loss in parsimony, does not strongly argue for separating the two sub-factors and using a seven-factor model instead of six.

Correlation coefficients were examined to determine the source of the incremental predictiveness of the WCL6 aggregates. Resiliency was the strongest stand-alone predictor for the mental health history, phobia, depression, and borderline criteria. 'Negative Valence' was the strongest predictor for lawbreaking behavior, compulsive drinking, risky behavior after drinking, and for dissociative tendencies. WCL6 Extraversion and Originality/Talent were the strongest

predictors for one criterion each. These results are consistent with a view that, among the six dimensions, Resiliency and NV produce the greatest portion of incremental prediction beyond the Big Five, at least for criterion variables of these kinds.

General Discussion

Lexical studies with inclusive variable-selection strategies yield varying sets of factors, as the varying labels in Table 3 imply. However, Study 1 indicated that these sets of factors can be understood in terms of six consensus dimensions, with the proviso that one of these factors (Agreeableness) tends often to divide into two subfactors (Even Temper and Accommodating).[6] Study 2 indicated that the narrow-band Cross-Language Six adds substantially to the Big Five in accounting for the six dimensions that tend to be recurrent in lexical studies with inclusive variable selections. Thus, from the standpoint of lexical studies examined here, and by Ashton et al. (2004), the Big Five seems to be one factor too few. Study 2 also established that a six-dimensional model adds substantial validity increments for predicting important criteria, and that Resiliency and Negative Valence factors are major sources of these increments.

Study 2 (and Table 5) implied that the Narrow-band Cross-Language Six would be even more comprehensive were further content from the so-called Negative Valence (NV) factor incorporated. Although NV content was partly assignable to NCL6 Agreeableness and Conscientiousness (*r* over .40 in both cases), it was more highly associated with NCL6 Honesty (*r* = -.55). Contentwise, Honesty and NV clearly share an emphasis on tendencies toward amoral/immoral (e.g., dishonest, corrupt, cruel, greedy) interpersonal behavior. This convergence on similar content across variable-selection strategies allays concerns that including highly evaluative terms might lead to factors without meaningful substantive content.

The narrow-band Cross-Language Six might be seen as a special case, observed with relatively narrow variable-selection strategies, of a more comprehensive and generalizable model -- the Big Six. The Big Six would appear in two major variants: the NCL6 in a narrow variable-selection context, and a closely related structure (the WCL6) in inclusive selections.

Resembling the narrow-band Cross-Language Six, the Big Six would be distinct as follows:

1. Replacing the factor labeled as Openness, the Big Six has a broad factor encompassing Originality/Talent that includes major components of Positive Valence (included in the WCL6 but not the NCL6). This is a factor of perceived abilities, of originality, and of intellectual and aesthetic interests. The association of Intellect with Positive Valence is not evident with a narrow selection of variables, because of the exclusion of PV terms (e.g., Impressive, Outstanding). As for Attractiveness descriptors, their place within the Big Six is unresolved – in the English-OR data they joined the Originality/Talent factor, but in the Turkish data they formed a separate factor apparently "beyond the Big Six."

2. Because in more inclusive variable selections NCL6 Emotionality tends to morph into a WCL6 Resiliency factor, the Big Six will have a fuller representation of descriptors of Internalizing Negative Emotionality (depression, anxiety, tendencies toward panic and phobias), that appear on their face to be related to internalizing disorder tendencies (Krueger & Markon, 2006). Why would a wide-band variable selection conduce to a Resiliency factor? Possibly, because including more words for emotional states (e.g., Sad, Angry, Disgusted, Frustrated) allows for clearer emergence of the distinction between externalizing versus internalizing forms of negative emotionality.

3. The Big Six involves addition of Negative Valence (NV) content to the (Dis)Honesty factor. Although there may be occasional cases (as with Filipino factors 1 and Greek factors 5 and 6; see Tables 2 and 3) where distinguishable NV and Honesty factors appear, with inclusive variable selections the NV variant seems as a general rule more likely to appear., With narrow selections the (Dis)Honesty variant is more likely to appear because prime NV terms are removed. The joint NV versus Honesty factor probably relates substantially to externalizing disorder tendencies (Krueger & Markon, 2006), although (low) Agreeableness and Conscientiousness are also likely to have some relation to externalizing. Across these eight studies, the most prototypical terms for NV (from Table 4) were Cruel, Corrupt, Disgusting, and Wicked. The descriptive content of NV resembles that for Honesty, focusing on tendencies toward violations of propriety, of moral

standards, and of the rights and trust of other people, in a way that can provoke a social effect -- moral outrage in others (indicated by terms like Disgusting and Wicked). Negative Valence is an overly ambiguous label for a pan-cultural factor with such content. Better labels would be Amorality, Psychopathy, Violativeness (Violation-Proneness), or to label by the favorable pole, Propriety or Non-Violativeness (better than "Negative Valence' while allowing retention of the NV label).

With these expansions, we can conceive a Big Six structural model that tends to generalize across a majority of 16 lexical studies, conducted in 14 languages -- and not generalizing only across those studies with narrow variable-selections. In lexical studies with narrow variable selections, the Big Six (represented by the narrow-band Cross-Language Six, presented in Ashton et al. [2004]) seems about as replicable as the Big Five. When we turn to lexical studies with inclusive selections, the Big Six seems to have a decisive advantage over the Big Five. In building a comprehensive classification of personality descriptors, the Big Six seems a more cross-culturally valid starting point than the Big Five. The Big Six might be operationalized from convergences between and combinations of the terms and themes found in Table 1 and in Table 4.

Is the Big Six structure a cross-cultural universal? By a very stringent standard, it is not. The structure has not appeared identically in all studies in all languages. Its latent pattern can be detected across the sixteen lexical studies reviewed here and by Ashton et al. (2004), but this pattern seems prone to have pieces missing in many single studies – as the blank cells in Table 3 will indicate. Better candidates for "strong universal" status, appearing with more invariance across lexical studies, are structures found in one- and two-factor solutions (Saucier & Goldberg, 2003): A "socially desirable qualities" dimension at the single-factor level, and dimensions of "social self-regulation" and "dynamism" at the two-factor level. However, these are less informative structures, so future research may prove them inferior for prediction purposes.

Figure 2 presents a hierarchical organization of personality-attribute structures, proposed to hold across narrow- versus wide-band variable-selection procedures. Factors and labels

invariantly applicable across variable selection are in bold. Those factors and labels applicable only for wide-band studies (including the frequent split of Agreeableness into even temper and accommodating components) are in italics. Omitting the Honesty/NV factor (shaded for emphasis), the remaining five factors at that level could pass for a Big Five representation – so this figure also integrates the Big Five into a structural hierarchy that can be tested in future studies.

These studies have two principal limitations. First, many terms appearing in the nine data sets could have been given alternative translations in the original investigations, which would then have affected the exact outcome of results here. However, any error so produced is unsystematic, and likely leads mainly to type II error – a heightened false negative rate. As an example, terms translated as Idiotic could probably just as well be translated as Stupid, and one study (Turkish) had a term translated as Idiotic but not as Stupid, leading to these results actually underestimating the pervasiveness of the concept Stupid. The recurrence of terms in Study 1 should be considered a conservative, lower-bound estimate on the pervasiveness of the concepts they represent.

Second, the validity comparison in Study 2 involved a disconnection in time between the predictors and the later-administered criterion variables. This does make the predictors interpretable as *forecasters* of the criteria. But a more refined approach would administer the criterion variables both simultaneously with the predictors and also several years later, allowing separate inferences for concurrent versus prospective validity. Present data did not allow that separation.

In a critique of the Big Five, Block (1995, p. 221) suggested there be "empirical and conceptual competition between alternative dimensional offerings to see which, predictively, best carves nature at its joints." The present studies provide such competition. If competition between dimensional offerings indicates an overall advantage for the six-factor model, this does not mean that attention to five-factor models has been a wasteful detour. A characteristic of good science – though not of dogma or ideology – is that exposure to empirical tests leads to upgrades in understanding, changing the standard model of reality. The Big Five model might be seen as an

initial platform that has beneficially enabled an expansion of knowledge about personality. In the practice of strong science, of course, such platforms do not devolve into pedestals.

Conclusions

Goldberg's (1981) clarion call – for studies to identify the most important dimensions of personality attributes represented in languages around the world – has now led to important observed convergences in factor-analytically derived personality structure between languages. Selection of variables contributes to the results of any factor analysis, and provides a potential source of discrepancies between studies. Fortunately, the present studies indicate that some major convergences in structure among lexical studies – identifiable with a Big Six model -- are also somewhat resilient across variable-selection strategies.

Lexical studies are important tools to discover basic dimensions for a science of personality. Natural-language dimensions have guaranteed social importance. When a set of these dimensions arises across a number of languages (not constrained to do so by a pre-selected, imposed set of variables or by pre-set model restrictions) it is something impressive. Such freely arising cross-cultural generalizability lights a way toward measures that will have adequate invariance across populations. With use of a cross-cultural generalizability criterion, we can exclude constructs whose definitions do not translate well from the language of origin to other language contexts, while retaining the more readily translatable ones.

Convergences among lexical studies point to personality dimensions that are relatively pervasive and universal. Pervasiveness and universality are two of four criteria often cited (e.g., by Costa & McCrae, 1992) in arguing that a five-factor structure is "basic." However, based on lexical-study evidence examined here, a more informative Big Six structure would appear to be at least as pervasive and universal as the currently popular five-factor model.

References

Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs, 47*, (1, Whole No. 211).

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

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

Benet, V., & Waller, N. G. (1995). The Big Seven factor model of personality description: Evidence for its cross-cultural generality in a Spanish sample. *Journal of Personality and Social Psychology, 69*, 701-718.

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

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117*, 187-215.

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

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

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13,* 653-655.

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

De Raad, B., Levert, E., Ostendorf, F., Mlacic, B., Di Blas, L., Hrebickova, M., Saucier, G., Szirmak, Z., Szarota, P., Perugini, M., Church, A. T., & Katigbak, M. S. (2008). *Cross-language replicable personality factors: A comparison of thirteen six-factor structures from twelve languages*. Manuscript submitted for publication.

De Raad, B., Perugini, M., Hrebickova, M., & Szarota, P. (1998). Lingua franca of personality: Taxonomies and structures based on the psycholexical approach. *Journal of Cross-Cultural Psychology, 29*, 212-232.

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

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology* (Vol. 41, pp. 417-440). Palo Alto, CA: Annual Reviews.

Durrett, C., & Trull, T. J. (2005). An evaluation of evaluative personality terms: A comparison of the Big Seven and Five-Factor Model in predicting psychopathology. *Psychological Assessment, 17*, 359-368.

Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E., & Amir, N. (1998). The validation of a new obsessive-compulsive disorder scale: The Obsessive-Compulsive Inventory. *Psychological Assessment, 10*, 206-214.

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

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

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

Goldberg, L. R. (1999). The Curious Experiences Survey, a revised version of the Dissociative Experiences Scale: Factor structure, reliability, and relations to demographic and personality variables. *Psychological Assessment, 11*, 134-145.

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

Grucza, R. A., & Goldberg, L. R. (2007). The comparative validity of 11 modern personality inventories: Predictions of behavioral acts, informant reports, and clinical indicators. *Journal of Personality Assessment, 89*, 167-187.

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

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.

Krueger, R. F. & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology, 2*, 111-133.

Leichsenring, F. (1999). Development and first results of the Borderline Personality Inventory: A self-report instrument for assessing borderline personality organization. *Journal of Personality Assessment, 73*, 45-63.

Marks, L. M., & Mathews, A. M. (1979). Brief standard self-rating for phobic patients. *Behavior Research and Therapy, 17*, 263-267.

McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology, 49*, 710-721.

McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist, 52*, 509-516.

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.

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]. Regensberg, Germany: S. Roderer Verlag.

Pendlebury, D. (1996). Which psychology papers, places, and people have made a mark? *APS Observer, 9*, 14-18.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385-401.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological*

Science, 2, 313-345.

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

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

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

Saucier, G., & Goldberg, L. R. (2003). The structure of personality attributes. In M. Barrick and A. M. Ryan (Eds.), *Personality and work* (pp. 1-29). New York: Jossey-Bass-Pfeiffer.

Simms, L. J. (2007). The Big Seven model of personality and its relevance to personality pathology. *Journal of Personailty, 57,* 65-94.

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

Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychological Inquiry, 4*, 122-130.

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

Zhou, Y., Saucier, G., Gao, D., & Liu, J. (in press). The factor structure of Chinese personality terms. *Journal of Personality.*

Notes

1. Goldberg's position spurred McCrae and Costa (1985) to add Agreeableness and Conscientiousness to an earlier three-factor (NEO) model to yield a questionnaire analogue – the Five-Factor Model (FFM). Studies (McCrae & Costa, 1997) have shown that translations of a five-factor inventory retain reasonably good psychometric characteristics (especially reliability and some factorial invariance under one style of confirmatory factor analysis) when translated into numerous languages. But studies that merely translate a measure are not lexical studies.

2. The Turkish lexical study (Goldberg & Somer, 2000) reported an Attractiveness factor. Only a few attractiveness terms were associated with this factor, which also had intellect-related content.

3. Terms having a common root and meaning but varying word form (e.g., Corrupt/Corrupted, Disciplined/Self-Disciplined, Hesitant/Hesitating, Bad-Tempered/Ill-Tempered/Quick-Tempered) were treated as a single term. This consolidation of root-sharing terms was carried out for 14 roots.

4. Cluster analyses based on as many 18 highest loading terms were examined, but were less clearly interpretable. This lenient inclusion criteria allows more variables far from the prototypical core of the factors, and clusters may then form based on these relatively peripheral terms.

5. The same sample was used in a lexical study (Saucier, 1997) examined in Study 1. That study only roughly reflected the consensus pattern among the eight lexical studies; its Agreeableness factor focused primarily on accommodating content, even temper content being divided between this factor and an Emotional Stability factor. The internal consistency and scale-intercorrelation statistics indicate, however, that these data can be constrained to produce the WCL6. The exploratory factor analysis of these 72 terms, which yielded six WCL6-like factors, allows the same inference.

6. Lexical studies with an inclusive variable-selection strategy have in each case emphasized seven-factor solutions. Indications are that one could identify something resembling the six consensus factors in English or Chinese or Turkish data's first six factors, but not as well for Filipino or Greek or Spanish; for Hebrew there is insufficient information to make an inference.

Table 1

Adjectival Personality Concepts Recurrently Associated with Six Prime Factors in Eight Conventional Narrow-Selection Lexical Studies Examined by Ashton et al. (2004)

Conscientiousness: Diligent, Precise (6); Conscientious, --Irresponsible (5); Disciplined, Orderly, --Lazy, --Reckless (4); Dutiful, Industrious, Meticulous, Organized, Thorough, --Absent-minded, --Careless, --Disorderly, --Frivolous, --Rash (3); Hard-working, Methodical, Scrupulous, Steadfast, --Chaotic, --Imprudent, --Inaccurate, --Inattentive, --Inconsiderate, --Inconstant, --Irrational, --Lax, --Negligent, --Undisciplined, --Untidy, --Wishy-washy (2).

Honesty/Humility: Honest (6), Sincere, --Hypocritical (5); Loyal, --Conceited, --Greedy (4); Just, --Boastful, --Calculating, --Dishonest, --Sly (3); Altruistic, Modest, Truthful, --Naughty, --Lying, --Pompous, --Pretending, --Pretentious, --Stingy, --Untruthful (2).

Agreeableness: Peaceful, Tolerant, --Aggressive, --Choleric (5); Mild, Patient (4); Agreeable, Good-natured, --Authoritarian, --Hot-headed, --Irritable, --Stubborn (3); Accommodating, Conciliatory, Kind-hearted, Lenient, Sympathetic, Tranquil, Warm, --Brusque, --Explosive, --Fierce, --Irascible, --Quarrelsome, --Quick-tempered, --Short-tempered (2).

Emotionality: Vulnerable (6); Emotional (5); Anxious, Sentimental, --Courageous, --Self-Assured, --Strong (4); Fragile, --Brave, --Imperturbable, --Independent, --Resolute (3); Delicate, Depressive, Fearful, Hypersensitive, Indecisive, Insecure, Melancholic, Oversensitive, Suggestible, Whining, Worrying, --Bold, --Intrepid, --Secure, --Stable, --Tough (2).

Extraversion: --Reserved (7); Sociable, --Introverted, --Silent (6); Lively (5); Cheerful, --Passive, --Quiet, --Shy, --Withdrawn (4); Extraverted, Talkative, Vivacious, --Solitary, --Taciturn (3); Exuberant, Hyperactive, Merry, Open, Vigorous, --Boring, --Distant (2).

Openness: Original (5); Creative, Intellectual, Intelligent, Sharp (4); Clever, Gifted, Ironic, --Conservative, --Conventional (3); Artistic, Bright, Critical, Educated, Inventive, Receptive, Smart, Talented, Wise, Witty; --Backward, --Ignorant, --Incompetent, --Obedient, --Uneducated, --Unintelligent (2).

Notes. Parentheses indicate net number of studies with the term represented on the factor. '–' in front of a term indicates a negative-loading direction on the factor.

Table 2

How Marker Items for 'Four Previous Personality-Structure Models Overlap With Salient Terms for Factors Found in Eight Lexical Studies

	Factor Numbered to Reflect Order of Appearance in the Lexical Study								
Study	1	2	3	4	5	6	7	Missing Constructs	Marker-Item Misses/Total
Big Five									
Filipino	2 II, *1 IV*	--	*1 II**	4 IV, *1 II*	2 I	*2 IV*, 1 I*	1 III	V	6/15 (3/15)
Chinese	5 I, *2 II*	3 III	3 II, *2 V*	*1 II**	2 IV, *2 I*	2 V	*2 IV*, 2 III, 2V*	--	13/28 (10/28)
Hebrew	2 I, *1 IV*	1 IV	--	*1 IV**	4 III	2 II, *1 III*	*1 II*, 1 I*	V	5/14 (3/14)
Turkish	6 I	5 III	2 II	3 IV	--	1 V	*1 II**	--	1/18 (0/18)
Greek	2 I, *2 II*	*2 IV**	3 III	4 IV, *1 I*	1 II, *1 III*	--	*1 III**	V	7/17 (4/17)
Spanish	--	--	3 I	2 IV	*1 IV**	--	1 V	II, III	1/7 (0/7)
English-MN	--	--	4 I	2 IV	3 III	--	--	II, V	0/9
English-OR	1 IV	7 II	8 III	--	3 V	6 I	*1 III**	--	1/26 (0/26)
							Mean number per study	**0.88**	**4.25/16.75** (2.5/16.75)
							As percentage	18% (of 5)	25% (15%)
Narrow-band Cross-Language Six									
Filipino	3 H	3 C	*1 H**	3 A	3 X	2 E, *2 X*	3 O, *1 H*	--	4/21 (3/21)
Chinese	9 X	8 C	4 H, *1 A*	--	2A, *1 X, 1 E*	2 O	5 E	--	3/33
Hebrew	2 E, *1 X*	3 H	2 O	--	3 C	1 A, *1 C*	2 X	--	2/15
Turkish	4 X, *1 C*	4 C	4 A, *1 H*	6 E	1 H	4 O, *2 E*	--	--	4/27
Greek	6 X	2 E	5 C	5 A, *1 X*	2 H, *1 C*	--	1 O	--	2/23
Spanish	--	--	2 X, *1 E*	1 E, *1 C*	2 C, *1 A*	2 A	1 O	H	3/11
English-MN	--	--	4 X	1 A	2 C	1 E, *1 A*	1 O, *1 C*	H	2/11
English-OR	1 E	--	3 C	--	4 O	4 X, *1 A*	--	H	1/13
							Mean number per study	**0.38**	**2.63/19.25** **(2.5/19.25)**
							As percentage	6% (of 6)	14% (13%)
Multi-Language Seven									
Filipino	1 CF	5 C, *1 NV*	2 NV	5 ET	5 G	6 SA	2 I	--	1/27
Chinese	2 G, *1 SA*	3 C	2 CF	2 NV	5 ET, *1 G, 1 SA*	4 I	1 SA	--	3/22
Hebrew	2 SA	1 NV	2 I	3 ET	3 C	2 CF	2 G	--	0/15
Turkish	3 G, *1 SA*	3 C	1 CF, *1 ET*	1 ET, *1 SA*	1 SA	--	1 I	NV	3/13
Greek	1 G	3 I, *2 SA*	--	*1 G*	*1 NV*	2 NV	--	CF, ET, C	5/10
Spanish	--	--	1 SA	--	1 C, *1 NV*	--	--	CF, ET, C, I	1/3
English-MN	1 I	--	3 G	1 ET	1 C	--	*1 C*	CF, NV, SA	1/7
English-OR	2 SA	3 CF	1 C, *1 NV*	--	4 I	3 G	4 NV	ET	1/18
							Mean number per study**	**2.00**	**2.33/10.5**
							As percentage	29% (of 7)	22%

Table 2 (continued)

Study	Factor Numbered to Reflect Order of Appearance in the Lexical Study							Missing Constructs	Marker-Item Misses/Total
	1	2	3	4	5	6	7		
Big Seven									
Filipino	--	1 C	1 NV	2 NE	3 PE	*1 NE*	--	CNV, A, PV	1/8
Chinese	3 PE	1 CNV	--	5 NV	2 NE, *2 PE*	1 PV, *2 CNV, 2 PE*	--	A, C	6/18
Hebrew	--	--	1 PV	1 NE, *1 PE, 1 A*	1 C	--	1 PE	CNV, NV	2/6
Turkish	2 PE	2 C	--	2 A	--	4 CNV, *1 PV*	--	NE, NV	1/11
Greek	2 PE	1 A	3 C	1 NE, *2 A, 1 PE*	--	1 NV	--	CNV, PV	3/11
Spanish	1 PV	2 NV	--	1 PE	1 C, *1 CNV*	2 A	1 CNV	NE	1/9
English-MN	10 PV	10 NV	11 PE	10 NE	8 C	10 A	11 CNV	--	0/70
English-OR	--	--	4 C	4 PV	--	2 PE	4 NV	CNV, NE, A	0/14
						Mean number per study* **		**2.14**	**2/11**
							As percentage	31% (of 7)	18%

Note. Constructs are labeled in abbreviated form as follows: I – Extraversion, II – Agreeableness, III – Conscientiousness, IV – Emotional Stability, V – Intellect/Imagination. C – Conscientiousness, H – Honesty, A – Agreeableness, E – Emotionality, X –Extraversion, O - Openness. NV– Negative Valence, CF – Concern for Others, ET – Even Temper, SA – Self-Assurance, G –Gregariousness, I - Intellect. NE – Negative Emotionality, PE – Positive Emotionality, PV – Positive Valence, CNV – Conventionality. Numbers just preceding construct labels are the number of marker-terms from the construct appearing as a salient term on the given factor. Missing Constructs – number of constructs in the model not accounted for by any salient terms. Marker-Item Misses/Total – number of marker items found on wrong factor, as a ratio to total number of marker items appearing on any factor (ratio in parentheses is ratio, if different, assuming loose rather than stringent expectations of the model, i.e., allowing a construct to be represented by multiple subfactors). *Factor labels and numbers of items printed in italics reference those marker-items that were counted as misses.* * - Under a loose replication standard, this could be counted as the second subfactor of the factor indicated ** - Because the Multi-Language Seven was developed by studying structures from these same Filipino and Hebrew studies, means and percentages are calculated across only the other six studies. *** - Because the Big Seven was developed from the "English-MN" data, means and percentages are calculated across only the other seven studies.

Table 3

How Factors from Eight Lexical Studies Map Onto Six Consensus Factors

	Hebrew	Greek	Chinese	Filipino	Spanish	English-OR	English-MN	Turkish
Conscientiousness	5 Dependability	3 Conscientiousns.	2 Conscientiousns.	--	5** Temperance	3* Conscientiousns.	5 Conscientiousns.	2 Conscientiousns.
Negative Valence (vs. Non-Violativeness)	2** Neg. Valence	6 Neg. Valence	4 Noxious Violativeness	3 Neg. Valence	2 Neg. Valence	7 Neg. Valence	2 Neg. Valence	--
Agreeableness (Even Temper and Accomm./Cooperation)	4 and 6 Neg. Emotlty. & Agreeability	4 Even Temper	5* and 3* Emot.Volatil. & Unselfishness	4* Tempera- mentalness	4* and 6 Engagement & Agreeableness	2* Agreeableness	4 and 6 Neg. Emotlty. & Agreeability	4* and 3 Emot. Stability & Agreeableness
Resiliency versus Internalizing Negat. Emotionality	1 Agentic PE	2* Prowess/Heroism	--	6** Self-Assurance	3 Pleasure	1 Emot. Stability	--	--
Extraversion (Gregariousness/ Cheerfulness)	7 Communal PE	1 Pos. Affect/ Agreeableness	1 Extraversion	5 Gregariousness	--	6* Extraversion	3 PE (Positive Emotionality)	1* Extraversion
Originality/Talent	3 Pos. Valence	7* Intellect	6 Intellect/ Pos. Valence	7 Intellect	1 Pos. Valence	5 and 4 Intellect & PSSV	1 and 7 Pos. Valence & Unconv.	6** Intellect
Factors Unassigned Above	--	5 Honesty (Honor)	7 Dependency/ Fragility	1 and 2 Egotism & Conscientiousns.	7 Openness	--	--	5 and 7 Neg. Valence & Attractiveness

Note. English-OR – Saucier (1997), English-MN – Tellegen & Waller (1987). PE – Positive Emotionality. PSSV- Positive Social Stimulus Value. Multiple numbers in a cell indicate that more than one factor in a study had the specified content, and none had a decisive majority (□2/3) of it. * In addition to multiple cluster-related terms (from Figure 1) one non-cluster-related term in the figure has a salient loading on this factor. ** In addition to multiple cluster-related terms (from Figure 1) two non-cluster-related terms in the figure has a salient loading on this factor.

Table 4

Adjectival Personality Concepts Most Recurrently Associated with Six Dimensions in Inclusive-Selection Lexical Studies

---

Conscientiousness (Consistency/Organization): Consistent, Neat, Organized, --Disorganized (3); Disciplined, Hard-working, Meticulous, Moderate, Orderly, Responsible, Systematic, Tidy, --Undisciplined, --Untidy (2).

Negative Valence (vs. Non-Violativeness): Cruel (4); Corrupt, Disgusting, Wicked (3); Awful, Bad, Beastly, Dangerous, Evil, Inhuman, Insane, Vicious (2).

Agreeableness: Calm*, Kind†, --Impatient*, --Impulsive*, --Irritable*, --Quick-Tempered* (3); Easygoing†, Generous†, Gentle, Peaceful†, Tolerant†, Understanding†, --Anxious*, --Brawling†, --Fiery*, --Irascible*, --Quarrelsome, --Stubborn† (2).

Resiliency vs. Internalizing Negative Emotionality: --Cowardly, --Depressed, --Fearful, --Frustrated, --Gloomy, --Sad (2).

Gregariousness/Cheerfulness: Talkative (4); Sociable (3); Cheerful, Friendly, Gregarious, Outgoing, Smiling, Vivacious, --Silent, --Withdrawn (2).

Originality/Talent: Impressive, Intelligent, Talented (3); Admirable, Brilliant, Important, Intellectual, Knowledgeable, Original, Outstanding, Unusual, Wise, --Average, --Ordinary, --Traditional (2).

Note. '–' in front of a term indicates a negative-loading direction on the factor. Number in parentheses is the number of studies in which the term was among the ten highest loading terms. * - Agreeableness term associated distinctly with Even Temper subfactor. † - Agreeableness term associated distinctly with Accommodating/Cooperation subfactor.

Table 5

Correlations of WCL6 Scales with Each Other and With Those for Big Five and NCL6

Measure	CO	NV	A	R/INE	G/C	O	
Conscientiousness (Consistency/Organization)	(.87)						
Negative Valence (vs. Non-Violativeness)	-.34	(.83)					
Agreeableness (Even Temper & Accommodating)	.26	-.43	(.80)				
Resiliency vs. Internalizing Negative Emotionality	.29	-.33	.47	(.76)			
Extraversion (Gregariousness/ Cheerfulness)	.14	-.21	.14	.38	(.84)		
Originality/Talent	.05	.00	.06	.20	.31	(.86)	
Big Five Conscientiousness	**.87**	-.37	.28	.35	.17	.12	
Big Five Agreeableness	.26	-.52	**.61**	.31	.46	.10	
Big Five Emotional Stability	.18	-.32	**.61**	**.63**	.20	.18	
Big Five Extraversion	.12	-.11	-.04	.35	**.81**	.35	
Big Five Intellect	.04	-.03	.09	.10	.23	**.72**	
Big Five Multiple R^2	*.77*	*.34*	*.60*	*.47*	*.75*	*.56*	*mean .58*
NCL6 Conscientiousness	**.86**	-.43	.29	.30	.15	.09	
NCL6 Honesty/Humility	.37	**-.55**	.47	.34	.25	.05	
NCL6 Agreeableness	.17	-.41	**.87**	.38	.17	.02	
NCL6 Emotionality	-.20	.08	-.19	**-.57**	-.17	-.43	
NCL6 Extraversion	.19	-.20	.08	.41	**.88**	.28	
NCL6 Openness	.00	-.01	.08	.19	.30	**.80**	
NCL6 Multiple R^2	*.74*	*.37*	*.78*	*.52*	*.79*	*.68*	*mean .65*
Change in R^2	*.05*	*.06*	*.20*	*.12*	*.06*	*.13*	*mean .10**

Note. N=520. Terms in parentheses reference internal consistency (values of coefficient Alpha). Correlations of .55 or more in magnitude are printed in boldface type. WCL6 - Wide-band Cross-Language Six. Multiple R^2 – Squared multiple correlation when the structural model is entered in the first block. Change in R^2 - increase in squared multiple correlation from adding NCL6 (Narrow-band Cross-Language Six) to Big Five scales as predictors. * The mean change in R-squared decreases to .09 when using instead the shorter (1995 administered only) versions of the aggregate indicators.

Table 6

Multiple Correlations of Big Five versus WCL6 Factors in Prediction of 12 Criterion Variables

Criterion Variable	Big Five R	WCL6 R	R^2 Change (*p*)	Largest significant *r*
Measured in 1999:				
Dissociative tendencies	.33*	.41*	<.001	.34 Negative Valence
Obsessive-compulsive symptoms	.40*	.43*	<.001	-.*33 Emotional Stability*
Measured in 2000:				
Borderline personality	.49*	.56*	<.001	-.48 Resiliency
Measured in 2002:				
Self-assessed health status	.13	.22*	<.01	.14 Originality/Talent
Depression (CES-D)	.31*	.39*	<.001	-.38 Resiliency
Measured in 2003:				
Medical history (mental health)	.18*	.28*	<.001	-.20 Resiliency
Medical history (other)	.17*	.23*	<.001	-.12 Extraversion
Measured in 2006:				
Lifetime smoking (>100 cigs.)	.11	.13	*ns*	
Compulsive drinking indicators	.24*	.29*	*ns*	-.24 Negative Valence
†Risk-posing behavior after drinking	.26*	.37*	<.001	.29 Negative Valence
Lawbreaking behaviors	.26*	.30*	<.05	.25 Negative Valence
Phobias	.31*	.36*	<.001	-.26 Resiliency

Note. N = 440, except for criterion variable marked with †, for which N=319 (because computed across only those individuals who responded directly to the risk-posing-behavior items, for which a prerequisite was both reporting having > 50 lifetime alcoholic drinks and reporting ever having ever had more than one drink in a week). WCL6 – inclusive-selection-based (wide-band) cross-language six dimensions. Big Five variables in the rightmost column are in italics, other variables listed (including Agreeableness) are all from the narrow-band Cross-Language Six (NCL6). Items for all predictor scales were administered simultaneously in 1995. R^2 change (*p*) is probability associated with increment in prediction from WCL6 entered in a second step after the Big Five in the first step. * $p < .05$. "Largest significant *r*" is the largest zero-order correlation between the criterion and a scale from either the Big Five or NCL6, provided $p < .05$.

Figures

Figure 1. Hierarchical clustering of 100 descriptor concepts found as a highly salient high-loading term in at least two of the eight studies. A – Agreeableness, C – Conscientiousness, NV – Negative Valence, O/T – Originality/Talent, Ex – Extraversion, R vs. INE – Resiliency versus Internalizing Negative Emotionality, accom. – Accommodating/Cooperation, PV – Positive Valence.

Figure 2. The Big Six and other recurrent dimensions as they tend to appear across lexical studies