The Lexical Big Seven: An Inclusive Psycholexical Investigation

Auke Tellegen

Professor Emeritus, University of Minnesota

Contents

Acknowledgments

I thank Niels Waller for his assistance, in Study 1, in locating suitable lexical descriptors on target pages, Beverly Kaemmer for her assistance, in Study 1, in the identification of adjectives that directly apply to a person, Gerard Saucier for suggesting an examination of the role of descriptor familiarity, the issue addressed in Study 3, and Yossef S. Ben-Porath and Annemarie Eigenhuis for responding with thoughtful questions and suggestions to an earlier version of this paper, which I like to think resulted in significant clarifications.

Abstract

According to the "lexical hypothesis", the well-known rationale for the Big Five taxonomy of major personality dimensions, the most important personality traits are more likely to be lexically encoded as one-word descriptors. Empirical tests of this conjecture when applied to the US English-language lexicon necessitated reduction of the many thousands of person descriptors to much smaller but still representative descriptor subsets. The Big Five research program achieved these reductions through a process of negative selections, or exclusions. Terms judged unfit for scientifically useful descriptions -- for example, those considered overly evaluative -- were eliminated. Unlike this exclusionary approach, the present lexical investigation was designed to be inclusive. To ensure inclusiveness but limit the research questionnaire to 400 items, the selections were spaced throughout the dictionary. Furthermore, as an alternative to the principal-component "scree test", an exploratory simple-structure method, the Smallest Factor Profile (SFP) procedure, was used to conduct a principal-factor "drop test" of the lexical hypothesis formulated explicitly as predicting a taxonomy, that is, a set of distinctly major dimensions. The results indeed indicate a taxonomy of seven and only seven major simple-structure rotated primary dimensions of personal characteristics named Positive Valence, Negative Valence, Positive Emotionality, Negative Emotionality, Disagreement Proneness, Unconventionality, and Dependability, or the Big Seven. Additional analyses, again including drop tests when applicable, show that the Big Seven are structurally robust, are recoverable from separate analyses of familiar and non-familiar descriptors, are not interpretable as artifacts, and in a new data set demonstrate the replicability of the Big Seven using previously identified markers.

The Lexical Big Seven: An Inclusive Psycholexical Investigation

Among contemporary individual-difference models of personality, the normal-personality trait constructs now well known as the Big Five have undoubtedly achieved and retained a commanding influence in the field of personality psychology. The marked increase in the volume and diversity of descriptive and explanatory Big Five investigations led John, Naumann, and Soto (2008) to declare a "paradigm shift". The Big Five had become the consensually pre-eminent descriptive dimensions of personality.

Viewed as critically contributing to this achievement was the merger, described by McCrae and John (1992), of the descriptive dictionary-based, or lexical, dimensional structure of the Big Five with Costa and McCrae's explanatory questionnaire-based dimensional Five Factor Model, or FFM. From her comprehensive review, Clark (2007) likewise concluded that among primary alternatives, the Five Factor Model was conceptually and empirically the most strongly supported structural model and was also most widely accepted as best representing the major dimensions of normal and abnormal personality. Widiger and Trull (2007) specifically recommended the Five Factor Model for implementing the shift in conceptions of personality disorders from a categorical to a dimensional model. Subsequently, Widiger and Costa (2012) again reviewed the empirical evidence supporting this dimensional model as an integrative structure capturing both normal and abnormal personality variations.

The related but independently developed Personality Psychopathology Five or PSY-5 (Harkness, McNulty, & Ben-Porath, 1995) also articulates a five-dimensional perspective on personality disorders. The current revised scales are included in the MMPI-2-RF (Ben-Porath & Tellegen, 2008/2011; Tellegen & Ben-Porath, 2008/2011). In addition, a five-dimensional model of personality disorders and a corresponding inventory, the PID-5, have been included among the Emerging Measures and Models of DSM-5 (Krueger & Markon, 2014). Integrations of the Five Factor Model and Big Five models and neuroscience are also being actively pursued (e.g., Allen & DeYoung, 2017).

Given the extraordinary impact of the Big Five and the FFM, it was perhaps to be expected that both have also been subjected to sometimes adamant criticism from alternative theoretical and methodological perspectives (e.g., Block, 1995, 2010; Epstein, 2010; Loevinger, 1993, 1994; Norem, 2010). Prominent lexical researchers have themselves periodically re-examined and raised questions about the sufficiency and cross-cultural generalizability of the lexical Big Five (e.g., De Raad et al., 2014; Saucier & Goldberg, 1998, 2001; Saucier, Hampson, & Goldberg, 2000; Saucier et al., 2014). However, while acknowledging differences of opinion and continuing scientific debate, Widiger (2017, p.1), consistent with John et al.'s (2008) earlier mentioned advocacy of the Big Five, has recently noted the continuing success of the Five Factor Model as an integrative structure capable of accommodating other predominant models of personality.

My own questions also concern lexical issues, but focus *exclusively* on the US English-language lexicon. I discuss how the Big Five were extracted from this source and outline an alternative approach, one I believe to be more attuned to identifying the content and number of major dimensions embedded in a lexicon and therefore targeted as such by the "lexical hypothesis". Before reporting the present investigation, I revisit this lexical conjecture to consider what I see as its substantive and potentially decisive data-analytic implications

The Lexical Hypothesis

Francis Galton was one of the first scientists to formulate what Goldberg (1990, 1993) has called "the fundamental lexical hypothesis". Goldberg quotes Galton's (1884) statement that "the most important individual differences in human transactions will come to be encoded as single terms in some or all of the world's languages", and that each identified term "shares a large part of its meaning with some of the rest" (Goldberg, 1990, p. 1216). I believe these two statements can be plausibly interpreted as jointly indicating that the most important individual differences are represented by the largest numbers of similar lexical terms. Given this interpretation, the two statements can be seen as foreshadowing and stimulating modern lexical research efforts first to complete the necessary first step of obtaining a representative *sample* of terms when the lexical domain to be studied happens to be very large, and then mapping the *structure* of interconnections underlying this individual-difference domain.

In the preface to their historical and conceptual analysis of lexical inquiries, Saucier and Goldberg (1996, p. 22) addressed both these points. They identified the lexical hypothesis and factor analysis among the necessary tools that only in the 20th century became available for solving the problems of (1) *sampling* human attributes and (2) *structuring* that sample of attributes. In the following section I consider both these 20th-century developments in the US.

First a terminological note. In what follows I do not use the terms "trait" or "personality" except when quoting or reporting on other researchers or reporting on instructions used in this investigation. Instead I refer to "personal characteristics". The reason is that the inclusive approach adopted in this investigation and shortly to be introduced precluded exclusions of any terms judged on conceptual grounds *not* to instantiate a characteristic of "personality" or a "trait".

Sampling the US English-Language Lexicon of Personal Characteristics

The English language abounds with terms denoting personal characteristics. Having revisited, updated, and augmented Allport and Odbert's (1936) famous list of "trait names" identified in Webster's second (1925) unabridged dictionary, Norman (1967, pp. 3-4) observed: "By rough estimate it appears that the total pool comprising both the total Allport-Odbert list and all potential additions from Webster's Third numbers nearly 40,000 terms". Goldberg (1982, p. 206) subsequently reported this number to be roughly 27,000. Clearly, lexical researchers setting out to empirically determine the number and content of major dimensions in this lexical domain would first have to recruit far smaller representative descriptor samples from its large pool.

How, then, were the reduced descriptor samples yielding the Big Five identified? The answer: through what I will call an *exclusionary* strategy. By this I mean that to be included as a lexical personal characteristic, it was necessary and sufficient *not* at some point to have been *excluded* from the initial list of candidate terms.

Goldberg's (1993) *American Psychologist* article introducing the Big Five to a general psychological readership includes among its antecedents the three major lexical research projects that ultimately converged on the Big Five. In the following I consider each of these projects, limiting descriptive details to those most relevant to the major exclusionary developments that shaped the Big Five.

Allport and Odbert (1936) made, without intending to, the first critically important contribution toward the Big Five. Goldberg (1982, p. 204) described their monograph as "an attempt to provide a truly comprehensive mapping of the domain of English personality-

descriptive terms". Allport and Odbert's (1936) initial culling yielded their final total list of 17,954 trait-names[1], which they then placed in four non-overlapping categories (pp. vi-viii) shown as columns in Part C of the monograph (pp. 38-171). In this comprehensive list of terms the four columns were headed as follows (with illustrations added from pp. 26-27): Column I, Personal Traits (e.g., *aggressive, introverted, sociable*); Column II, Temporary States (e.g., *abashed, gibbering, rejoicing)*; Column III, Social Evaluations, the longest of the four (e.g., *insignificant, acceptable, worthy*); Column IV, Metaphorical and Doubtful (e.g., *pampered, roly-poly, able*). In their Preface (p. viii) the authors clarified that "in the strictest sense of the word, only Column I contains *trait-names*". In the main text (pp. 25-26) they added that the terms in Column I, the Personal Traits, "seem to symbolize most clearly 'real' traits of personality". In other words, Allport and Odbert's four non-overlapping columns served to implement an exclusionary approach to identifying what they considered "real" personality traits. As a result of these exclusions, membership of the Personal Trait category was limited to 4,504 terms (p. 26).

Goldberg (1982), in his section (pp. 205-207) on the continued culling and classification of lexical "personality-descriptive terms", introduced the next major contribution crucial to understanding the exclusionary origins of the Big Five as follows: "To that end, Warren Norman began again, just where Allport and Odbert left off" (p. 205). As summarized by Goldberg, Norman (1967) first updated Allport and Odbert's list and somewhat enlarged it to 18,125 terms. Norman's research team then proceeded to apply four criteria (labeled Obscurity, Ambiguity, Appearance, and Pure Evaluation) for excluding terms. The resulting removal of 10,831 terms, which Goldberg (p. 206) characterized as "dross", about 60% of the initial pool, left Norman and his team with 7,294 terms, which they further divided into three categories. Briefly, they were: I, Stable Traits (e.g., *calm, autocratic, cool*); II, Temporary States and Activities (e.g., *confused, demanding, despondent, bawling)*; III, Social Roles, Relationships, and Effects (e.g., *dominant, attractive, droll, appetizing*); for more details, see Goldberg (1982, p. 207). Only the 2,797 Stable Trait terms were retained, and then only provisionally. Subsequently, as reported by Goldberg (1990, p. 1217), Norman collected additional descriptive information on each of these terms (for example, the percentage of student participants indicating they did not know its meaning), and used this information to exclude ambiguous and unfamiliar adjectives. These exclusions, combined with the removal of nouns, shrank the number of retained lexical terms to 1,431.

In turn, following up on Norman's contributions and responding to early questions from Waller and Ben-Porath (1987), Goldberg (1990) conducted the three factor-analytic studies from which the now familiar Big Five structure finally and repeatedly emerged. What sorts of variables did Goldberg analyze? All three sets of measures consisted of composites made up of adjectives that had been judged similar in meaning. In Study 1 (pp. 1217-1222), Goldberg analyzed 75 composite scores corresponding to 75 content categories into which Norman had previously sorted the 1,431 lexical items on the basis of how he understood their similarities. For Study 2 (p. 1222), composites or clusters were more consensually assembled with the aid of dictionaries, synonym finders, and with the item-desirability ratings Norman had provided. Using these tools, Goldberg assembled and analyzed "a set of 133 synonym clusters based on 479 commonly used trait adjectives" (p. 1222). In Study 3 (pp. 1222-1223), the 133 clusters, on the basis of a series of internal-consistency analyses, were in turn replaced with an improved set of 100 clusters (p. 1223). I consider the 479 "commonly used trait adjectives" that formed the

[1] Allport and Odbert (1936) reported that 17,953 terms were listed in their monograph, but the correct number is 17,954 (Goldberg, 1982, p. 206, footnote 1)

basis of the 133- and 100-variable sets of clusters to again be the result of an exclusionary approach -- the exclusion of terms judged *not* to be "commonly used trait adjectives".

In sum, the item composites analyzed in Goldberg's three studies consist in each case of quasi-synonymous adjectives from lists that had been selected through exclusionary methods. This observational basis and the proven replicability of the lexical Big Five led me to conclude: *The Big Five content domain is the replicated outcome of an exclusionary approach to identifying the lexical body of personal characteristics to be sampled.*

Structuring (Mapping) the US English-Language Lexicon of Personal Characteristics

According to Galton's earlier quoted lexical hypothesis, each lexical term encoding one of the major human traits shared "a large part of its meaning" with some of the terms encoding the same major trait. In other words, Galton's hypothesis was equally consistent with undifferentiated and with cluster-like quasi-categorical distributions of meaning proximities. However, once exploratory factor analysis with simple-structure rotations targeting cluster-like constellations in factor space had become popular[2], 20th-century lexical researchers were among those who adopted a quasi-categorical model. Because the Big Five structure was discovered under this model, it has repeatedly been referred to as a "taxonomy" (e.g., by Goldberg, 1993; John, 1990; John, Naumann, & Soto, 2008; John & Srivastava, 1999; and Norman, 1963), a categorical term.

The label "quasi-categorical" is meant to convey that the terms "taxonomy" and "categorical" are not to be taken literally. Goldberg (1992, p. 27) notes that lexical trait terms are "not neatly clustered in multivariate space", but like stars on a clear night "can be found nearly everywhere one looks in semantic space". However, he then goes on to stress (p. 27) that, like the stars, there are "some semantic galaxies within the widespread distribution of personality descriptors" that as such "can be viewed as natural categories (e.g., Rosch, 1978)".

Saucier and Goldberg (1996, pp. 26-27) likewise suggest that each of "the most important phenotypical personality attributes" is to be viewed as "a dense cluster of loosely synonymous terms", and collectively as non-redundant "bundles of related concepts". Citing a "considerable body of research", they characterize the Big Five factor structure as a "very promising candidate to fill the role of a set of 'reasonable locations' for possible universal lexical dimensions" (p. 36). They actually conclude from the published evidence that any additional factors, which they label "outlier dimensions", have "*far* fewer adjectives defining them" and "thus the lexical hypothesis would suggest they are less important" (p. 39). However, they also note that the Big Five factors themselves vary in importance and replicability (p. 39), and that "personality descriptors are not organized neatly into tight and discrete clusters of variables" and instead "most variables fall in the interstitial areas between the factor poles" (p. 40). Yet "as a general rule", each Big Five factor "represents a major concentration in a continuous distribution of attributes in descriptive space" (p. 40).

The investigations Saucier and Goldberg (1996) cite in support of this conclusion include the three reported in Goldberg's (1990) "Alternative Description of Personality" introduced in the previous section. Focusing on the US lexicon, as does the present investigation, these studies were designed to provide a compelling corroboration of the lexical Big Five. I introduced

[2] Cattell (1973, p. 125) was a notable dissenter. He referred to simple-structure rotations such as Varimax as "cluster-chasing", and instead favored simple-structure rotations maximizing the "hyperplane count", i.e., the percentage of factor loadings not exceeding |.10|

Goldberg's article in that section because, as noted there, it includes an account (summarized on p. 1217) of how he dealt with the preliminary task of identifying in the US English-language of "trait-descriptive adjectives" much smaller but still representative descriptor *samples* amenable to factor-analytic investigations, namely, by applying exclusionary criteria. I reintroduce Goldberg (1990) in this section focusing on *structure* because the author also reports here the three consequential series of exploratory factor analyses with simple-structure rotations that he found to converge on the Big Five, namely, his analyses of the 75 synonym cluster scores Norman had previously assembled from his reduced set of 1,431 terms, and of his own sets of 133 and 100 synonym cluster scores. In his account of the first two studies Goldberg (pp. 1217-1222) in addition reported extracting and rotating *more* than five dimensions. The results provided evidence that the Big Five dimensions were not only *prominent* but also *robust* (i.e., remained prominent when additional factors were extracted and rotated), which are the two properties specifically and systematically documented in, respectively, Study 1 and Study 2 of the present lexical investigation.

Jointly, Goldberg's (1990) landmark study and Saucier and Goldberg's (1996) integrative review advanced a factor-analytic version of the original lexical hypothesis. The vivid depictions and detailed descriptions encountered there led me to conclude that each major lexical dimension was thought to be identifiable in factor space as a loose but dimensionally distinctive cluster of descriptors. More specifically, I concluded that from an inclusive perspective these two contributions could in factor-analytic terms be interpreted as encouraging or even implying the following lexical conjecture. *Among the one-word lexical terms denoting personal characteristics (i.e., those that are informative about "what kind of person" someone is thought to be) all and only the characteristics that are the most important will be recovered through exploratory factor analysis as prominent simple-structure rotated dimensions"*. This, in any case, is the conjecture adopted in the five studies that make up the present investigation. Study 1 includes in its Data Analysis and Results section a description of the exploratory factor-analytic procedure I have labeled the Smallest Factor Profile or SFP procedure and have used to test this simple-structure version of the lexical hypothesis. To select a subset of representative lexical terms consistent with this approach but small enough for data-collection and data-analytic purposes, a "page-sampling" approach was adopted and is described in the first of the five studies reported here. The 400-item self-report inventory assembled in this manner was used or drawn on in each of these five studies to pursue the following five main objectives (using the same data base in the first four studies and a new data set in the fifth):

(1) determine the number of major lexical dimensions through analyses designed to implement the simple-structure version of the lexical hypothesis, and examine the content of these lexical dimensions,

(2) examine the structural robustness of these dimensions, named the Big Seven,

(3) rule out systematic differences between familiar and unfamiliar lexical descriptors,

(4) rule out measurement artifacts, and

(5) evaluate the replicability of the Big Seven structure.

In the following, unless specified otherwise: (1) the factor-analytic findings are based on product-moment item inter-correlations and refer to (orthogonal) varimax simple-structure rotated principal factors; (2) "markers" or "marker items" of a given factor are items with larger absolute loadings on that factor than on any other factor from the same rotated principal-factor

solution; and (3) the "strongest" markers of a given factor are the ones with absolute loadings higher than those of the other markers of the same factor.

Study 1. Determining the Number and Examining the Content of the Major Lexical Dimensions[3]

Development of a 400-item Self-Report Research Questionnaire (S400)

The first step was to assemble an adequate lexicon-based questionnaire, one that would inclusively sample the domain of adjectival personal characteristics. Accordingly, any adjective qualified for potential selection if one of its listed lexical definitions would qualify for inclusion in an updated edition of Allport and Odbert's (1936) four lists of "trait names". In other words, all adjectives describing "Personal Traits", as such referring to clearly dispositional terms (Category I), qualified. So did any adjective, A, that under Allport and Odbert's rules would be assigned to "Temporary States" (Category II), because I considered these terms, clarified in this investigation as "tends to be A" or "is often A", to be inherently a personal characteristic. Neither was any adjective excluded on account of being considered a "Social Evaluation", i.e., judgmental, evaluative (Category III), since such terms play a prominent role in perceptions of self and others. "Metaphorical and Doubtful" terms (Category IV) were also retained, clarified by their lexical definitions. I assumed that the domain of adjectival personal characteristics would require relatively few -- no more than nine or ten -- major distinctive dimensions. Based on this working assumption, 400 adjectives, appropriately sampled, were expected to yield enough item-level representations of all major dimensions. Rather than using Allport and Odbert's (1936) and Norman's (1967) exhaustive lexical sources, I relied on the American Heritage Dictionary of the English Language (1982/1955), which I expected to be sufficiently comprehensive, as did Goldberg (1982), who used an earlier edition.

To achieve the target of 400 items I adopted a "page-sampling" procedure. The 1,342 pages comprising the lexical substance of the dictionary were divided into 53 25-page sections plus one 17-page section. In each 25-page section, 7 or 8 pages, chosen randomly with the constraint of being spaced at least two pages apart, were designated "target pages"; in the same way, 3 target pages were selected from the final 17-page section. On each of these pages Niels Waller, who participated as the second judge, and I first independently examined that page and then consensually either did not identify or did identify and select (1) the first adjective on that page denoting a personal characteristic and (2) the first listed lexical definition of that characteristic. In this manner, 408 questionnaire items were first assembled, of which 8 items were then randomly dropped.

[3]Portions of the subject-matter covered in Study 1 are referred to or are partially reported in a number of previous publications, including Tellegen (1993), Tellegen and Waller (1987), and Waller (1999). Unique to this investigation are a reformulation of the lexical hypothesis and the introduction to and applications of the empirical tests provided by the Smallest Factors Profile (SFP) method included in Study 1. Studies 2, 3, 4, and 5 have not previously been reported.

The 400 items selected in this manner for the project's research questionnaire, the S400, were worded to conform closely to their dictionary entries and definitions. Changes were made only to the extent necessary to avoid possible unclarities. This is the reason why a little over three quarters (311) of the questionnaire items include the selected lexical adjective as well as a definition, but why the other items do not include the former. Example: to avoid semantic unclarity, the questionnaire item "open to change, adaptable", while defining the lexical adjective "labile", does not include "labile" itself. In other cases, language such as "apt to be…", "tends to be…", "given to…" was added to ensure interpretation of the item as dispositional. Additional adjustments were made as needed. Example: the lexical target term "semiformal" was changed to "formal" to avoid the ambiguity of having to apply the graded rating scale adopted for the S400 items (see below) to a dictionary entry ("semiformal") that itself included a quantifier.

The following four-choice response format was chosen for the S400: 1. Very inaccurate, very uncharacteristic; 2. Slightly inaccurate, slightly uncharacteristic; 3. Slightly accurate, slightly characteristic; 4. Very accurate, very characteristic. The participants were asked to respond to each item according to how well it applied to her/him "in general rather than just this moment". The 400 items were divided into two booklets, both of which were to be completed by each participant.

Participants and Instructions

Undergraduates taking the introductory psychology course volunteered to participate in one of several data collection sessions scheduled throughout the academic year, for which they received course credits. Of the 585 individuals who began the session, 21 did not finish and did not hand in an answer sheet for the second booklet; another 21 returned answer sheets indicative of one or more skipped booklet pages or otherwise lacking a total of more than eight answers. The remaining 543 participants served as the study sample of 229 men, 279 women, and 35 individuals who had not identified their gender.

Before the booklets were handed out, participants were informed of the basic purpose of the study, namely, to take a fresh look at how individuals describe self and others. They were told that for this reason the research questionnaire was composed of representative personality adjectives taken directly from the dictionary. It was pointed out that some items would therefore seem bookish and would occasionally include unfamiliar words, but that in all cases careful reading of the whole item would make the meaning clear. It was also emphasized that any data provided would be kept strictly confidential, that the sole purpose of the study was to analyze the data set as a whole, not to study individual records.

Data Analysis and Results

Needed at this point was a test of the adopted simple-structure version of the lexical hypothesis. Its corroboration would accomplish the necessary first step of disclosing the *number* of major lexical dimensions. To make this determination, instead of relying on the familiar "scree test", based on the slope of 30 eigenvalues plotted in Figure 1[4], and arguably suggesting five major dimensions, I completed the series of factor analyses named earlier the Smallest Factors Profile (or SFP) procedure. The critical results from this set of analyses, each in turn conducted on the S400 data base, serve also to describe the method and are plotted in Figure 2. Displayed, in order, are 29 factor sizes (sums of squared factor loadings): the size of

[4]In all figures, *open* circles represent *unrotated* principal components or factors and *filled* circles represent *rotated* common factors

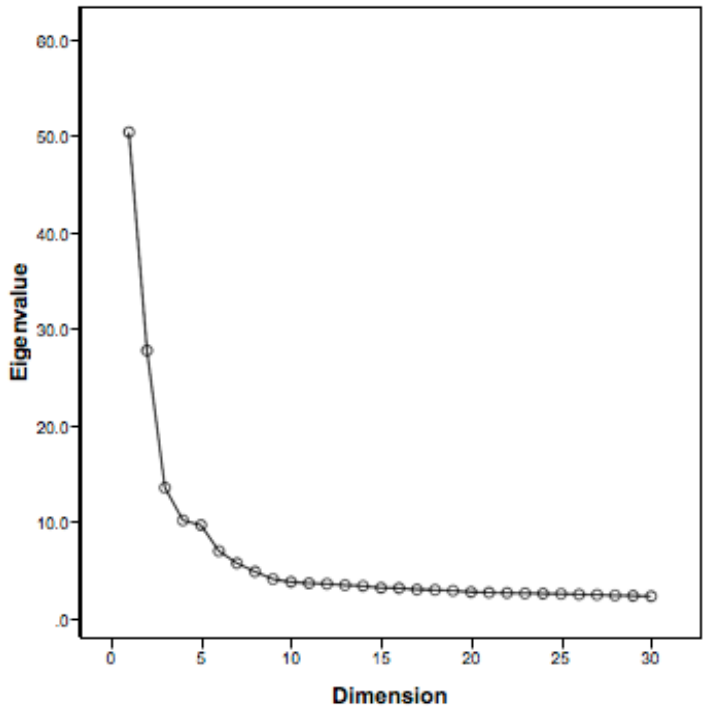

Figure 1. Study 1: Eigenvalue plot from the principal-component analysis of the S400 psycholexical inventory items.

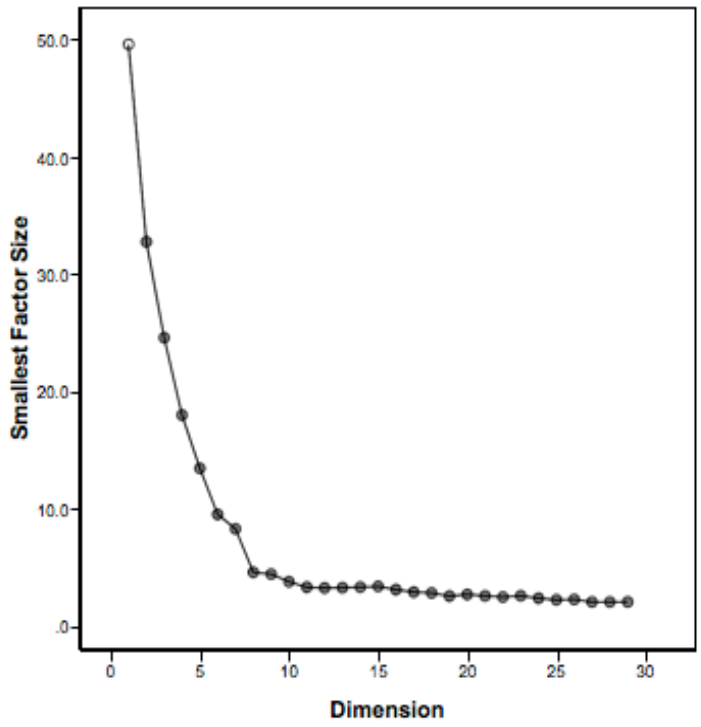

Figure 2. Study 1: Factor-size plot, produced by the SFP procedure, of the first (unrotated) factor (open circle) and subsequent smallest rotated factors (filled circles) resulting from principal factor analyses of the S400 psycholexical inventory items.

the (unrotated) first principal factor (open circle) and of each of 28 rotated-factor sizes (filled circles). The latter are: the size of the smallest rotated factor from the 2-factor solution, from the 3-factor solution, and so on, ending with the smallest rotated factor from the 29-factor solution. The resulting 29-point graph served as the basis for a "drop test," that is, to determine whether, in accordance with the simple-structure version of the lexical hypothesis, a final major drop was discernible, and, if so, at what point. I concluded that such a drop had occurred after seven dimensions, evidence therefore of seven major lexical factors, the "Big Seven".

Figure 1 features, of course, one more eigenvalue than Figure 2 features common-factor sizes. The same difference will be noted in all three subsequent pairs of plots.

Interpretation of the Big Seven

Having arrived at seven major lexical dimensions, it was possible to undertake the task of examining and interpreting these Big Seven. To explain and allow the reader to judge these interpretations, I begin by introducing the 10 strongest markers, numbered 1 through 10, of each factor. Table 1 presents, along with additional information, all 70 markers in abbreviated form, each followed by its seven rotated-factor loadings. In the text below, the same markers are reproduced in quotes and as presented in the S400 inventory, except that for greater clarity semicolons have been replaced with commas within items and semicolons have been inserted between items, and the items are presented in full instead of abbreviated as they are in Table 1.

Although negative Big Seven markers (those with negative marker loadings) are proportionally few (50 out of 400), tend to be weaker, are therefore usually not included in Table 1, and factorially more complex, they too include a number of informative descriptors. I am therefore also providing information about the three or fewer strongest negative markers of all six factors of which the strongest ten markers are all positive (Big Seven Unconventionality is the exception). To facilitate interpretation of these added complex item-level indicators, each is not only reproduced in full, but the text is followed by the labels of and loadings on those two Big Seven factors on which it received its (always negative) marker loading and its second-strongest loading. In each subset of positive and negative markers each item is listed in the order of its absolute loading on the factor of which it is a marker.

In addition, to facilitate comparisons among the Big Seven and with other construct sets, the content of the ten strongest markers of each Big Seven dimension is displayed as follows. In each set of ten S400 markers, *all adjectives directly applied to a person*, including repetitions and irrespective of whether the adjective appears alone or in a phrase separated by commas or conjunctions ("or", "and"), are shown in italics[5]. Among these italicized terms, the ones that *in addition* were placed in Allport and Odbert's (1936) Category III (Social Evaluations), their large subset of evaluative terms, are shown in bold italics, and the ones *not* included in their list of 17,954 categorized trait names are shown in parentheses. For each Big Seven dimension, I summarize by reporting both the total number of these Category III adjectives (the terms shown

[5]Beverly Kaemmer and I first independently completed the task of identifying among each of the ten strongest markers of each Big Seven dimension the adjectives that directly applied to a person. On four of the dimensions our choices were in complete agreement. Across the remaining three dimensions a total of four initial discrepancies were encountered, which were resolved through discussion.

Table 1.
Factor Loadings and Unfamiliarity Ratings of the 10 Strongest Marker Items of Each of the Big Seven Dimensions of the 400-item Lexical Inventory

Items	Factor Loadings							Unfamiliarity
Negative Valence (NV)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
1. lousy*	**.68**	-.13	-.04	.09	.03	.00	.04	2.53
2. (heinous) reprehensible*	**.67**	-.08	.01	.04	.04	-.02	-.05	2.60
3. (vicious) depraved*	**.67**	-.12	-.04	.02	-.05	.05	-.05	1.34
4. pinheaded	**.67**	-.05	-.15	.13	-.16	-.02	-.04	1.79
5. (cursed) detestable*	**.66**	-.12	-.06	.04	.03	.03	.03	2.75
6. wicked*	**.66**	-.14	-.03	-.02	.00	-.06	-.06	1.51
7. jackbooted*	**.66**	-.14	.00	.03	-.01	-.07	-.08	3.85
8. oafish	**.61**	-.03	-.13	.18	-.14	.01	.03	2.53
9. squalid*	**.60**	-.19	-.06	.07	.01	.02	-.06	3.74
10. lunatic	**.60**	-.08	-.10	.15	-.04	.20	.08	2.02
Positive Emotionality (PEM)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
1. (crank) spirited*	-.06	**.73**	.11	-.05	-.09	-.02	.043	3.23
2. (gingery) vigorous*	-.07	**.71**	.21	-.16	-.02	.03	.06	3.27
3. bouncing*	-.09	**.70**	.16	-.00	.04	-.02	-.07	2.98
4. (on fire) enthusiastic*	-.09	**.67**	.19	-.08	-.08	.07	.12	3.02
5. alive*	-.16	**.62**	.15	-.07	-.05	.10	.02	1.81
6. talkative*	.03	**.61**	.02	-.04	.14	.10	-.04	1.36
7. jovial*	-.09	**.60**	.15	-.03	-.09	.09	.10	2.15
8. spirited*	-.09	**.59**	.25	-.18	.04	.18	.10	1.34
9. boon*	-.02	**.58**	.15	-.02	-.02	.07	-.04	3.88
10. (rosy) optimistic*	-.22	**.57**	.14	-.12	-.12	.01	.12	2.28
Positive Valence (PV)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
1. (prime)excellent*	-.17	.11	**.65**	-.12	.02	.12	.22	3.17
2. high-level*	-.01	.16	**.64**	-.09	.04	-.02	.07	3.56
3. (premier) important*	-.00	.20	**.63**	-.13	.08	.06	.11	3.28
4. sensational*	-.02	.16	**.62**	-.15	.02	.07	.10	2.06
5. tiptop*	-.12	.13	**.61**	-.16	.02	.13	.19	3.30
6. (uppermost) influential	.12	.15	**.58**	-.20	.11	-.00	.14	3.55
7. copacetic*	-.09	.04	**.58**	-.14	.03	.09	.19	3.46
8. especial*	.05	-.02	**.57**	-.13	.08	.23	.03	3.56
9. momentous*	-.16	.25	**.57**	-.10	-.05	.07	.15	2.87
10. (sovereign) powerful	.07	.13	**.56**	-.19	.16	.04	.07	2.94
Negative Emotionality (NEM)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
1. (queasy) easily troubled*	-.03	-.08	-.11	**.55**	.19	.04	.01	3.23
2. guidable*	.00	.10	-.03	**.54**	-.21	-.22	-.04	3.85
3. perplexed*	.25	-.12	-.20	**.54**	.07	.03	.02	2.09
4. maneuverable*	.29	-.01	-.09	**.54**	-.17	-.08	-.03	3.04
5. (psychoneurotic) anxious*	.18	-.18	-.15	**.53**	.10	.14	.05	1.72

Table 1. Factor Loadings and Unfamiliarity Ratings, continued

Items	Factor Loadings							Unfamiliarity
	NV	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
6. bluffable*	.11	.01	-.05	**.52**	-.04	-.08	-.08	2.54
7. (trembly) anxious*	.12	-.03	-.11	**.51**	.12	.02	.02	3.26
8. (foggy) bewildered	.30	-.08	-.11	**.49**	-.00	.09	.03	2.65
9. Ill at ease*	.18	-.29	-.13	**.49**	.08	-.03	.04	2.15
10. (funky) panicky*	.25	-.03	-.15	**.49**	.04	-.06	.06	2.00
Negative Emotionality (NEM)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
1. (queasy) easily troubled*	-.03	-.08	-.11	**.55**	.19	.04	.01	3.23
2. guidable*	.00	.10	-.03	**.54**	-.21	-.22	-.04	3.85
3. perplexed*	.25	-.12	-.20	**.54**	.07	.03	.02	2.09
4. maneuverable*	.29	-.01	-.09	**.54**	-.17	-.08	-.03	3.04
5. (psychoneurotic) anxious*	.18	-.18	-.15	**.53**	.10	.14	.05	1.72
6. bluffable*	.11	.01	-.05	**.52**	-.04	-.08	-.08	2.54
7. (trembly) anxious*	.12	-.03	-.11	**.51**	.12	.02	.02	3.26
8. (foggy) bewildered	.30	-.08	-.11	**.49**	-.00	.09	.03	2.65
9. Ill at ease*	.18	-.29	-.13	**.49**	.08	-.03	.04	2.15
10. (funky) panicky*	.25	-.03	-.15	**.49**	.04	-.06	.06	2.00
Disagreement-Proneness (DSP)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
1. stiff-necked*	.12	.01	.06	-.11	**.63**	.01	-.01	2.79
2. (wayward) willful*	.02	.01	.02	-.01	**.59**	.10	.02	3.21
3. dogged*	-.04	-.04	-.00	-.05	**.58**	.20	.04	3.46
4. opinionated*	.03	.04	.07	.04	**.56**	-.03	.02	1.04
5. obstinate*	.02	-.03	.05	.05	**.54**	.12	.10	2.15
6. bullheaded*	-.06	.05	.01	-.01	**.53**	.01	-.03	1.17
7. (pettish) ill-tempered*	.26	-.15	.02	.12	**.47**	-.04	.01	3.40
8. dubious*	.07	-.16	.09	.09	**.45**	.20	.00	2.56
9. choleric*	.21	-.02	.02	.05	**.45**	-.07	.00	3.44
10. sarcastic	.02	-.01	.10	.07	**.43**	.27	-.05	1.13
Unconventionality (UCV)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
1. divergent*	.11	-.06	.02	-.02	.22	**.58**	-.01	2.81
2. heterodox*	.14	-.13	.12	-.01	.07	**.57**	.00	3.85
3. orthodox*	-.04	.00	-.06	.26	.05	**-.52**	.12	2.06
4. buttoned-down*	.12	-.14	-.07	.18	.06	**-.49**	.16	2.96
5. progressive*	.01	.02	.15	.00	-.07	**.48**	.05	2.04
6. weird	.30	.03	.13	.13	.20	**.47**	-.05	1.23
7. kicky*	.15	.24	.25	-.05	.16	**.47**	-.03	3.65
8. (sans-culottic) revolutionary*	.32	-.08	.12	-.08	.01	**.46**	.02	3.96
9. comme il faut*	-.19	.05	.14	.18	-.06	**-.44**	.21	3.94
10. (semiformal) formal*	-.12	.04	.00	.23	-.07	**-.43**	.20	2.91

Table 1. Factor Loadings and Unfamiliarity Ratings, continued

Items	Factor Loadings							Unfamiliarity
Dependability (DPD)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*	
1. reasonable*	-.29	.03	.15	-.10	-.09	-.06	**.47**	1.15
2. deliberate*	-.17	-.00	.16	-.07	-.08	-.20	**.46**	2.33
3. staunch*	-.11	.12	.14	-.08	.14	.03	**.44**	3.02
4. ratiocinative*	-.09	-.10	.16	-.04	.02	-.22	**.42**	3.83
5. tried	-.27	.18	.10	-.07	-.00	-.02	**.41**	3.26
6. (seamless) consistent*	-.13	.01	.14	-.15	-.05	-.19	**.41**	3.09
7. duteous*	-.11	.09	-.01	.22	-.10	-.26	**.39**	2.88
8. attentive*	-.19	.13	.16	-.15	-.14	-.05	**.38**	1.27
9. (prosaic) straightforward*	-.11	.08	.11	-.18	.12	.14	**.37**	3.85
10. methodical*	-.07	-.05	.06	.10	.18	-.16	**.37**	1.77

Note. N = 543; The first column lists 70 lexical terms from the 400-item research questionnaire, which also included the *first lexical person-descriptive definition* of each term; only when the lexical terms are shown here in parentheses are they followed by an abbreviated version of the lexical definition; in these cases it was felt that research participants might interpret the two terms as discrepant in meaning (in the research questionnaire all such lexical terms were omitted and only the lexical definitions were included). Item i is defined as a "marker" of factor F if its absolute loading on F is larger than on any other factor; the table displays the 10 strongest markers of each factor in the *orthogonal (varimax) 7-factor rotation, based on the product-moment item intercorrelation,* with all strongest loadings shown in boldface; "*" identifies among these 70 markers the items that are also among the 10 strongest markers in the *15-factor rotation*. The unfamiliarity values of each lexical term are averages based on ratings of how sure the raters (N = 47 or 48) were of its meaning, using a 4-point scale ranging from "very sure" (1) to "don't know" (4). See text for additional details.

in bold italics) and the total number included among Allport and Odbert's (1936) categorized trait names (all italicized terms not in parentheses). This information about Category III terms is critical in comparisons of the Big Seven with construct sets (including those advocated by other psycholexical researchers) from which evaluative terms have generally been excluded, a topic I return to in one of the concluding subsections of this article (Relations to Other Constructs).

Negative Valence (NV). This is by a substantial margin the largest lexical factor, accounting for 28.2 % of the total Big Seven common variance. As worded in the S400 inventory and as numbered in Table 1, its 10 strongest markers (all positive) are: 1. "***Nasty***, ***contemptible***, ***lousy***"; 2. "***Wicked*** or ***reprehensible***"; 3. "***Evil*** or ***immoral***, ***depraved***"; 4. "***Stupid***, ***pinheaded***"; 5. "Deserving to be cursed, ***wicked***, ***detestable***"; 6. "*Vicious*, ***depraved***, ***wicked***"; 7. "Cruelly and violently ***oppressive***, *(jackbooted)*"; 8. "***Stupid***, ***oafish***"; 9. "Having a dirty or wretched

appearance, ***squalid***"; and 10. "Suffering from lunacy, ***insane***, ***lunatic***". **Summary**: 20 of 21 categorized terms are evaluative. The two strongest negative NV markers are: 1. "Receptive to learning, teachable", NV: -.44, DPD: .32; and 2. "Of sound mind, sane", NV: -.40, UCV: -.19. The label "Negative Valence" is intended to reflect the wide range of extremely negative adjectival attributes represented by the major markers of this dimension. The two negative markers support a construal of low NV as including intellectual and mental soundness.

Positive Emotionality (PEM). This factor, the second largest, accounts for 18.3% of the Big Seven common variance. As worded in the S400 inventory and as numbered in Table 1, its 10 strongest markers (all positive) are: 1. "*Lively*, *cheerful*, *spirited*"; 2. "*Lively*, *vigorous*"; 3. "*Spirited*, *lively*, *bouncing*"; 4. "Filled with enthusiasm or excitement"; 5. "Full of life, *lively*, *alive*"; 6. "Having an inclination to talk, *loquacious*, *talkative*"; 7. "Marked by hearty conviviality, *jovial*"; 8. "Full of animation, vigor, or courage, *spirited*"; 9. "*Jolly*, *convivial*, ***boon***"; and 10. "*Cheerful*, *optimistic*". **Summary**: 1 of 19 categorized terms is evaluative. Of the seven dimensions, PEM subsumes the largest number of negative markers: 21 in all. The three strongest are: 1. "Not speaking, silent, tacit", PEM: -.55, NEM: .24; 2. "Lacking in liveliness or interest, dull", PEM: -.54, NEM: .27; and 3. "Not affectionate or friendly, aloof, cold", PEM: -.53, NV: .28.

The label "Positive Emotionality" was chosen to reflect the positive and energetic temperamental quality of its main markers. Positive markers 6 and 7 above highlight PEM's interpersonal aspect and its relatedness to Big Five Extraversion as measured by Goldberg's (1992) 100-item Big Five inventory, which includes "talkative" and "untalkative" among its markers. However, I do not think the familiar Extraversion label fully captures PEM's distinctive affectively colored substance. The three negative markers underscore the lack of social engagement and spirit associated with low PEM.

Positive Valence (PV). This third largest lexical dimension accounts for 15.1% of the Big Seven common variance. As worded in the S400 inventory and as numbered in Table 1, its 10 strongest markers (all positive) are: 1. "***Excellent***, of high quality or value"; 2. "***Important***, *(high-level)*"; 3. "Having status, ***important***"; 4. "***Outstanding, spectacular, sensational***"; 5. "***Excellent, first-rate, tip-top***"; 6. "High in position, place, rank, or influence"; 7. "***Excellent, first-rate***, *(copacetic)*"; 8. "Standing above or apart from others, ***exceptional, especial***"; 9. "***Important*** or ***significant***, ***momentous***"; and 10. "Having rank or power". **Summary**: all 16 categorized terms are evaluative. PV has one negative marker: "Not special or outstanding, average, run-of-the-mill", PV: -.31, NEM: .28.

Together, the positive markers convey a generalized sense of excellence, of being first-rate and exceptional. The label "Positive Valence" for this dimension is intended to accommodate its breadth. To some readers "Narcissism" may have occurred as a fitting alternative to "Positive Valence". However, the passive self-absorption associated with its mythological provenance, as well as the sense of not being duly recognized that is associated with its contemporary diagnostic use, argue against the Narcissism label. In one of the concluding subsections of this article (The Clinical Relevance of the Big Seven) I return to the narcissism topic. The negative marker suggests viewing low PV as a perceived lack or absence of outstanding qualities.

The NV and PV dimensions clearly represent the subset of evaluative terms excluded from the body of lexical terms that were targeted by the Big Five.

Negative Emotionality (NEM). This dimension, the fourth largest, accounts for 13.9 % of the Big Seven common variance. As worded in the S400 inventory and as numbered in Table

1, its 10 strongest markers (all positive) are: 1. "Easily troubled"; 2. "Readily controlled or influenced, *(guidable)*"; 3. "*Puzzled*, *bewildered*, *confused*, *perplexed*"; 4. "Readily manipulated by others, *(maneuverable)*"; 5. "*Psychoneurotic*, suffering from anxiety and fears"; 6: "Apt to be misled, apt to be deceived, apt to be hoodwinked, ***bluffable***"; 7. "Feeling or expressing fears or anxiety"; 8. "*Bewildered*, *perplexed*"; 9. "*Nervous* and *uncomfortable*, *(ill-at-ease)*"; and 10. "*Frightened*, *panicky*". **Summary**: 1 of 12 categorized terms is evaluative. NEM's three negative markers are 1: "Decisive and firm, resolute", NEM: -.30, DPD: .30; 2. "Possessing strength of character, backboned", NEM: -.29, DPD: .27; and 3. "Courageous, lionhearted", NEM: -.28, PV: .26.

NEM, like PEM, has a strongly affective quality. Here, the predominant disposition is to have emotionally negative experiences, hence the "Negative Emotionality" label. The seven positive NEM markers 1, 3, 5, and 7 through 10 above refer directly to such experiences. Viewed together with positive markers 2, 4, and 6, the three negative markers, while again factorially complex, can be seen as contributing a plausible further explication of low NEM as a tendency not to feel weak and vulnerable.

The meaning of several of the attributes in positive markers 5, 7, 9, and 10 matches that of the adjectival Neuroticism markers "anxious", "nervous", and "fearful" in Goldberg's (1992) 100-item Big Five inventory, and positive Big Seven marker 5 includes the adjective "psychoneurotic". I nonetheless believe "Negative Emotionality" is preferable over the widely used but in my view antiquated and descriptively uninformative "Neuroticism" label.

The remaining three Big Seven factors, although substantially smaller than the preceding four "valenced" ones, are in view of the earlier reported drop test results still to be considered major dimensions.

Disagreement Proneness (DSP). This lexical dimension accounts for 8.9% of the Big Seven common variance and is the fifth largest. As worded in the S400 inventory and as numbered in Table 1, its 10 strongest markers (all positive) are: 1. "*Stubborn*, *unyielding*, *stiff-necked*"; 2. "*Stubborn*, *willful*"; 3. "Not yielding readily, *willful*, *stubborn*, *dogged*"; 4. "Holding stubbornly and often unreasonably to one's own opinions, *opinionated*"; 5. "Stubbornly adhering to an attitude, opinion, or course of action, *obdurate*, *obstinate*"; 6. "*Stubborn*, *obstinate*, ***bullheaded***"; 7. "*Ill-tempered*, *peevish*, unreasonably *irritable*"; 8. "Reluctant to agree, *skeptical*, *dubious*"; 9. "Easily angered, *(bad-tempered)*, *choleric*"; and 10. "Given to expressing sarcasm, *sarcastic*". **Summary**: 1 of 21 categorized terms is evaluative. The strongest negative DSP marker is "Harmonious, agreeing, concordant", DPD: -.36, PEM: .25.

Pondering a label, I considered "Disagreeableness", but also took note of the labels of both poles of Big Five Factor II, namely, "Pleasantness or Agreeableness" (Goldberg, 1992, p.42), or, alternatively, "Agreeableness or Pleasantness" (Goldberg 1993, p. 27) vs. "Disagreeableness" (Goldberg, 1990, p.1221). For this Big Seven dimension, to put more emphasis on dissent and less on unpleasantness given the overall content of its major markers, I eventually chose "Disagreement-Proneness" in preference to "Disagreeableness". The content of the negative marker suggests interpreting low DSP as a consensus-oriented interpersonal disposition, consistent with the proposed interpretation of the positive DSP markers.

Unconventionality (UCV). This Big Seven dimension, the sixth largest, accounts for 8.7% of the Big Seven common variance. It includes four negative markers among its ten strongest ones, so that its interpretation did not call for listing additional negative markers.

The six positive markers, as worded in the S400 inventory and as numbered in Table 1 are: 1. "Departing from convention, (*divergent)*"; 2. "Holding unconventional, unorthodox

opinions, *heterodox*"; 5. "Promoting or favoring political or social reform, ***liberal***, ***progressive***"; 6. "Of an ***odd*** and ***inexplicable*** character, ***strange***, ***fantastic***, ***weird***"; 7. "Providing a kick by being ***unusual*** or *unconventional*, *(kicky)*"; and 8. "*Revolutionary*, politically *radical*". The four negative markers, again as worded in the test booklet, and numbered as in Table 1, are 3. "Adhering to what is commonly accepted, customary, or *traditional*, *orthodox*"; 4. "*Conservative*, *conventional*, or *unimaginative*, (*buttoned-down)*"; 9. "In accordance with conventions or accepted standards, ***proper***, *(comme-il-faut)*"; and 10. "Following or adhering to acceptable norms, conventions, or regulations, *formal*". **Summary**: 9 of 19 categorized terms are evaluative.

The six positive Unconventionality markers cover a broad range of attributes, including unconventionality, holding politically radical views, being strange and weird. Lack of imagination and socially conservative attitudes are among the attributes represented by the negative markers. They indicate some overlap with the Intellect dimension of Goldberg's (1992, Table 3) 100-item lexical Big Five inventory, which includes the polar-opposite Intellect markers "unimaginative" vs. "imaginative". Goldberg (1993, p. 30, footnote 5) accordingly suggested "Imagination" as a possible alternative label for this dimension. The overlap is somewhat greater with the Fantasy and Values facets of the Openness dimension of the Five Factor Model as operationalized by the NEO PI-R (Costa & McCrae, 1992). However, both overlaps are limited. The full content range of this Big Seven dimension suggested labeling it "Unconventionality" instead of "Openness".

Although this Big Seven UCV dimension is not a third Big Seven Valence dimension (does not join Big Seven NV and PV, which clearly are), the inclusive sampling of lexical terms did lead to the inclusion of certain distinctive Category III terms. This result led Widiger (1993a, 1993b) to suggest certain possible clinical correlates, which are discussed in more detail in the Discussion section of Study 4.

Dependability (DPD). This Big Seven dimension, the smallest, accounts for 6.9% of the Big Seven common variance. As worded in the S400 inventory and as numbered in Table 1, its 10 strongest markers (all positive) are 1. "Governed by reason or sound thinking, *reasonable*"; 2. "*Careful* and *thorough* in making decisions and determinations, *deliberate*"; 3. "*Firm* and *steadfast*, *true*, *staunch*"; 4. "Relying on orderly, systematic, and logical reasoning, *ratiocinative*"; 5. "Tested and proved to be ***good*** or ***trustworthy***, ***tried***"; 6. "*Consistent*"; 7. "*Obedient* or *dutiful*, *duteous*"; 8. "Pays attention, *observant*, *attentive*"; 9. "*Matter-of-fact*, *straightforward*"; and 10. "Characterized by ordered or systematic habits or behavior, *methodical*". **Summary**: 3 of 21 categorized terms are evaluative. DPD's strongest negative marker is "Careless in handling money, wasteful, thriftless", DPD: -.29, NEM: .23.

Of the positive Dependability markers, markers 2 and 10 include descriptors matching the positive Conscientious markers "Careful", "Thorough", and "Systematic" of Goldberg's (1992, Table 3) 100-item lexical Big Five inventory. Furthermore, Dependability marker 6 corresponds to his negative Big Five Conscientiousness marker, "Inconsistent". These significant overlaps notwithstanding, I could not adopt the Big Five label for this Big Seven factor, and chose "Dependability", a label closer to Goldberg's (1993, p. 27) "Conscientiousness or Dependability", and one I considered to be more accurately descriptive. However, for reasons similar to Loevinger's (1994) critique I too avoided "Conscientiousness" because of the connotative connection between "being conscientious" and "having a conscience". Adoption of the "Conscientiousness" label for the Big Seven Dependability dimension would have implied adoption of a disputable summary characterization of its major markers. The specific focus of the

negative marker on imprudent money management was also judged more consistent with the Dependability label than with the Conscientiousness construct.

Discussion

The arguments, procedures – among these the inclusive sampling of descriptors and the data-analytic SFP method -- and the resulting findings presented in this first study led to the identification of what I believe to be the major lexical US English-language dimensions of personal characteristics, the Big Seven. This positive outcome also encouraged continued reliance on the SFP procedure when called for in subsequent analyses.

As noted in the Data Analysis and Results section, the familiar scree test could have been interpreted as indicative of five rather than seven major dimensions. A five-factor solution was in fact the first one to be presented in some detail (Tellegen & Waller, 1987). In that solution, the four largest factors clearly tapped into the four largest Big Seven dimensions and were given the same or similar labels: Negative Valuation, Positive Emotionality, Positive Valuation, and Negative Emotionality. The fifth factor, named Normativeness-Conformance, comprised a mixture of positive and negative DSP and UCV markers (all keyed according to the negative poles of these two Big Seven dimensions).

The choice of a five-factor solution was only part of the story. In this early account mention was also made of an interpretable solution achieved with a larger number of dimensions and of additional analyses in progress. Furthermore, from the few additional analyses conducted at that time it already appeared that rotated factors beyond the first seven were comparatively small. These observations led soon thereafter to a still mostly content-based adoption of the Big Seven as the major S400 lexical dimensions. The present investigation finally yielded the more conclusive outcome achieved when the reformulated lexical hypothesis required replacement of the scree test (based on the series of unrotated principal component sizes plotted in Figure 1) with the drop test (based on the series of smallest rotated factor sizes plotted in Figure 2).

The Big Seven findings reported so far encouraged the four follow-up studies announced at the close of the earlier discussion of the lexical hypothesis and intended to address the following four questions. 1. How structurally robust are the Big Seven when additional factors are extracted and rotated? 2. Will separate analyses of the familiar and unfamiliar S400 items still give rise to Big Seven-like factor structures? 3. Could the inclusion of extremely desirable and undesirable descriptors in the S400 have resulted in artifactual factors? 4. Will the Big Seven prove to be replicable in a new data set? The following four studies address each question in turn.

Study 2: Examining the Structural Robustness of the Big Seven

What happens if more than seven, say, up to fifteen dimensions are extracted and rotated? To what extent will the Big Seven survive these over-extractions and continue to emerge as major dimensions? In other words, how structurally robust are the Big Seven?

Data Analysis and Results

As a first step toward addressing the robustness issue, I inspected the *seven largest* factors of each of the rotated 8-factor through 15-factor solutions -- in all, eight solutions. I found that in *each* of these eight 7-factor sets at least nine of the ten strongest markers of each factor marked only one Big Seven dimension, and that each Big Seven dimension was represented in this way. Furthermore, each of these factors shared at least seven of its markers with the ten strongest markers of its Big Seven counterpart (the ten listed in Table 1).

Following up on this overall appraisal, I subjected the robustness of the most extreme over-extraction to a still more demanding test. Just how well did the seven largest dimensions of the 15-factor rotation still represent the Big Seven? A direct comparison of the two sets of seven dimensions, one not contaminated by correlated measurement error, ruled out the use of overlapping markers. To avoid inclusion of too many weak markers as a result of not using any items twice, I selected just eight markers to represent each of the 14 dimensions (which is half the number of markers loading ≥ |.30| on the smallest Big Seven dimension, DPD). To represent the Big Seven, I first identified and selected the strongest eight markers of each Big Seven factor (included in Table 1) and only then, "without replacement" (that is, without re-selecting previously included Big Seven markers), selected from the remaining items the eight strongest markers of each of the largest seven dimensions of the 15-factor solution. Favoring in this manner (i.e., with respect to strength) the Big Seven markers over markers representing the first seven factors of the 15-factor solution meant that a clear robustness finding for the 15-factor solution would be a particularly convincing positive outcome.

The next step was to separately factor-analyze the 56 Big Seven markers and the 56 15-factor-based markers, in both analyses extracting and rotating seven factors and saving the (orthogonal) Anderson-Rubin (1956) factor scores. The two sets of 7 factor scales were then jointly analyzed. The 14 eigenvalues were obtained, and the SFP method was used to generate from 1 to 13 principal-factor solutions. The 14 eigenvalues are plotted in Figure 3, and the results of the SFP procedure, namely the size of the (unrotated) first principal factor (open circle) and of the smallest rotated factor from each of the subsequent 12 multifactor solutions (filled circles) are plotted in Figure 4. Even the slope of 14 eigenvalues suggested seven major dimensions. More pertinently, the drop test afforded by the plot of factor sizes shown in Figure 4 was judged to indicate a final major drop after seven factors.

Discussion

In view of these results, I believe it is reasonable to conclude that the seven largest factors of even the 15-factor rotation remained very similar to the Big Seven. It is in this sense that the Big Seven can be said to be structurally quite robust.

Study 3: Examining the Role of Familiarity

As noted in the section on the lexical hypothesis, terms judged unfamiliar were systematically excluded from the lexical English-language investigations from which the lexical Big Five emerged. It is therefore not surprising that the inclusion of unfamiliar terms in the present study has met with criticism. Saucier (1997, pp. 1298-1299), in particular, commenting, for example, on Tellegen (1993) and Waller & Zavala (1993), noted and questioned the inclusion, as a result of the page-sampling method, of obviously unfamiliar (albeit dictionary-defined) terms such a "rhadamanthine" and "tenebrous".

The purpose of this section is to deal with two major questions relating to the role of familiarity. First of all, how many familiar and non-familiar adjectives made up the S400 inventory? Second, how closely did the dimensions representing the seven-factor structures of the two subsets of terms approximate the Big Seven and even parallel each other? The first step toward answering the first question was to obtain familiarity ratings.

Development of Familiarity Booklets

The S400 items were divided into two booklets of 202 and 198 items, respectively. However, in both booklets only the originally sampled adjectives were reproduced. The often

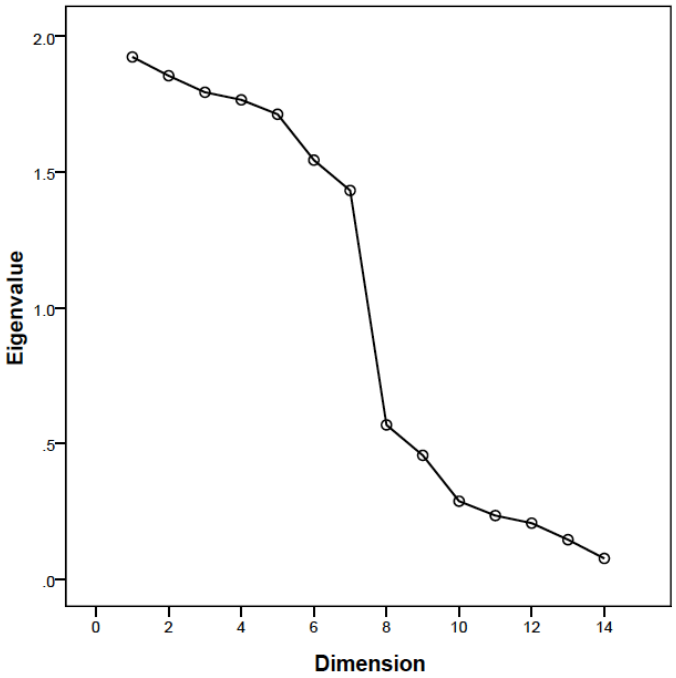

Figure 3. Study 2: Eigenvalue plot from 14 Anderson-Rubin factor scores produced by separate 7-factor principal-factor analyses of subsets of S400 psycholexical inventory marker items, namely, items marking the 7-factor (Big Seven) solution and non-overlapping items marking the 7 largest factors of the 15-factor solution.

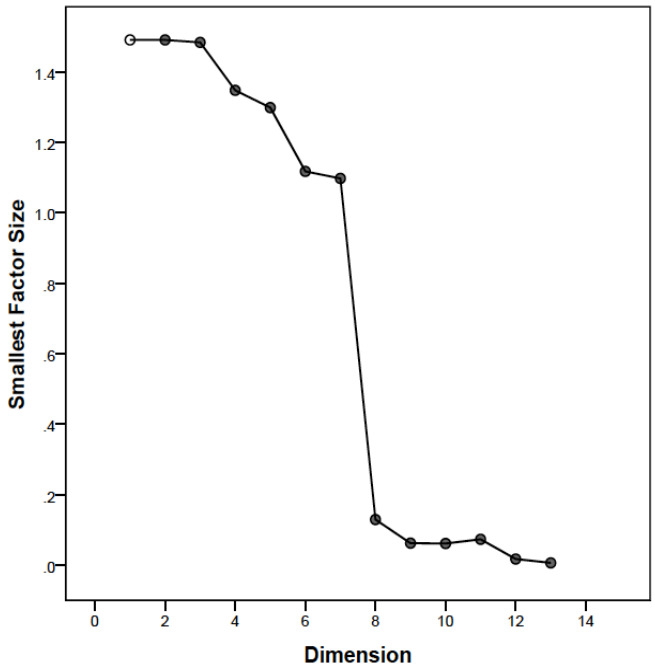

Figure 4. Study 2: Factor-size plot, produced by the SFP procedure applied to 14 Anderson-Rubin factor scores, of the first (unrotated) factor (open circle) and of the smallest rotated factor from each of 12 subsequent multi-factor solutions (filled circles). The 14 Anderson-Rubin factor scores were obtained from separate 7-factor principal-factor analyses of the same subsets of S400 psycholexical inventory marker items as in Figure 3, namely, items marking the 7-factor (Big Seven) solution and non-overlapping items marking the 7 largest factors of the 15-factor solution.

crucially clarifying dictionary definitions were left out. For example, in the second familiarity-rating booklet, "rhadamanthine" replaced its S400 counterpart, "strictly and uncompromisingly just; rhadamanthine".

Participants and Instructions

The undergraduate research participants, who were recruited from an introductory personality course and received course credits, were each handed one of the two booklets and a four-option multiple-choice answer sheet. Copies of the booklets were distributed to 27 and 28 participants, respectively. The instructions provided in the booklets explained that the current study was part of a larger investigation of how people describe other people or themselves, and that the booklet entries were adjectives sampled from a dictionary and included both familiar and unfamiliar terms. The instructions then directed participants to rate how sure they were of the meaning of each adjective when used as a personality descriptor and using the four response options provided as follows:

A. Am very sure what it means
B. Am moderately sure what it means
C. Have a vague idea what it means
D. Don't know what it means

Data Analysis and Results

The four item-rating choices were numerically coded 1, 2, 3, and 4, respectively, indicating increasing degrees of unfamiliarity. Three inter-rater reliability coefficients (Cronbach's alpha, Spearman-Brown, Guttman split-half) were computed for both booklets, and all six coefficients turned out to equal .98. One booklet item (of which, instead of the sampled adjective, the dictionary clarification had mistakenly been entered) had to be excluded from these calculations. The distribution of the 399 average item unfamiliarity values is plotted in Figure 5. Treating the midpoint scale value of 2.5 as the unfamiliarity level distinguishing reasonably familiar adjectives (numbering 149) from the non-familiar ones (numbering 250) made it clear, numerically and visually, that the non-familiar terms substantially outnumbered the familiar ones. Despite the small number of familiar terms in particular it still seemed reasonable to determine if empirical analyses of the two sets of items would yield similar seven-factor structures. To that end I conducted separate factor analyses of the familiar and unfamiliar S400 terms, performing in both analyses a seven-factor rotation.

I inspected these two solutions to determine whether both still resembled the Big Seven. I did so by examining in both these seven-factor structures the ten strongest marker items of each of the seven factors. I found that in *both* structures *each* Big Seven dimension was represented by one of the seven factors, such that at least six of its ten marker items were markers of that Big Seven dimension and no more than two marked a different Big Seven member. These findings suggested that both the familiar and unfamiliar items had yielded seven major dimensions approximating the Big Seven.

The availability of familiar and unfamiliar items in the same data set allowed, as a final step, a direct assessment of the actual similarity of the two Big Seven-like structures recovered from these two subsets of items. For this kind of analysis I was able to use the same procedure as the one conducted for the same purpose as part of the earlier reported structural robustness analysis.

First, familiar and non-familiar items were separately factor-analyzed. In both analyses seven factors were extracted and rotated and the Anderson-Rubin (1956) factor scores were saved. The resulting 14 factor scales were then jointly analyzed, the 14 eigenvalues were

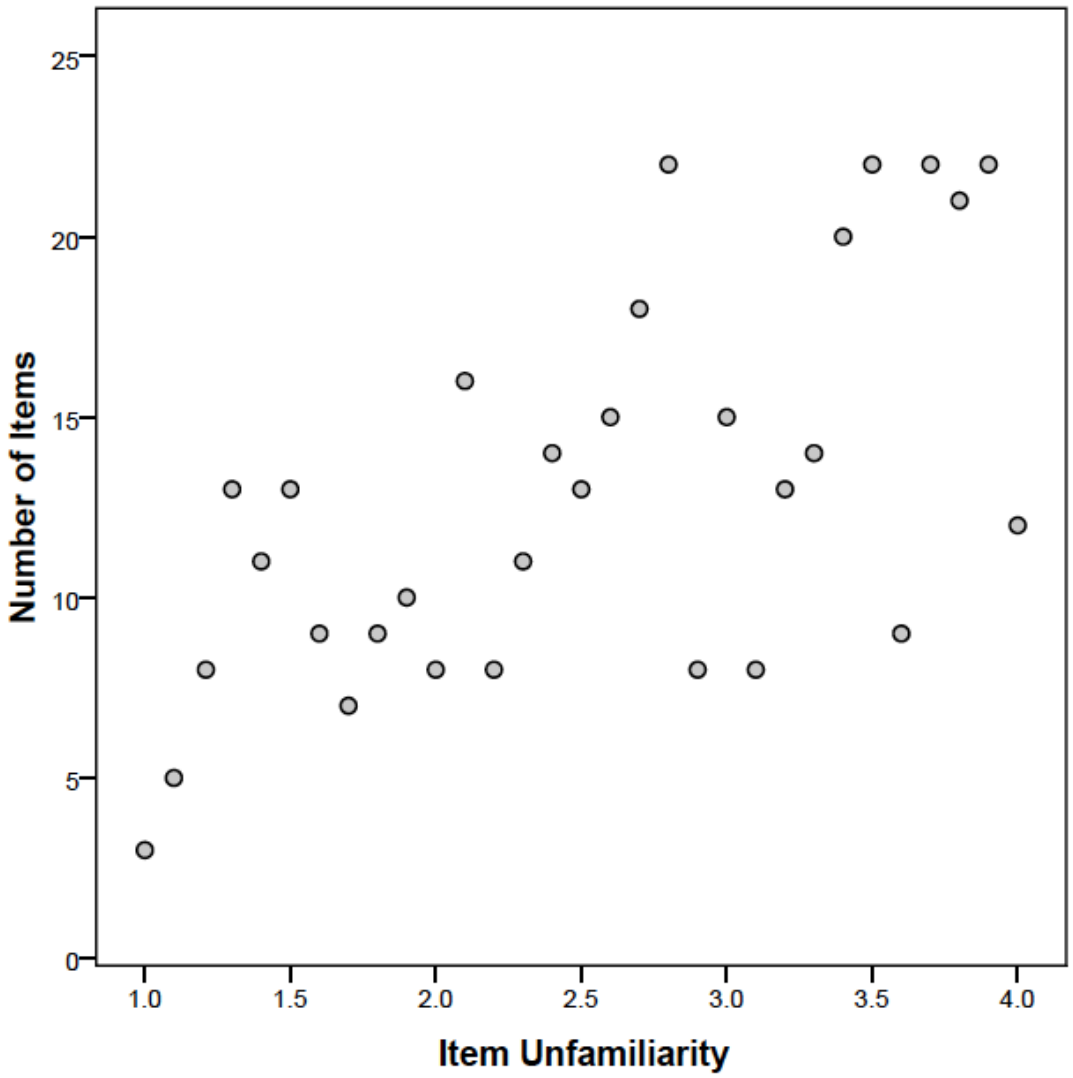

Figure 5. Study 3: Plot of the number of S400 psycholexical inventory items with mean unfamiliarity ratings ranging from 1.0 to 4.0.

obtained and the SFP procedure was applied to obtain from 1 to 13 principal-factor solutions. The 14 eigenvalues are plotted in Figure 6 and the results of the SFP procedure, namely, the size of the (unrotated) first principal factor (open circle) and of the smallest rotated factor from each of the 12 subsequent factor rotations (filled circles) are plotted in Figure 7. The drop test based on the latter was judged to indicate a final major drop after seven factors, indicating marked similarity of the seven-factor structures of the familiar and non-familiar items. As was true for the robustness findings, even the slope of 14 eigenvalues suggested seven major dimensions.

Discussion

The marked variations in familiarity among the S400 items understandably raised questions about how well the subsets of familiar and unfamiliar terms would still embody a Big Seven-like structure, and about the dimensional congruence of the familiar and unfamiliar terms. I believe the preceding analyses allay both these concerns. The large overall familiarity difference between the analyzed familiar and unfamiliar items did not result in the large marker departures from the Big Seven or the large factor-analytic difference one might very well have anticipated.

Study 4: The Significance of Evaluative Descriptors

In the earlier discussion of the lexical hypothesis I noted that in addition to insufficiently familiar descriptors, overly evaluative terms had also been excluded from the body of potential lexical Big Five markers. In the introduction, I also mentioned the merger of the Big Five and the Five Factor Model, or FFM, today's dominant personality trait taxonomy. However, FFM theorists McCrae and Costa (1995) at least temporarily suspended the very exclusion rules that were critical to the emergence of the Big Five and that by the same token also restricted the FFM content domain. This suspension made it possible to take a fresh look at what these authors, referring to the Big Seven NV and PV factors, called the "unabashedly evaluative constructs that have previously been intentionally omitted from descriptive systems" (p. 444). On the basis of a wide-ranging conceptual and empirical analysis, which included an experimental Big Seven inventory, the IPC7 (Tellegen, Grove, & Waller, 1991), McCrae and Costa (p. 458) concluded that while the large body of evaluative terms did not as such contribute to the description of personality, it merited continued efforts toward understanding its ubiquitous uses to characterize self and others.

Especially important for this section is McCrae and Costa's (1995, p.447) suggestion, in agreement with McCrae and John (1992), that NV and PV might be "infrequency" or "difficulty" factors resulting from skewed distributions of individual items. They also reported (p. 445) Widiger's (1993a) observation on this point that PV and NV appear to be extreme variants of traits located within the Five Factor Model. Quoting Gorsuch (1974, p. 260), they noted (p. 447) that such artifacts arise because "variables with similar marginal splits will correlate more with each other than they will with variables which do not have that same marginal split even when they all measure the same thing". That is, a problem may occur when the marker items of "the same thing" (the same non-artifactual dimension) consist of two subsets with difficulty levels that are similar within subsets and differ between them, so that item intercorrelations are stronger within the subsets. An artifactually separate difficulty factor may then emerge, one representing the subset of infrequently endorsed or difficult items, positively skewed given the S400 item scoring and the artifact source. While the label "difficulty factor" is used in reference to binary (e.g., True-False) items, a graded multiple-choice item format like that of the S400 items allows similar correlational contrasts with similar artifactual results. In sum, any two Big Seven factors

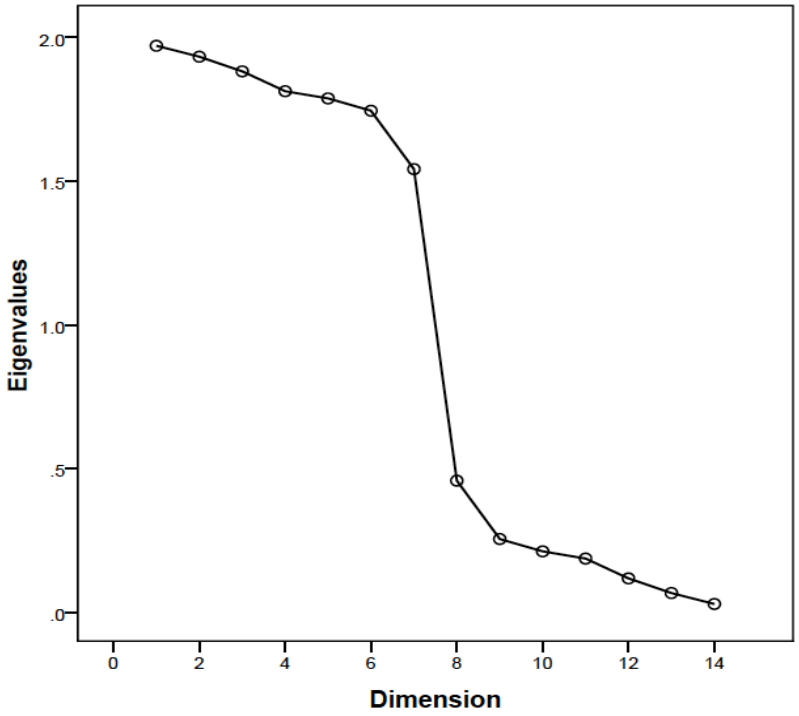

Figure 6. Study 3: Eigenvalue plot from 14 Anderson-Rubin factor scores produced by separate 7-factor principal-factor analyses of non-overlapping subsets of familiar and unfamiliar S400 psycholexical items.

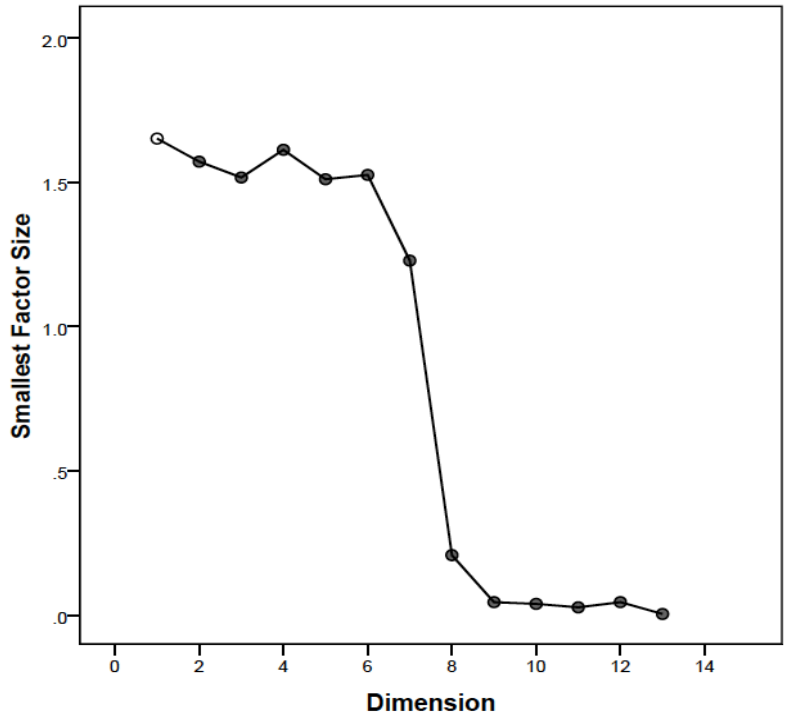

Figure 7. Study 3: Factor-size plot, produced by the SFP procedure applied to 14 Anderson-Rubin factor scores, of the first (unrotated) factor (open circle) and of the smallest rotated factor from each of 12 subsequent multi-factor solutions (filled circles). The 14 Anderson-Rubin factor scores were obtained from separate 7-factor principal-factor analyses of the same non-overlapping subsets of familiar and unfamiliar S400 psycholexical inventory items as in Figure 6.

could be descendants of a single lexical dimension, one of the two being artifactual, representing a distinctive subset of infrequently endorsed, hence positively skewed marker items.

McCrae and Costa (1995) did more than raise the possibility of NV and PV being skewness artifacts. They empirically examined the IPC7-based Big Seven items and found that the NV scores were indeed very skewed, but that the PV scores appeared more normally distributed (p. 449). They followed up by creating three IPC7 item parcels with "reasonable distributions" (p. 449) to represent each of the Big Seven dimensions. A principal component analysis of the 21 parcels yielded unrotated eigenvalues dropping sharply after the first seven, and the corresponding Varimax rotation clearly recovered all seven IPC7 factors (p. 449).

Although these analyses of the IPC7 indicated that NV and PV were not mere artifacts, this experimental inventory includes many items that were imported from non-lexical sources or were modified versions of lexical items. Skewness-related artifact issues involving the actual S400 lexical items still required independent empirical scrutiny.

A familiar approach to addressing skewness problems has been to replace product-moment correlations with tetrachoric or polychoric correlations (depending on whether binary or graded multiple-choice items are analyzed). These substitutions are based on the assumption that the underlying variables are normally distributed. Commenting on the tetrachoric correlation, Gorsuch (1983, p. 297) has characterized this assumption as often unwarranted. I therefore continued to use product-moment correlations when dealing with skewness issues involving NV and PV. I did so by creating item composites, just as McCrae and Costa (1995) had done, but in my own examination I proceeded to assemble S400 lexical item composites instead of IPC7 proxies.

Data Analysis and Results

I began by inspecting in the S400 data set the skewness statistics actually associated with Big Seven NV and PV. The distributions of the 15 strongest NV markers were all found to be unimodal and positively skewed, peaking at the *lowest possible* level of endorsement (response option 1: "Inaccurate, Uncharacteristic -- Very") with a median endorsement frequency of 86.6 % (range: 72.2% to 90.4%). The distributions of the 15 strongest PV markers turned out to be unimodal as well, and also peaked at the same endorsement level, but in contrast to the NV markers, at a *closer to intermediate* level (response option 3: "Accurate, Characteristic -- Slightly") with a median endorsement frequency of 53.0% (range: 36.1% to 61.5%). In sum, the PV marker ratings were not nearly as skewed as the NV counterparts, a result very similar to what McCrae and Costa's (1995) had found with the IPC7. This outcome suggested the (admittedly post-hoc) explanation that to some extent the combined PV attributes of desirability and improbability tended to influence respondents' degree of acceptance as self-characterizations in opposite directions, in other words, tended to cancel out each other.

Given these results, I proceeded to focus primarily on skewness-related questions about the Big Seven NV factor. Would closer structural scrutiny show it to be a mere skewness artifact? Even if not, could the extreme skewness of its marker items still have inflated NV's distinctiveness as a lexical dimension?

To answer these two questions, I assembled two composite scales conforming to the graded four-choice item format of the S400 items, but aggregated to represent the NV and PV factors each with an indicator that would be minimally skewed, namely, maximally rectangular. The NV composite was assembled from the 30 strongest individual NV markers after the first three (which were set aside for a different function explained below). Why were so many NV markers included? The goal of approximating a rectangular distribution, given not only the

pronounced skewness of NV's marker items but also the large number of these items (reflecting NV's dominant size) made it possible and desirable to aggregate an especially large number of NV markers. The PV composite was formed in the same way, except that only the strongest 15 markers following the first three were combined.

For the NV composite, it was possible to arrive at successive response-option proportions of .250, .247, .262, and .241, resulting in a skewness of -.004. For its PV counterpart, the corresponding results were successive response-option proportions of .243, .245, .280, and .232, and a skewness of -.028. In view of these roughly rectangular response-option distributions, the two composites are in what follows simply referred to as "rectangular composites".

The next step was to jointly factor-analyze the two rectangular composites and the earlier mentioned three strongest marker items of each Big Seven factor. The simple-structure seven-factor solution is reproduced in Table 2. From the factor pattern of the 21 strong Big Seven marker items it was evident that each of the Big Seven was represented by its own distinctive factor.

These results are informative about both NV and PV. They show that the rectangular NV composite with its NV loading of .50 is a clearly distinctive but substantially *weaker* marker of the same factor as the three leading highly skewed NV marker items. Therefore, in answer to the two earlier raised questions about NV being artifactual or artifactually inflated, these outcomes, although not supporting interpretations of NV as merely a skewness *artifact*, do demonstrate artifactual inflation and as such a skewness *effect*. For PV, the results are simpler, essentially as anticipated. Because leading PV marker items turned out not to be highly skewed to begin with, the rectangular PV composite was likewise expected to emerge as a strong marker. Beyond confirming this expectation, its loading of .79, strongest of the four in this solution, can be interpreted as in part reflecting its enhanced reliability as an aggregate indicator. Both these results are strong indications that PV is a genuine lexical dimension.

The 21 Big Seven markers were also factor-analyzed *without* the two rectangular composites to determine whether inclusion of the composites had appreciably altered the embedding basic Big Seven structure. The marker item factor pattern provided by this analysis is not shown here but clearly indicated that it had not. A comparison of the two sets of 3 x 7 x 7 = 147 marker item loadings showed that loadings on the same individual items differed from one another by at most |.03|, with just one exception: the .70 loading of the PV marker listed first in Table 2 was .65 in the composite-less factor pattern.

Discussion

The preceding analyses, culminating in the factor matrix reproduced in Table 2, were designed to deal with conjectures that NV and PV might be skewness artifacts. The findings presented here did not corroborate and in fact disconfirmed these ideas. Neither NV nor PV was found to be a mere skewness factor. However, as just noted, a comparison of the NV loading of the rectangular NV composite with those of the three NV markers showed the latter to be inflated, indicating a skewness effect.

However, even if both NV and PV had been shown to be skewness artifacts, which they are not, the Five Factor Model might nonetheless have benefited from the same kind of inclusive approach as the one adopted for the present investigation. This, at least, is how I have interpreted the position of a leading Five Factor Model theorist. Widiger (1993b, p. 138) declared that he considered the exclusion of evaluative terms an error in judgment. He suggested that the

Table 2

Factor Analysis of the Three Strongest Marker Items of Each of the Big Seven, Augmented with Near-Rectangular Four-Choice NV and PV Composite Items

Single/Composite Items	Factor Loadings						
	NV	PV	PEM	NEM	DSP	UCV	DPD
Rectangular NV Composite	***.50***	*.05*	*-.11*	*.26*	*.13*	*.13*	*-.27*
lousy	**.67**	-.02	-.08	.08	.04	-.03	-.14
reprehensible	**.78**	-.03	-.03	.02	.00	.04	-.05
depraved	**.73**	-.10	-.08	-.03	.00	.08	-.06
Rectangular PV Composite	*.01*	***.79***	*.22*	*-.11*	*.01*	*.12*	*.15*
excellent	-.14	**.70**	.07	-.11	.03	.11	.19
high-level	.00	**.77**	.12	-.03	.04	.00	.01
important	-.01	**.69**	.13	-.10	.06	.05	.08
spirited	-.05	.11	**.81**	-.06	-.08	-.04	.10
vigorous	-.09	.24	**.73**	-.15	.00	.01	.10
bouncing	-.13	.18	**.81**	.00	.04	-.06	-.07
easily troubled	.00	-.09	-.09	**.57**	.05	.00	-.02
guidable	.01	-.04	.05	**.47**	-.20	-.31	-.05
perplexed	.23	-.18	-.11	**.54**	.01	.02	-.14
stiff-necked	.11	.09	-.02	-.06	**.83**	.00	-.06
willful	.05	.03	.00	.06	**.76**	.06	.02
dogged	-.04	.00	-.02	-.03	**.71**	.21	-.02
divergent	.07	.02	.00	.01	.18	**.58**	-.07
heterodox	.10	.10	-.05	.06	.04	**.67**	-.03
orthodox	.02	-.07	.00	.15	.00	**-.55**	.06
reasonable	-.20	.06	.03	-.09	-.06	-.05	**.68**
deliberate	-.05	.13	.02	-.03	-.09	-.13	**.56**
staunch	-.10	.17	.06	-.05	.22	.02	**.38**

Note. N = 543; all strongest loadings are shown in boldface; loadings of the composite items are in italics. See text for additional details.

reformulation of the Openness dimension as Unconventionality as a result of the inclusion of certain evaluative terms indicative of the maladaptive variant of extreme openness (such as "odd" and "unusual") might improve the construct validity of the Five Factor Model and clinically also provide a better representation of schizotypal personality traits.

Study 5: Replication of the Seven Lexical Dimensions

For the purpose of examining the replicability of the Big Seven structure, a subset of previously identified Big Seven markers was selected and administered to a new participant sample.

Development of a Replication Inventory (S60)

Eight to ten previously identified marker items were selected from the S400 inventory to represent each of the Big Seven dimensions. The items were chosen to cover a *wider range of content*, especially of the five largest Big Seven dimensions, than do the 70 strongest markers reproduced in Table 1. To this end, the number of eligible Big Seven markers was expanded by also including markers with lower absolute Big Seven loadings. From this pool of eligibles a content-diverse subset was selected and assembled in a 60-item questionnaire, the S60. Of the 56 S60 items included in the data analysis (see below), 23 Big Seven markers items are as a result of these item selections not among those listed in Table 1.

Participants and Instructions

The S60 was completed by 696 volunteer introductory psychology student participants comprising 383 women, 295 men, and 18 individuals who had not identified their gender. Course credits and data collection schedule were the same as in Study 1. The instructions largely followed those provided with the S400. That is, before the S60 booklet was handed out, participants learned that the purpose of the study was part of an ongoing investigation intended to take a fresh look at how individuals describe self and others; that the research questionnaire consisted of personality adjectives taken directly from the dictionary; that although some items would seem bookish and would occasionally include unfamiliar terms, in each case careful reading of the item would clarify its meaning; and that the information provided would be kept strictly confidential.

Data Analysis and Results

In the S60 each Big Seven dimension was represented by at least eight markers, and when more markers had been included, only the eight previously identified strongest ones were used for a total of 56 markers. Factor analyses of the 56 markers were conducted to yield the first 30 eigenvalues, those plotted in Figure 8, and generate through the SFP procedure the series of 29 principal-factor sizes plotted in Figure 9. The factor sizes were plotted according to the same rules as before and, accordingly, the first circle open and the remaining 28 filled. The drop test based on this factor-size plot was judged to clearly indicate seven dominant dimensions. Considering the pool of preselected Big Seven markers, it is not too surprising that even the eigenvalues suggested seven major dimensions (unlike the eigenvalue series based on the entire S400 item pool plotted in Figure 1).

Following up on this result, an exploratory principal factor analysis of the 56-item intercorrelation matrix was specified to generate a seven-factor Varimax-rotated orthogonal simple structure, which yielded the factor structure summarized in Table 3. The pattern of

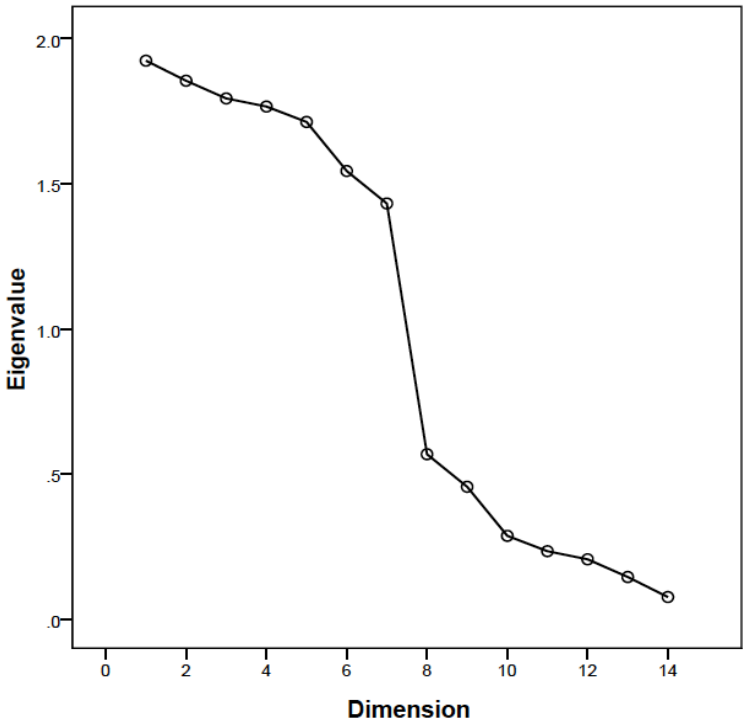

Figure 8. Study 5: Eigenvalue plot from the principal-component analysis of 56 S60 psycholexical inventory items.

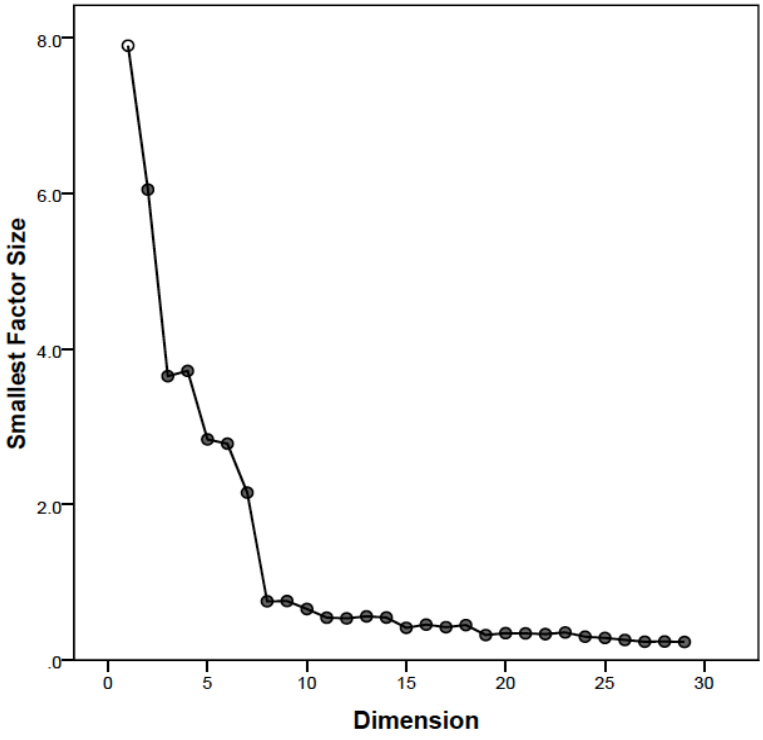

Figure 9. Study 5: Factor-size plot, produced by the SFP procedure, of the first (unrotated) factor (open circle) and subsequent smallest rotated factors (filled circles) resulting from principal-factor analyses of 56 S60 psycholexical inventory items.

Table 3

Replicated Factor Loadings of Previously Identified Marker Items of Each of the Big Seven Dimensions (7-Factor Orthogonal Rotation)

Items	Factor Loadings						
Negative Valence (NV)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*
lousy	**.64**	-.16	.03	.15	.16	.00	-.08
(heinous) reprehensible	**.69**	-.03	.05	.10	.21	-.02	-.09
(vicious) depraved	**.67**	-.08	.05	.07	.10	.04	-.13
thick	**.41**	-.07	-.03	.33	-.03	-.04	-.12
dotty	**.45**	-.13	-.07	.39	-.01	.16	-.18
dangerous	**.68**	-.11	.10	.02	.19	.04	-.04
discommendable	**.51**	-.13	.00	.31	.07	-.03	-.10
(compos mentis) sane	-**.33**	.12	.07	-.20	.08	-.15	.27
Positive Emotionality (PEM)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*
bouncing	.01	**.81**	.19	-.08	-.06	.13	.10
(on fire) enthusiastic	-.10	**.71**	.27	-.06	-.12	.05	.12
talkative	.01	**.64**	.14	-.02	.18	.08	-.04
boon	-.11	**.70**	.14	-.02	-.11	.03	.07
(rosy) optimistic	-.15	**.65**	.14	-.13	-.16	.03	.23
tacit	**.**15	**-.56**	-.05	.31	.03	.01	.10
(barren) not lively	.27	**-.47**	-.09	.34	.10	-.09	.09
aloof	.27	**-.42**	-.04	.27	.17	.02	.06
Positive Valence (PV)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*
(premier) important	.10	.15	**.56**	-.18	.06	-.12	.17
sensational	.05	.18	**.76**	-.15	.04	.06	.02
especial	.01	.00	**.66**	-.07	.14	.17	.13
excellent	-.17	.10	**.62**	-.13	-.03	.07	.26
foudroyant	.15	.23	**.65**	-.03	.03	.14	.02
elegant	.04	.16	**.41**	.01	-.12	-.11	.18
lofty	.09	.09	**.50**	-.12	.10	.00	.16
run-of-the-mill	.12	-.03	-**.43**	.36	-.02	-.24	.06
Negative Emotionality (NEM)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*
(queasy) easily troubled	.01	-.16	-.19	**.63**	.15	.05	-.02
bluffable	.10	.06	-.04	**.53**	-.10	-.07	-.04
(foggy) bewildered	.27	-.02	-.02	**.55**	.09	.07	-.07
Ill at ease	.22	-.26	-.09	**.61**	.18	-.03	-.03
(funky) panicky	.20	-.11	-.01	**.66**	.10	-.05	-.12
sheepish	-.04	-.15	-.12	**.39**	.08	-.12	.03
defenseless	.15	-.02	-.09	**.39**	**-**.13	-.08	-.03
apprehensive	-.01	-.06	-.18	**.58**	.17	-.04	-.03

Table 3. Replicated Factor Loadings, continued purposeful

Items	Factor Loadings						
Disagreement-Proneness (DSP)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*
(wayward) willful	-.03	-.01	.01	.02	**.60**	.08	.08
dubious	.16	-.17	.05	.04	**.40**	.14	.15
choleric	.35	-.14	.09	.29	**.45**	.00	-.02
cat-and-mouse	.31	.06	.11	.06	**.42**	-.05	-.03
counteractive	.32	-.10	.15	.06	**.47**	.09	.01
interruptive	.27	.06	.03	.25	**.38**	.11	-.06
concordant	-.04	.25	.07	.17	-**.37**	-.18	.16
leery	.26	-.19	-.03	.22	**.37**	.09	.08
Unconventionality (UCV)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*
divergent	.03	-.04	.02	.03	.17	**.55**	.05
orthodox	.11	.02	.06	.18	-.03	-**.64**	.34
buttoned-down	.18	-.24	.10	.14	.05	-**.48**	.22
progressive	-.02	.05	.06	-.03	-.06	**.46**	.08
weird	.16	.04	.16	.11	.20	**.48**	-.12
(sans-culottic) revolutionary	.25	.05	.17	-.04	.13	**.43**	.02
(semiformal) formal	.04	-.02	.05	.17	-.04	-**.59**	.29
(unco) uncanny	.20	.18	.22	.03	.18	**.42**	-.09
Dependability (DPD)	*NV*	*PEM*	*PV*	*NEM*	*DSP*	*UCV*	*DPD*
reasonable	-.15	.07	.01	-.09	.01	-.04	**.61**
deliberate	-.13	-.02	.10	-.12	-.05	.00	**.53**
staunch	-.07	.12	.26	-.23	.15	.04	**.37**
ratiocinative	-.03	-.09	.03	.08	.10	-.10	**.63**
(seamless) consistent	-.04	.04	.13	-.09	-.04	-.28	**.53**
attentive	-.17	.10	.16	-.05	-.08	.07	**.47**
methodical	.05	-.04	.12	.15	.07	-.19	**.49**
purposeful	-.08	.14	.28	-.14	.05	.05	**.42**

Note. N = 696. Shown are the groups of 8 items previously identified in and selected from the 400-item lexical research questionnaire as Big Seven markers. Item content is summarized in the same way as in Table 1. The strongest loading of each item is shown in bold-face. Additional Information is provided in the text.

highest absolute item loadings, printed in bold-face, shows each Big Seven dimension, as represented by its eight previously identified markers, to be most strongly associated with one and only one of the seven rotated factors, which was labeled accordingly. In other words, from the factor loadings matrix displayed in Table 3 it is clear that each Big Seven dimension, when represented by a diverse subset of previously identified markers, was successfully replicated in a new college sample.

Discussion

The analyses presented here were intended to answer the question raised earlier about the replicability of the Big Seven structure. The final results, displayed in Table 3, show this to be the case in a student sample recruited from the same population as were the participants in the original S400 study. This limits of course the generalizability of the current findings. Replicability appraisals of the Big Seven in other populations and settings will require additional empirical studies.

Summary and General Discussion

Although other studies of the US English-language lexicon of personal characteristics were also intended to test the lexical hypothesis, they excluded descriptors judged, for example, overly evaluative or unfamiliar, unlike the present investigation, which was designed to be inclusive. The page-sampling method adopted to implement this inclusive approach produced the S400 lexical inventory, the self-report instrument used or drawn on in each of the five studies reported here.

An inclusive sampling of descriptors is one of the two major distinctive methodological features of this investigation, namely, the one that defined its content domain. The second distinctive feature is the data-analytic Smallest Factor Profile (SFP) procedure. This method made it possible through its drop test, replacing the scree test, to determine in Study 1 the existence and number of major dimensions in the S400 data base.

The identification of the Big Seven in Study 1, and continued use of the SFP method when called for, made it possible to undertake and complete the reported four follow-up studies. The findings from these additional inquiries indicated, respectively, that the Big Seven were structurally robust, that two very similar sets of seven major dimensions approximating the Big Seven were recoverable from the familiar and unfamiliar S400 items, that the two Big Seven valence factors, NV and PV, were actual lexical dimensions, not artifacts, and that the Big Seven were replicable in a new sample.

Distinctions between the Big Seven

Interpretations of the findings reported here have thus far primarily focused on core characteristics of each individual Big Seven factor. Not explicitly addressed was the possibility of these core features also contributing to conceptually meaningful *distinctions between* the seven lexical dimensions. I believe this to be the case as well. First of all, the four valenced dimensions, PV, NV, PEM, and NEM, are clearly demarcated in terms of both the basic positive-vs.-negative and evaluative-vs.-affective distinctions. And the three remaining Big Seven dimensions, DPD, DSP, and UCV (unlike, it would seem, Big Five Conscientiousness, Agreeableness, and Openness) could be said to capture (in this study as self-rated attributes) the equally basic triad of, respectively, predictability, controllability (reversed), and explainability (reversed).

The Clinical Relevance of the Big Seven

In the discussion section of Study 4, I noted Widiger's (1993b) observation that because of the inclusion of certain evaluative terms, the Big Seven Unconventionality dimension might provide an improved representation of schizotypal personality traits. I would now add that the availability of Big Seven PV marker items could also allow the claims of excellence and exceptionality needed to capture the exaggerated self-appraisals and grandiosity considered to be associated with narcissistic personality and bipolar disorders.

I should note, however, that the inclusion of PV markers also allows psychologically informative appraisals of self and others as outstanding and exceptional that would *not* be narcissistically exaggerated and grandiose. An in-depth study of US presidents by Watts et al. (2013) may help illustrate this point. Its authors used extensive personality and behavior ratings of 42 US presidents provided by expert American scholars drawing on a wealth of information. Among other tasks, these experts had completed an observer-report version of the Five Factor Model (FFM) personality inventory, the NEO-PI-R (Costa & McCrae, 1992). Watts et al. (2013) drew on this information to derive sets of FFM-based ratings, including estimates of each president's "grandiose narcissism" (considered a sub-dimension of narcissistic personality disorder, and, as the title of their article indicates, its main focus). These ratings, quantified as *z*-scores and ranked from highest to lowest, ranged from 1.652 to -1.366 (Watts et al., 2013, p. 7). The five top-ranked US presidents were: L. B. Johnson ($z = 1.652$), T. Roosevelt ($z = 1.641$), Jackson ($z = 1.511$), F. D. Roosevelt ($z = 0.952$), and Kennedy ($z = 0.890$).

Equally interesting, however, were the low-ranking presidents, among whom Washington ($z = -0.212$), Jefferson ($z = -.481$), and Lincoln ($z = -0.495$) ranked 19th, 30th, and 32nd, respectively. To illustrate what I believe to be the need for also obtaining appraisals from a descriptive rather than a narrowly narcissism-oriented perspective on PV, I chose for two reasons the example of Abraham Lincoln: his very low narcissism rating and Doris Kearns Goodwin's (2005) revealing Lincoln biography. On the basis of the large body of biographical information about Lincoln and his rivals, Kearns Goodwin concluded that the self-taught Lincoln was a political and literary genius (p. 749). She also concluded (p. xvi) that he was not only shrewder and cannier than his rivals and fiercely ambitious, but also profoundly self-confident, in other words, also someone with a very high level of self-appraised PV. Furthermore, I would not expect Abraham Lincoln to be the only US president with both a low grandiose narcissism rating and a high PV self-rating.

For choosing leaders in today's world it would seem to be even more important to distinguish the high or low grandiose narcissism levels of available candidates from their PV self-appraisals. If in the Watts et al. (2013) study a set of valid observer-based ratings of the latter could also have been obtained, they would have been worth reporting as well, highlighting, for example, Lincoln's non-grandiose self-confidence. I would expect such self-appraisals to play an important role in many human endeavors (e.g., artistic, athletic, humanitarian, mathematical, political, religious, scientific, technological).

Attending to Study Limitation Issues

The Page-Sampling Method. An important methodological question is how best to select for one's lexical research inventory a representative subset of terms denoting personal characteristics. Saucier (1997), aware of Tellegen and Waller's (1987) description of the page-sampling method, argued that the roughly rectangular distribution of descriptors targeted by this method could not be expected to match the uneven spacing of these terms across the lexicon (pp. 1298-1299). As a case in point, he showed (footnote 2, pp.1298-1299) that many familiar descriptors begin with one of several prefixes (e.g., "dis-" and "un-") that are concentrated on relatively few dictionary pages and are therefore systematically under-sampled with a page-sampling method. The following comparison of the S400, a page-sampling product, with Goldberg's (1992) lexical Big Five inventory of 100 unipolar markers, therefore an instrument reflecting exclusions based on content but not on locations in a dictionary, served to provide a concrete test of Saucier's critique.

First, I identified in both the S400 and Goldberg's (1992) 100-item inventory all terms

beginning with one of Saucier's list of 11 prefixes. I found the total percentage of these terms to be indeed substantially lower in the S400 than in Goldberg's inventory: 5.25% vs. 35.00%, for a mean difference of 29.75%. However, I also found that the percentage differences varied widely across the 11 prefixes. By far the largest difference was associated with the prefix "un-", namely 18.75% (1.25% vs. 20.00%), the next largest being only 3.75% (.25% vs. 4.00%). Nonetheless, these results strongly suggest that the inclusive sampling method adopted in the present investigation remains a significant issue. Studies addressing it would have to begin by assembling a contemporary version of Allport and Odbert's (1936) exhaustive list of "trait names", from which an indisputably representative and reasonably sized sample of terms could then be selected and serve as a lexical research inventory. Such follow-up studies would seem worth the effort so long as the lexical hypothesis formulated here seems worth testing.

In that case, however, I believe that an optimal representation of prefixed terms would call for certain additional analyses. One major concern: non-affirmative responses to terms starting with negative prefixes may require "double-negative" responses to multiple-choice item formats (e.g., "I am not slightly non-envious") that could be challenging and problematic for many respondents. I believe such response options would need to be empirically examined before being included in a research inventory. Another concern: the number of included pairs of the same lexical terms with and without a negative prefix (e.g., "envious" and "non-envious"). A great many of such item pairs could overly narrow the content diversity of the inventory, and even increase the likelihood of "bloated specifics", a topic taken up shortly.

Choice of Participants. The five studies reported here were made possible by the contributions of college-student participants completing a self-report questionnaire under conditions of confidentiality. Broader demonstrations of the recoverability of the Big Seven will of course require a more inclusive sampling of participants and self-report conditions

Cross-Cultural Studies. Two cross-cultural tests of the Big Seven have also been reported, again with college-student participants. However, even setting aside this participant sampling limitation, the outcomes of both studies, while suggestive, are, for different reasons, not conclusive.

Benet and Waller (1995) reported a Spanish replication of the Big Seven structure, using a Spanish translation of the IPC7. As noted in the discussion of Study 4, this instrument, although its content is in part clearly related to that of the lexical Big Seven, is not a lexical inventory. Furthermore, the authors used for their 60-item Spanish-language questionnaire only the 10 best markers of each of the seven major factors from a previous analysis of the *English*-language IPC7, a selection that could have favored a corroboration of the English-language-based Big Seven.

Almagor, Tellegen, and Waller (1995) did conduct a lexical investigation, one intended to identify the major trait descriptor dimensions in the Hebrew language. They reported recovering six of the seven US English-language Big Seven factors. The seventh, UCV, was considered to have emerged clearly in the 15-factor solution, namely, reversed as the 13th rotated factor. However, the lexical inventory analyzed in this study had been markedly curtailed, in hindsight arguably too severely. Applying the page-sampling method to target every fourth page of their Hebrew dictionary, the authors first arrived at an initial list of 326 adjectives, but then proceeded to eliminate what they considered redundancies, to ensure that the final list would not contain words identical in meaning. The result of this pruning process was a 252-item lexical inventory. However, subsequent to the Hebrew study, in the present investigation, the version of the lexical hypothesis implemented by the SFP procedure was applied to the often content-

redundant S400 inventory and then yielded the clearly demarcated Big Seven structure. I now believe that this result, if obtained prior to the Hebrew study, could have discouraged the use of a minimally content-redundant lexical inventory. It could have been deemed especially unlikely that dimensions resembling DSP, UCV, and DPD, the three smallest Big Seven factors, even if embedded in the original Hebrew lexicon, would have emerged among the seven largest dimensions from such a curtailed inventory. To demonstrate the generalizability of the Big Seven or any of its individual dimensions across different cultures will continue to require full-scale empirical studies.

On the other hand, some observers might consider the removal of certain content-redundant terms a justified effort to avoid the intrusion of the kind of artifactual factors defined by near-synonymous markers Cattell (1973) has warned against and has labeled "bloated specifics". As DeYoung (2011, p. 718) and Oltmanns and Widiger (2016, p. 424) have pointed out, bloated specifics are direct results of an investigator's own efforts to construct a test. It seems therefore reasonable to ask how many major item dimensions a drop test would identify in any of today's Big Five and FFM inventories, and how many, if any, of these dimensions, would have to be judged bloated specifics. Empirical studies designed to answer these questions, while outside the scope of the present investigation, would seem worth the effort.

Could similar artifacts simply caused by sampling error (not by the possible factor-analytic artifacts examined in Study 4) also have vitiated the S400, causing the SFP-based drop test to cause the inclusion of one or more overly specific factors? It is possible but does not seem likely, especially not considering the highly congruent seven-factor solutions obtained from the markedly different subsets of lexical items examined in Studies 2, 3, and 5.

Relations to Other Constructs. As I have indicated previously (Tellegen, 1993, pp. 123-128), I believe it is a plausible conjecture to view the major psycholexical dimensions as "personal constructs" (Kelly, 1955), but ones that are *shared* rather than idiosyncratic, and as such are interpretable as "folk-concept" dimensions. Obtaining a person's self-appraisals on such dimensions could be an important part of any psychological assessment. The content of these different common self-appraisal concepts resembles in some ways that of various psychological-science (personological and psychopathological) concepts (demoralization, grandiosity, inferiority complex, internalization, schizotypy, self-esteem, etc.).

However, I also believe that the two kinds of concepts need to be considered *sui generis*. That is, folk concepts as well as science concepts, including psychological-science concepts, need to be acknowledged, but without blurring the distinction between the two. The need to maintain this distinction is especially clear in the case of the two Big Seven Valence dimensions. As shown earlier (in the interpretive section of Study 1), the adjectival content of both the Negative Valence and Positive Valence dimensions consists predominantly of terms listed in Allport and Odbert's (1936) Category III (Social Evaluations). As noted in the introductory discussion of the lexical hypothesis, these are the sort of terms psycholexical researchers have excluded from the lexical body of personal characteristics considered scientifically acceptable. On the other hand, science constructs could in principle serve the function of informatively *explaining* these folk concepts viewed as shared personal constructs.

The suggested distinction between folk concepts and science concepts is not unique to personology, nor are explanations of the former in terms of the latter. Rather similar conceptual distinctions and explanatory concerns can also be found in other areas of inquiry. Scientists accept the well-known "moon illusion" as a consensual perceptual experience, but have proposed "real-world" explanations. Meteorologists acknowledge the shared geocentric experiences of the

sun “rising” and “setting”, but again only as phenomena to be scientifically explained.

Another theoretically and empirically relevant issue is how congruent Big Seven self-ratings such as those obtained in the present investigation are with ratings of the same attributes as inferred by informed observers. The needed empirical studies have thus far not been undertaken, nor have any wider-ranging behavior-observational, biographical, and neuroscience studies.

Also, still to be clarified are relations between Big Seven constructs and similar but not interchangeable ones. For example, the Positive and Negative Emotionality and other constructs represented by the Multidimensional Personality Questionnaire (MPQ, Tellegen & Waller, 2008), and the clinical variants of these constructs represented by the MMPI-2-RF (Ben-Porath & Tellegen, 2008/2011; Tellegen & Ben-Porath, 2008/2011).

How about the dimensionality and content of future personality folk dimensions? The rapidly expanding network of near-instant global communications leads one to wonder about the possibility of increasingly world-wide similarities in the structure of folk dimensions.

However, the five studies comprising the current investigation focus solely on the US English-language lexicon of personal characteristics with the specific aim of empirically testing an inclusive interpretation of the lexical hypothesis. Although the results have been encouraging -- among these an arguably meaningful transformation of the lexical Openness construct into Unconventionality -- significant issues arising from study limitations obviously still need to be addressed -- among these the reliance on student participants and questionnaires being anonymously answered. In this connection, a special cautionary note may be called for to stress that the abbreviated Big Seven inventory (S60) analyzed in Study 5, while useful for its stated replication purpose, is by no means comparable to existing instruments representing alternative lexical models. Because of the enumerated study limitations, the development of such an instrument representing the Big Seven would of course have been premature.

References

Allen, T. A., & DeYoung, C. D. (2017). Personality neuroscience and the Five Factor Model. In T. A. Widiger (Ed.), *Oxford handbook of the Five Factor Model*. New York, NY: Oxford University Press.

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

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

American Heritage Dictionary (1982/1955). *Second college edition.* Boston: MA Houghton Mifflin.

Anderson, R. D., & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability, 5*, 111-150.

Ben-Porath, Y. S., & Tellegen, A. (2008/2011). *MMPI-2-RF (Minnesota Multiphasic PersonalityInventory-2 Restructured Form) manual for administration, scoring, and interpretation*. Minneapolis: University of Minnesota Press.

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.

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

Block, J. (2010). The five factor framing of personality and beyond: Some ruminations. *Psychological Inquiry*, *21*, 2-25.

Cattell, R. B. (1973). *Personality and mood by questionnaire.* San Francisco, CA: Jossey Bass.

Clark, L. A. (2007). Assessment and diagnosis of personality disorder: Perennial issues and an emerging reconceptualization. *Annual Review of Psychology*, *58,* 227–257.

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

De Raad, B., Barelds, D. P. H., Timmerman, M. E., De Roover, K., Mlacic, B., & Church, A. T. (2014). Towards a pan-cultural personality structure: Input from 11 psycho-lexical studies. *European Journal of Personality, 28,* 497-510.

DeYoung, C. G. (2011) Intelligence and personality. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 711-737). New York, NY: Cambridge University Press.

Epstein, S. (2010). The Big Five model: Grandiose ideas about surface traits as the foundations of a general theory of personality. *Psychological Inquiry, 21*, 34-39.

Galton, F. (1884). Measurement of character. *Fortnightly Review 36*, 179–185.

Goldberg, L. R. (1982). From ace to zombie: Some explorations in the language of personality. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in personality assessmen*t (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. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4,* 26-42.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26-34.

Goodwin, D. Kearns. (2005). *Team of rivals: The political genius of Abraham Lincoln.* New York, NY: Simon & Schuster.

Gorsuch, R. L. (1974). *Factor analysis.* Philadelphia, PA: Saunders.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Harkness, A. R., McNulty, J. L., & Ben-Porath, Y. S. (1995). The Personality Psychopathology Five (PSY-5): Constructs and MMPI-2 scales. *Psychological Assessment, 7,* 104-114.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York, NY: Guilford Press.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114-158). New York, NY: Guilford Press.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York, NY: Guilford Press.

Kelly, G. A. (1955). *The psychology of personal constructs* (Vols 1 & 2). New York, NY: Norton.

Krueger, R. F., & Markon, K. E. (2014). The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annual Review of Clinical Psychology, 10,* 477-501.

Loevinger, J. (1993). Measurement of personality: True or false. *Psychological Inquiry, 4,* 1–16.

Loevinger, J. (1994). Has psychology lost its conscience? *Journal of Personality Assessment, 62,* 2-8.

McCrae, R. R., & , Costa, P. T. (1995). Positive and negative valence within the Five-Factor Model. *Journal of Research in Personality, 29*, 443-460.

McCrae, R. R., & John, O. P. (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality, 60*, 175-215.

Norem, J. K., (2010). Resisting the hegemony of the Five-Factor Model: There is plenty of personality outside the FFA. *Psychological Inquiry, 21*, 65-68.

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.* Ann Arbor, MI: University of Michigan, Department of Psychology.

Oltmanns, J. R., & Widiger, T. A. (2016). Self-pathology, the Five-Factor Model, and bloated specific factors: A cautionary tale. *Journal of Abnormal Paychology. 125,* 423-434.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48) Hillsdale, NJ: Erlbaum.

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., & Goldberg, L. R. (1996). 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, NY: Guilford Press.

Saucier, G., & Goldberg, L. R. (1998). What is beyond the Big Five? *Journal of Personality, 66*, 495-524.

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

Saucier, G., Hampson, S. E, & Goldberg, L. R. (2000). Cross-language studies of lexical personality factors. In S. E. Hampson (Ed.), *Advances in personality psychology* (Vol. 1, pp. 1-36). Hove, England: Psychology Press.

Saucier, G., Thalmayer, A. G., Payne, D. L., Carlson, R., Sanogo, L., Ole-Kotikash, L.,… Zhou, X. (2014). A basic bivariate structure of personality attributes evident across nine languages. *Journal of Personality, 82*, 1-14.

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

Tellegen, A., & Ben-Porath,Y. S. (2008/2011). *MMPI-2-RF* (*Minnesota Multiphasic Personality Inventory-2 Restructured Form) technical manual.* Minneapolis, MN: University of Minnesota Press.

Tellegen, A., Grove, W. M., & Waller, N. G. (1991). *Inventory of Personal Characteristics #7 (IPC7)*. Unpublished materials. Minneapolis, MN: Department of Psychology, University of Minnesota.

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

Tellegen, A., & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In G. J. Boyle, G. Mathews, & D. H. Saklofske (Eds.), *Handbook of personality theory and assessment, Vol. 2: Personality measurement and testing* (pp. 261-292). London, UK: Sage.

Waller, N. G. (1999). Evaluating the structure of personality. In C. R. Cloninger (Ed.), *Personality and psychopathology* (pp. 155-200). Washington, D.C.: American Psychiatric Press.

Waller, N. G., & Ben-Porath, Y. S. (1987). Is it time for clinical psychologists to embrace the five-factor model of personality? *American Psychologist, 42,* 887-889.

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

Watts, A. L., Lilienfeld, S, O, Smith, S. F., Miller, J. D., Campbell, W. K., Waldman, I. D.… Faschingbauer, T. J. (2013). The double-edged sword of grandiose narcissism: Implications for successful and unsuccessful leaders among U.S. presidents. *Psychological Science, 24,* 2379-2389.

Webster's (1925). New international dictionary of the English language. Springfield, MA: G. & C. Merriam.

Widiger, T. A., (1993a). The DSM-III-R categorical personality disorder diagnoses: A critique and an alternative. *Psychological Inquiry, 4*, 75-90.

Widiger, T. A., (1993b). Reply to commentators: From B to Z. *Psychological Inquiry, 4*, 135-141.

Widiger, T. A., (2017). Introduction. In T. A. Widiger (Ed.), *The Oxford handbook of the Five Factor Model* (pp. 1-7). New York, NY: Oxford University Press.

Widiger, T. A., & Costa, P. T. (2012). Integrating normal and abnormal personality structure: The Five-Factor Model. *Journal of Personality, 80*, 1471-1506.

Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *American Psychologist, 62*, 71-83.