



## Identification and measurement of a more comprehensive set of person-descriptive trait markers from the English lexicon

Dustin Wood<sup>a,\*</sup>, Christopher D. Nye<sup>b</sup>, Gerard Saucier<sup>c</sup>

<sup>a</sup> Department of Psychology, Wake Forest University, USA

<sup>b</sup> Department of Psychology, University of Illinois, USA

<sup>c</sup> Department of Psychology, University of Oregon, USA

### ARTICLE INFO

#### Article history:

Available online 13 February 2010

#### Keywords:

Personality structure  
Personality measures  
Personality traits  
Big Five  
Trait structure  
Reliability  
Gender differences

### ABSTRACT

We suggest some refinements to earlier approaches to generating “comprehensive” personality inventories and address some methodological concerns that accompany their use. By applying cluster analysis to Saucier's (1997) list of the 504 most frequently used trait adjectives, we identified 61 clusters that can be used to represent the lower-order structure of individual differences found in the lexicon. We show that very short measures of these clusters have acceptable reliabilities, that single items can regularly be identified that correlate with standard measures of Big Five dimensions above .70, and finally, using gender and life satisfaction as examples, illustrate how comprehensive inventories can reveal relationships between personality and variables of interest that are masked by the use of Big Five scales alone.

© 2010 Elsevier Inc. All rights reserved.

### 1. Introduction

The most common approach to the development of individual difference measures in contemporary personality psychology can be labeled the *superfactor* approach, in which investigators search for a small number of broad factors underlying a set of personality-relevant descriptors (e.g., Ashton, Lee, & Goldberg, 2004; Goldberg, 1990; Tellegen & Waller, 1987). Although attempts to measure the broad dimensions identified in factor analysis represent the most common means of developing personality measures, a smaller tradition exists which can be labeled the *comprehensive* approach, in which investigators attempt to create measures to assess all important aspects of personality (e.g., Block, 1961; Peabody, 1987; Westen & Shedler, 2007). Unlike superfactor inventories, the primary goal of such inventories is not to measure the *major* ways that people vary, but to measure the *many* ways that people vary.

Investigators who use comprehensive inventories have done so in large part due to recognition of certain disadvantages of superfactor measures. Factor analysis proceeds by finding factors that can explain the largest amount of covariation between different traits ratings, and consequently a number of more distinctive dimensions that people regularly find useful in the description of themselves and others are frequently not represented in the factors

that are extracted. For instance, some of these characteristics that have been considered to be largely uncaptured by the Big Five factors include deceptiveness, honesty, sexuality, masculinity and femininity, frugality, religiosity, arrogance, humor, and physical dimensions such as height, weight, and attractiveness (Paunonen & Jackson, 2000; Saucier & Goldberg, 1998).

Another limitation of superfactor measures concerns how they are typically used. For instance, the domain of conscientiousness is described as covering a diverse array of content related to orderliness, conventionality, industriousness, and dependability. However, when these distinct elements are aggregated to form a single “broad” measure, or when only the “core” of the superfactor is measured, it is often unclear which distinguishable aspect of the measure is most related to the variable of interest. Although the Big Five and other superfactor solutions have been related to an enormous range of important outcomes (e.g., Ozer & Benet-Martinez, 2006; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), the frequent failure to examine how the narrower aspects of these superfactors differentially relate to these outcomes has slowed the accumulation of knowledge about the processes linking personality and behavior. In particular, we expect that distinguishable traits within a particular superfactor domain will generally vary in their associations with variables of interest – even frequently being associated in opposite directions – leading investigators to conclude that the relationship between the trait and the variable is large, small, positive, negative, or zero depending on the somewhat arbitrary content emphases of the scale (Paunonen, Rothstein, & Jackson, 1999; Saucier & Ostendorf, 1999). Indeed, differences in

\* Corresponding author. Address: Department of Psychology, 438 Greene Hall, Wake Forest University, Winston-Salem, NC 27109, USA.

E-mail address: [dwood@wfu.edu](mailto:dwood@wfu.edu) (D. Wood).

the content emphases of personality scales appear to have produced opposite conclusions regarding how traits are related to variables of interest in some investigations. For instance, the [Feingold \(1994\)](#) and [Lynn and Martin \(1997\)](#) meta-analyses of personality traits and gender both found gender differences on extraversion measures, but in opposite directions. As women have been found to score higher on the sociability aspects of extraversion than men but lower on assertiveness aspects, it seems likely that these discrepancies were produced by meta-analyzing scales with different content emphases in the domain of extraversion ([Costa, Terracciano, & McCrae, 2001](#)).

As we detail below, these problems can be circumvented by approaches that construct the lower-order structure of personality without reference to the Big Five or other superfactor structures. In the current studies, we delineate procedures that may be employed for the development and use of comprehensive personality assessments, and then present empirical demonstrations for why it may generally be beneficial to use such inventories in basic personality research. We begin by outlining procedures described by [Block \(1961\)](#) and [Peabody \(1987\)](#) in their attempts to develop a comprehensive taxonomy of individual differences, and then suggest some improvements.

## 2. Considerations in the development of more comprehensive personality assessments

Both [Block \(1961\)](#) and [Peabody \(1987\)](#) were interested in the development of comprehensive personality tools. Block's efforts to develop comprehensive assessment devices resulted in the development of the California Adult Q-Sort (CAQ; [Block, 1961](#)) and later the California Child Q-Set ([Block & Block, 1980](#)), both of which were designed to measure an individual's standing on as wide a range of attributes as possible. As described by [Block \(1961\)](#):

“The purpose of the [CAQ] is to provide a “Basic English” for clinical psychologists, psychiatrists, and personologists to use in their formulations of individual personalities. Ideally – and the set is not the ideal – the items should permit the portrayal of any kind of psychopathology and of any kind of normality. . . . To the extent the set fails in this aspiration, to the extent that it is deemed unable to reflect the discriminations and integrations of the observer, the method is to be judged deficient. (pp. 37–38).”

In an departure from the principal goal of factor analytic work on personality structure, which could be described as trying to locate a small number of the most important personality dimensions, Block evaluated the worth of his instrument by the extent to which it was able to capture all of the important discriminations people make about one another. Other investigators have cited comprehensiveness as a central goal in the development of similar instruments. For instance, [Peabody \(1987\)](#) argued that a thorough trait taxonomy should be able to describe “all perceptual variations in performance and appearance between persons or within individuals over time” (p. 59), and the [Shedler–Westen Assessment Procedure \(SWAP; Westen & Shedler, 1999\)](#) for assessing psychopathology was developed using procedures adapted from Block with the goal of allowing clinicians to be able to “describe everything considered psychologically important about their patients” ([Westen & Shedler, 2007](#), p. 812).

Although Block and Peabody appeared to outline their procedures for constructing comprehensive taxonomies independently, both arrived at a similar set of considerations that must be addressed to accomplish this task (see also [Stephenson, 1953](#)). First, the researcher needs to provide a defensible pool of items that could be used to define or represent the universe of content within

the domain. Second, if the complete set of items is large, then the researcher would need to use a method to represent the content found in the complete pool in a smaller number of items in order to make a practical assessment device.

Block's long-standing skepticism about the ability to define the universe of content related to personality and individual differences using an atheoretical method ([Block, 1961; Block, 1995](#); in press) led him instead to enlist the help of psychologist and psychiatrist “experts” to assist in identifying the important content of personality. His procedures in developing the CAQ have subsequently served as the primary guide for the development of other tools aspiring to comprehensiveness in other domains, such as the [Riverside Behavioral Q-Sort \(RBQ; Funder, Furr, & Colvin, 2000\)](#), and the [SWAP \(Westen & Shedler, 1999\)](#). In contrast, [Peabody \(1987\)](#) suggested that a comprehensive inventory could be developed through a relatively atheoretical rationale. In particular, [Peabody's \(1987\)](#) strategy for creating a comprehensive framework involved a two-step process of first classifying adjectives contained within a lexical pool into narrow groupings “according to their similarity in meaning,” which apparently involved Peabody's own discriminations, followed by the selection of terms within these larger clusters that could be used to represent the overall group.

## 3. The current approach to developing a comprehensive inventory

Inventories such as the CAQ, SWAP, and RBQ are used by a number of researchers to achieve comprehensive assessments of individual differences. However, there are some limitations to the construction of these instruments which may limit their use for this goal. For instance, it has been suggested that the reliance on psychiatrists to select content for the CAQ likely resulted in an over-representation of clinical terms ([Bem & Funder, 1978; Block, 1961](#)). More generally, we suspect that the generation of items primarily through the subjective decisions of subject matter experts will tend to result in instruments that over-represent certain aspects of personality and under-represent others. In keeping with a lexical tradition, we argue that it is preferable to adopt an approach to delineating the structure of traits which minimizes the number of subjective decisions researchers must make concerning the particular content that should be considered to exist within a pool. Although some of these decisions may ultimately be unavoidable, we propose a modified approach which aims to minimize these decisions. Following Peabody, we depart from [Block \(1961\)](#) in our consideration that certain methods of sampling terms from the lexicon (such as those used by [Saucier \(1997\)](#), [Tellegen and Waller \(1987\)](#)) serve as a reasonable means of generating item pools that represent the range of content associated with personality and individual differences.

An additional place where subjectivity should to be removed is through the secondary classification of the larger set of terms into a smaller set of homogenous categories or groups, which has almost invariably followed from intuitive considerations in the development of past comprehensive inventories ([Block, 1961; Peabody, 1987; Westen & Shedler, 2007](#)). We argue here that cluster analysis may be employed to accomplish this goal. Cluster analysis is designed to group similar items into a smaller set of clusters by optimizing the grouping of highly-correlated items with one another ([Cattell, 1944](#)). Due to this optimizing function, cluster analysis may be well suited for identifying the large number of distinguishable dimensions of individual differences that are indicated by several terms in language. This aim may be operationalized concretely as a decision to consider a dimension “well-represented” in the lexicon if there are a specified number of terms that show average or minimum inter-correlations above some specified magnitude. In essence, we can apply a standard criterion used for identifying a cluster

(e.g., Cattell, 1944) to the task of identifying the distinguishable trait aspects that can be thought to exist in the lexicon. If the goal of the investigation is to form a comprehensive summary of the lexicon, the investigator may aim to extract as many clusters from the lexicon as possible that meet certain homogeneity standards.

#### 4. Psychometric concerns with the use of comprehensive inventories

Given that comprehensive assessment tools are generally employed to relate a variable of interest to as many attributes as possible, developers of comprehensive inventories have often opted to measure each dimension by a single item to keep their assessments practical (e.g., Funder et al., 2000; Westen & Shedler, 1999). In fact, to make his instrument maximally comprehensive while making it practically administrable, Block (1961) chose to remove items that reflected content already contained in the inventory and replace them with more distinct items in order to have the resulting 100-item instrument reflect 100 distinguishable individual differences. In turn, this convention has led to a unique way in which the CAQ and similar instruments have often been used in personality research: investigations are frequently done by simply correlating all items of the inventory with a variable of interest, and then reporting the single items and their correlations together as the central results of the investigation (e.g., Bem & Funder, 1978; Block & Block, 2006; Colvin & Longueuil, 2001; Funder et al., 2000). Decisions such as these represent decisions that conflict with some commonly-held assumptions of how trait measures should be developed or used. In particular, it is commonly assumed that a measure of a dimension needs to be lengthy in order to be reliable and valid (e.g., John & Soto, 2007), as it is sometimes assumed that the reliability of single-item scales cannot be estimated (Gosling, Rentfrow, & Swann, 2003), or is precipitously low (Loo, 2001).

We believe the psychometric problems of single-item trait measures have been somewhat exaggerated. It should be remembered that reliability estimates are usually obtained in order to indicate the precision, dependability, or repeatability of scores (McDonald, 1999; Nunnally & Bernstein, 1994; Watson, 2004). Reliability coefficients are supposed to serve as an estimate of the expected stability in the rank-ordering of scores if we were to administer an equivalent test a second time (Cohen, Cohen, West, & Aiken, 2003). Suspicions of the low reliability of single-item scales likely derive in part from the modest reliabilities sometimes seen in broad scales. However, as the items of broad personality scales are generally selected to be fairly distinct from one another in order to cover different parts of the domain (e.g., the items *energetic*, *bold*, *talkative*, in an extraversion scale), item inter-correlations almost certainly underestimate the reproducibility of scores that would be observed if we simply assessed two close synonyms twice, or even the same trait item twice over a very short time span. As test–retest correlations over short time periods become better appreciated as appropriate reliability assessments (Cattell, Eber, & Tatsuoka, 1970; Watson, 2004), our understanding of the reliability of very short scales might stand to change dramatically. For instance, two-item Big Five measures have been found to have only slightly lower test–retest correlations than longer, standard scales (Gosling et al., 2003; Rammstedt & John, 2007). And single-item measures of self-esteem (Robins, Hendin, & Trzesniewski, 2001), life satisfaction (Pavot & Diener, 1993), and religiosity (Saucier & Skrzypinska, 2006) have all been reported as having test–retest correlations exceeding .70 over several weeks and longer. All of these values are far above what should be observed if single-item scales have unacceptable levels of reliability.

One or two-item scales also appear to be surprisingly valid. For instance, Gosling and colleagues (2003) found their two-item

scales to show correlations with matching Big Five Inventory (BFI) factors ranging from .65 to .87. Robins and colleagues (2001) found the item “I have high self-esteem” correlated with the complete Rosenberg self-esteem scale between .72 and .76, and Wood, Gosling, and Potter (2007) reported a .80 correlation between the single item “I feel depressed” with the average of the remaining 19 items of the Center for Epidemiological Studies Depression scale (CES-D; Radloff, 1977). A message emerging from these studies is that very simple, direct, and face-valid single items can have high levels of validity when related to longer standard inventories of the same construct.

#### 5. Goals of the present study

The overarching goals of the present study are to outline procedures that can be used to create more comprehensive inventories of personality characteristics and then to illustrate some uses of comprehensive inventories. The primary focus of Study 1 was to identify the distinguishable content that could be identified among English person-descriptor adjectives. To do this, we conducted a reanalysis of Saucier's (1997) common person-descriptor adjectives as assessed in a large community sample, which had been selected to consist of the most frequently used person-descriptor adjectives in the English lexicon. This analysis resulted in the identification of 61 content clusters. Study 2 then consists of the creation of single-item measures of the 61 clusters identified in Study 1, demonstrations of the reliability and validity of these items, and illustrations of what might be gained by using such measures to explore the relation between personality characteristics and other variables of interest.

#### 6. Study 1: Identification of clusters in the English lexicon

We begin with a reanalysis of a lexical dataset originally analyzed by Saucier (1997). As the goal was to generate a relatively comprehensive summary of the content contained within the item pool, we identified a cluster as any group of three or more items which was regularly extracted across a variety of clustering algorithms, and extracted as many clusters as possible. We then explored the adequacy of the clusters as a summary of the content found in the full item pool by comparing the factor structure of the clusters to the factor structure of the broader pool. Finally, a benefit of comprehensive measures is their ability to provide a quick snapshot of what the superfactors measure (Cattell, 1944). We thus correlated the clusters with NEO and AB5C measures of the Big Five in order to see how the clusters relate to the Big Five framework commonly used by personality psychologists.

##### 6.1. Method

###### 6.1.1. Participants and procedure

A total of 700 participants from the Eugene-Springfield Community Sample (ESCS) completed the materials examined in this study as part of an ongoing study on the structure of personality. Of the 639 participants who indicated their gender, 56% were female. Since we also report some correlations estimated from this sample later in Study 2, we also refer to this sample as “S1.” At the time they were recruited, participants ranged in age from 20 to 87 years, with a mean age of 53.8 (SD = 12.8) and median age of 52, and were of all levels of education. See Goldberg (2006a) for additional details about this sample.

###### 6.1.2. Saucier's common trait adjectives

To generate the clusters, we used data from participant self-ratings of the 504 adjectives identified as the most frequently used

person-descriptive adjectives in the English language by Saucier (1997). Participants rated the adjectives on a scale ranging from 1 (*very uncharacteristic or untypical of me*) to 7 (*very characteristic or typical of me*) with a midpoint of 4 (*uncertain, neutral, or meaning is unclear*). Additional analyses using these adjectives as assessed in this sample have been reported elsewhere (Saucier, 2003; Saucier, 2010; Wood et al., 2007).

### 6.1.3. Personality instruments

Participants also completed a number of different questionnaires designed to assess the Big Five. In the current study, we utilized some of these measures to validate the clusters that were identified. Specifically, we focused on the abridged Big Five Circumplex (AB5C) scales from the International Personality Item Pool (Goldberg, 1999) and the NEO-PI-R (Costa & McCrae, 1992). The AB5C measure was administered approximately a year prior to the administration of Saucier's common trait adjectives; the NEO was administered approximately 9 months prior to the administration of the adjectives.

## 6.2. Results

### 6.2.1. Identification of major clusters

To identify the major clusters of trait content in the lexicon, we used a hierarchical cluster analysis. With this method, items are grouped into a smaller number of clusters defined by aggregating items that are "similar" or "close" to one another using a particular distance measure (e.g., Pearson correlations or Euclidean distances). These relations can be summarized graphically through the use of *dendrograms* – diagrams which detail the distances between clusters of items. Cluster analyses typically generate unipolar clusters with antonyms appearing on separate clusters due to the fact that negatively correlated items are considered "distant" or "dissimilar" (e.g., *tall* will fall on a separate cluster than *short*, *tiny*). To allow the formation of clusters containing antonyms, we analyzed a set of adjectives that included all 504 items in their original form, and also created reverse-scored variables for all items, resulting in a cluster analysis of 1008 items. As the inclusion of all reversals in the cluster analysis resulted in dendrograms with two symmetrical halves (i.e., one set of clusters in the first half and the same set of clusters reverse-scored in the second half), only the clusters in one half of the dendrogram were examined. To facilitate the generation of bipolar clusters, we ipsatized the scores before creating reversals (i.e., subtracted the person's mean and divided by their standard deviation across all items). In general, this appeared to reduce the positive correlations between synonymous terms (e.g., *outgoing* and *sociable*), and increase the negative correlations between antonymous terms (e.g., *outgoing* and *shy*), which is consistent with the use of ipsatizing as a means of removing general acquiescence biases (Hofstee, ten Berge, & Hendriks, 1998; Soto, John, Gosling, & Potter, 2008). All analyses conducted to identify clusters were done using these ipsatized scores.

Different clustering algorithms can be used to create clusters, and optimize different criteria. The differences between these algorithms can cause terms to be placed in different clusters in the dendrogram. Consequently, we used three clustering algorithms to extract initial cluster groups, which were then compared across solutions. The first algorithm used was *between-groups linkage* (or *average-linkage*) clustering, which estimates the distance between two clusters as the average distance between all inter-cluster pairs. The second algorithm used was the *furthest-neighbor* (or *complete-linkage*) clustering, which estimates the distance between two clusters as the distance between the two furthest (least related) items of each cluster. The third algorithm used was *within-group linkage* clustering, which attempts to minimize the distance between items within the same cluster. The complete-linkage and

between-group linkage algorithms tend to create clear clusters more readily than other common clustering algorithms (Aldenderfer & Blashfield, 1984), whereas within-group linkage generally results in fewer clusters but also tends to maximize the similarity of items contained within the clusters.

For initial cluster extraction, a group of items was determined to form a cluster if three items were found to cluster together before a particular distance in the dendrogram. The distances we selected were chosen to correspond roughly to the point at which the inter-item correlations exceeded .30. These decision rules were used for a couple reasons. First, as one goal of the analysis was ultimately to create an inventory where all clusters could be assessed in a relatively practical manner, these decision rules were chosen in order to ultimately identify a moderate number of clusters (e.g., between 30 and 80 clusters) from the larger lexical pool. Second, requiring three items to be associated beyond the chosen threshold would make us more confident that the trait dimension underlying the cluster was "well-represented" by terms in the lexicon and would likely be identified in other lexical datasets than if only two items were required or a smaller intercorrelation was necessary, which could lead to less internally consistent clusters. Using these decision rules, a total of 55 clusters were extracted from the between-linkage clustering, 81 from the complete-linkage clustering, and 62 from the within-linkage clustering algorithms.<sup>1</sup>

Given that different clustering algorithms can place similar items in fairly different places in their dendrograms, the next task was to identify clusters that had been extracted across multiple solutions. The cluster membership of each item was compared across the three methods and a cluster was considered to be identified across multiple solutions if it met any of the following criteria:

- (a) Any set of three or more items that were defined as part of the same cluster in all three analyses.
- (b) Any set of four items that were defined as part of the same cluster in one solution, with three out of four of these items defined as part of the same cluster in both of the other two solutions.
- (c) Any set of four or more items that were defined in the same cluster in any two solutions.
- (d) Any set of three or more items in which two items were placed on the same cluster in all three solutions and additional items existed which had sufficiently high correlations with the other items (all inter-item correlations  $>.30$ , and average inter-item correlation  $>.40$ ).
- (e) Finally, to ensure that the items within the cluster showed an adequate level of homogeneity, a group of items was only considered to form a cluster if it consisted of at least three items with all inter-item correlations  $>.30$  and with an average inter-item correlation  $>.40$ , or if it consisted of at least four items with nearly all inter-item correlations  $>.30$ , and an average inter-item correlation  $>.35$ .

It should be noted that these rules allowed items forming a large cluster in one solution to be split into two or more smaller clusters if they formed smaller clusters with the other cluster algorithms. For inclusiveness, when a group of items could be considered to be defined by more than one rule, the rule which resulted in a cluster with more items was used. After comparing the clusters across solutions using these rules, a total of 62 clusters

<sup>1</sup> The complete dendrograms and the file used to compare clusters across solutions are available from the first author.

were identified across analyses, which included 302 of the complete set of 504 terms.

We took additional steps in order to minimize the presence of redundant clusters, and to identify finer distinctions that might have existed within large cluster groups. We focused on items: (a) that were placed on a single cluster in two dendrograms but were split into two clusters in a different dendrogram (which would result in two clusters by the rules above), or (b) that fell on a cluster containing six or more items. The items defining such clusters were placed into a factor analysis using principle axis factoring and direct oblimin rotations. If two distinct dimensions could be identified within the analysis which each had at least three items with loadings of .40 or above, two clusters would be considered from the larger group, and if not, only one cluster would be considered. Ultimately, this resulted in splitting one large cluster into two smaller clusters, and combining four clusters into two larger clusters. A total of 61 clusters were ultimately identified following these procedures; the person-descriptor adjectives which defined each cluster are provided in Appendix A.

### 6.2.2. Factor structure of the 61 clusters

We next explored the structure of the 61 clusters by constructing a “factor tree” (Goldberg, 2006b; Wood et al., 2007) to explore the major content areas represented as a progressively larger number of factors were extracted. A central goal of this investigation was to examine the extent to which the cluster extraction method resulted in a set of clusters which represented major dimensions of content (e.g., extraversion) in proportion to its level of representation in the larger pool. Because the dataset used to extract the clusters was the same as that used in Saucier’s (1997) investigation, we were particularly interested in the nature of the seven-factor solution: if the seven-factor structure identified here is similar to the seven-factor structure reported by Saucier (1997), it would suggest that our 61 clusters represented content roughly in proportion to its density in the larger lexical pool of 504 terms.

Since the clusters differed in the number of items that were contained in each, which would serve to create differential reliability across clusters if all items were averaged as a function of differential scale length, the length of the cluster scales was standardized by averaging two items in the cluster. Our goal in selecting markers was to use adjectives that measured the content closest to the core of the cluster, and to select two terms that were highly synonymous, so that the markers communicated similar information about the core content of the cluster. In most cases, the two terms we selected to represent the cluster were the two items with the highest factor-loadings when the cluster items were placed into a factor analysis and one factor was extracted (the items are presented in this order in Appendix A). Occasionally markers besides the two with the highest factor-loadings were used; this was done: (1) primarily in order to use cluster markers that were more synonymous (e.g., the markers *smart*, *intelligent* was used instead of *smart*, *intellectual*), which was determined in part by looking at inter-item correlations, (2) to avoid the use of negation terms and “un-” words when possible (e.g., the marker *awkward*, *clumsy* was used instead of *awkward*, *ungraceful*), and (3) to utilize more behavioral terms within the cluster than evaluative or reputational terms when possible (e.g., *influential* was used instead of *well-known*).

As the two-item cluster markers were selected to use highly synonymous terms, we expected the inter-item reliabilities of these two-item marker scales to be fairly high. We estimated Cronbach’s alpha for each of the 61 two-item cluster markers (see Table 2); these coefficients can be interpreted as the expected correlation between the two-item marker scales with two different but equally synonymous terms. Across the 61 cluster markers, alphas ranged from .56 to .94 (average  $\alpha = .76$ ).

The two markers selected to represent the clusters were averaged and then entered into a factor analysis. Varimax rotations of factors extracted using principal axis factoring were used to extract factors from the ipsatized data; the first 12 eigenvalues were 10.0, 5.5, 3.7, 2.4, 2.3, 2.2, 1.6, 1.5, 1.4, 1.3, 1.2, and 1.1. Although these eigenvalues provide some indication of a six-factor structure, we chose to construct a factor tree detailing the nature of the factors identified as successively more factors were extracted, as this can frequently provide more information about the factor structure of a dataset (see Goldberg, 2006b). This is shown in Fig. 1; only the first seven factors are shown, as the eight-factor solution produced a factor with only a single cluster marker loading above .40 (*unreliable/undependable*). To aid in interpreting the factors, one adjective from at least the first four highest-loading cluster markers for each factor is given.

A couple of findings are noted here. First, the two-factor solution consisted of a factor defined by terms relating to agency, dynamism, and positive emotionality (*exciting*, *sad<sub>R</sub>*, *admirable*, *confident*) and a factor defined by terms relating to communion and social self-regulation (*pleasant*, *kind-hearted*, *courteous*, *stable*). Within the five-factor solution, four of the factors showed close resemblance to the traditional Big Five factors: agreeableness, emotional stability, extraversion, and conscientiousness-related factors could be identified, however, a strong positive evaluation/attractiveness dimension was found in the place of an intellect or openness factor. By the six and seven-factor solutions, however, all of the traditional Big Five factors were clearly recognizable. Interestingly, this analysis neatly separated intellect and openness interpretations of the fifth factor found in five-factor structures (e.g., DeYoung, Quilty, & Peterson, 2007) into separate factors in the seven-factor solution, with an openness factor splitting off from a broader conscientiousness factor and an intellect factor splitting off from a broader positive evaluation factor.

Although the eigenvalues associated with the 61 clusters did not suggest a seven-factor structure as clearly as Saucier’s (1997) original investigation using unaggregated trait adjectives from this dataset, our results indicated that the seven-factor solution nonetheless appeared very similar to the seven-factor solution reported by Saucier, with both showing Big Five-like factors and a physical attractiveness factor. The only significant divergence was the identification of a “negative valence” or “low-base-rate” factor in Saucier’s (1997) seven-factor solution, which was replaced here with a traditionalism or low openness factor. The similarity of the two seven-factor solutions indicates that the 61 clusters represented major dimensions of content in close proportion to its level of representation within the broader set of 504 items.

### 6.2.3. Relationships between the 61 clusters and Big Five measures

We next explored how the cluster markers were associated with the NEO and AB5C scales in order to see how the clusters related to common Big Five measures; the results are reported in Table 1. To provide more structure to the presentation of the results, we chose to order the clusters in descending order of the Big Five trait they showed the highest association with, and to group clusters dealing with the highly evaluative/reputational and physical content at the bottom of the table. We limit our discussion of how AB5C and NEO scales were associated with the cluster markers to associations that exceeded  $|r| = .30$ .

**6.2.3.1. Extraversion.** The NEO and AB5C extraversion scales were most associated with clusters indicating sociability (*outgoing/sociable*, *bashful/shy*), indicating this was the content closest to the core of standard extraversion scales. The scales also tended to correlate with clusters indicating positive emotionality and affectivity (*happy/joyful*, *positive/optimistic* vs. *sad/unhappy*), various indicators of

**Table 1**  
Correlations between cluster markers and AB5C, NEO, and BFI scales.

Cluster markers	Extraversion			Agreeableness			Conscientiousness			Emotional stability			Openness/intellect		
	AB5C(S1)	NEO (S1)	BFI (S3)	AB5C(S1)	NEO (S1)	BFI (S3)	AB5C(S1)	NEO (S1)	BFI (S3)	AB5C(S1)	NEO (S1)	BFI (S3)	AB5C(S1)	NEO (S1)	BFI (S3)
Outgoing, sociable	<b>.66</b>	<b>.62</b>	<b>.78</b>	<b>.35</b>	.18	.21	.12	.13	.12	.11	.23	.23	.14	.16	.12
Enthusiastic, excited	<b>.59</b>	<b>.59</b>	<b>.59</b>	.29	.10	<b>.34</b>	.07	.11	.23	.09	.17	.26	.28	<b>.30</b>	.16
Bold, assertive	<b>.60</b>	<b>.46</b>	<b>.57</b>	-.07	-.23	-.13	.11	.15	.13	.12	.21	.20	.28	.14	.12
Happy, joyful	<b>.41</b>	<b>.43</b>	<b>.46</b>	<b>.40</b>	<b>.31</b>	<b>.43</b>	.10	.18	.23	<b>.35</b>	<b>.41</b>	<b>.34</b>	.03	.08	.11
Loud, noisy	<b>.44</b>	<b>.32</b>	<b>.50</b>	-.12	-.18	-.19	-.16	-.19	-.15	-.22	-.15	.05	.05	.05	.03
Funny, amusing	<b>.45</b>	<b>.38</b>	<b>.39</b>	.09	-.07	.11	-.10	-.03	.10	-.01	.07	.22	.24	.19	.18
Brave, adventurous	<b>.35</b>	<b>.36</b>	<b>.38</b>	.01	-.05	.06	-.04	.04	.10	.13	.18	.22	<b>.31</b>	.24	<b>.31</b>
Bashful, shy	-.54	-.41	-.74	-.11	-.01	-.09	-.14	-.21	-.09	-.19	-.32	-.33	-.10	-.04	-.09
Kind-hearted, caring	.21	.28	.21	<b>.67</b>	<b>.47</b>	<b>.61</b>	.17	.12	.24	.12	.10	.09	.16	.20	.20
Giving, generous	.24	.30	.18	<b>.58</b>	<b>.42</b>	<b>.57</b>	.11	.09	.22	.09	.10	.10	.14	.18	.16
Pleasant, agreeable	.14	.21	.25	<b>.49</b>	<b>.39</b>	<b>.49</b>	.14	.17	.29	.20	.24	.24	.08	.10	.12
Thankful, grateful	.17	.23	.23	<b>.46</b>	<b>.36</b>	<b>.50</b>	.16	.12	<b>.32</b>	.15	.14	.14	-.09	.00	.14
Affectionate, loving	<b>.41</b>	<b>.38</b>	.29	<b>.50</b>	<b>.30</b>	<b>.49</b>	.08	.04	.20	-.01	.05	.07	.22	.27	.16
Courteous, polite	.05	.14	.09	<b>.49</b>	<b>.34</b>	<b>.46</b>	<b>.31</b>	.20	<b>.36</b>	.13	.13	.11	-.02	.02	.15
Truthful, honest	.00	.06	.17	<b>.38</b>	.27	<b>.36</b>	.25	.23	<b>.36</b>	.15	.17	.08	-.02	-.05	.14
Rude, inconsiderate	-.03	-.10	-.07	-.48	-.41	-.65	-.30	-.24	-.28	-.27	-.24	-.12	.02	-.01	-.12
Selfish, self-centered	.01	-.05	-.01	-.47	-.46	-.57	-.23	-.20	-.21	-.29	-.23	-.11	.10	.07	-.08
Unfriendly, cold	-.34	-.33	-.26	-.52	-.40	-.57	-.19	-.17	-.26	-.19	-.21	-.16	-.09	-.11	-.12
Angry, hostile	-.09	-.13	-.05	-.40	-.41	-.56	-.23	-.23	-.17	-.50	-.48	-.28	.09	.05	-.05
Conceited, egotistical	.27	.15	.02	-.32	-.44	-.56	-.12	-.08	-.17	-.16	.01	-.10	.21	.13	-.07
Cruel, abusive	.04	.01	-.09	-.37	-.35	-.48	-.13	-.15	-.24	-.31	-.26	-.20	-.02	.01	-.11
Controlling, dominant	<b>.38</b>	.25	.26	-.24	-.40	-.41	.10	.08	-.01	-.22	-.10	-.06	.20	.04	-.04
Dependable, reliable	.06	.14	.10	.29	.16	<b>.31</b>	<b>.36</b>	<b>.34</b>	<b>.56</b>	.09	.15	.04	.03	-.07	.07
Practical, sensible	.01	.05	.01	.23	.13	.20	<b>.41</b>	<b>.39</b>	<b>.40</b>	.17	.19	.14	.00	-.14	.01
Competent, capable	.24	.26	.23	.24	.04	.27	<b>.38</b>	<b>.34</b>	<b>.42</b>	.22	<b>.31</b>	.14	.28	.11	.24
Disorganized, messy	-.02	-.07	.01	-.15	-.11	-.09	-.63	-.55	-.61	-.15	-.18	-.14	.09	.18	.07
Undependable, unreliable	-.09	-.10	-.08	-.26	-.15	-.38	-.39	-.35	-.54	-.12	-.21	-.13	-.02	.05	-.10
Awkward, clumsy	-.13	-.17	-.09	-.10	-.06	-.13	-.35	-.31	-.29	-.34	-.37	-.21	.01	.06	.07
Calm, relaxed	-.02	.00	.09	.19	.21	.19	.04	.10	.05	<b>.51</b>	<b>.45</b>	<b>.60</b>	.00	-.05	.05
Confident, self-assured	<b>.45</b>	<b>.37</b>	<b>.48</b>	.18	.07	.19	.27	<b>.33</b>	.29	<b>.45</b>	<b>.55</b>	<b>.40</b>	.16	.02	.14
Stable, well-adjusted	.25	.24	.26	<b>.32</b>	.21	.29	<b>.30</b>	<b>.36</b>	<b>.33</b>	<b>.45</b>	<b>.52</b>	<b>.42</b>	.08	-.06	.08
Positive, optimistic	.37	.40	.45	<b>.33</b>	.21	.41	<b>.15</b>	<b>.23</b>	<b>.24</b>	<b>.39</b>	<b>.45</b>	<b>.41</b>	.14	.14	.15
Tense, anxious	-.10	-.11	-.22	-.10	-.15	-.22	-.09	-.19	-.04	-.58	-.56	-.81	-.07	-.02	-.08
Afraid, scared	-.23	-.19	-.27	-.11	-.10	-.19	-.21	-.29	-.18	-.54	-.61	-.55	-.09	.02	-.09
Sad, unhappy	-.33	-.32	-.39	-.30	-.23	-.36	-.17	-.27	-.28	-.52	-.57	-.57	-.02	-.02	.00
Touchy, temperamental	-.06	-.09	-.04	-.22	-.25	-.39	-.07	-.19	-.18	-.53	-.51	-.44	.01	.02	-.08
Crabby, grouchy	-.12	-.16	-.22	-.37	-.37	-.52	-.17	-.24	-.24	-.46	-.47	-.44	-.02	-.02	-.09
Lonely, lonesome	-.28	-.27	-.40	-.19	-.15	-.29	-.17	-.23	-.22	-.38	-.45	-.46	.09	.02	.02
Creative, imaginative	<b>.31</b>	.19	.20	.10	-.06	.15	.03	.06	.04	.08	.16	.14	<b>.50</b>	<b>.39</b>	<b>.72</b>
Intelligent, smart	.17	.12	.20	.07	-.08	.16	.12	.13	<b>.31</b>	.19	.26	.14	<b>.51</b>	<b>.30</b>	.23
Radical, rebellious	.12	.06	.18	-.27	-.33	-.33	-.33	-.27	-.29	-.25	-.22	-.01	<b>.36</b>	<b>.32</b>	.13
Skilled, skillful	.22	.17	.25	.03	-.04	.20	.25	.25	<b>.33</b>	.25	.29	.22	<b>.34</b>	.19	<b>.33</b>
Traditional, conservative <sup>a</sup>	-.18	-.11	-.04	.04	.03	.14	<b>.37</b>	.26	.17	.05	.02	-.04	-.33	-.46	-.19
Narrow-minded, close-minded	-.19	-.20	-.14	-.26	-.27	-.33	.02	-.02	-.12	-.25	-.23	-.17	-.27	-.33	-.24
Exciting, fascinating	<b>.59</b>	<b>.53</b>	<b>.54</b>	.12	-.07	.23	.08	.16	.18	.15	.25	.24	<b>.40</b>	.26	.25
Prominent, influential	<b>.44</b>	<b>.43</b>	<b>.44</b>	-.02	-.11	.11	.17	.25	.23	.19	.27	.18	.16	.08	.23
Well-liked, likeable	<b>.40</b>	<b>.43</b>	<b>.35</b>	<b>.42</b>	.25	<b>.42</b>	.20	.17	.25	.21	.27	.21	.09	.07	.15

(continued on next page)

Cluster markers	Extraversion			Agreeableness			Conscientiousness			Emotional stability			Openness/intellect		
	AB5C(S1)	NEO (S1)	BFI (S3)	AB5C(S1)	NEO (S1)	BFI (S3)	AB5C(S1)	NEO (S1)	BFI (S3)	AB5C(S1)	NEO (S1)	BFI (S3)	AB5C(S1)	NEO (S1)	BFI (S3)
Admirable, impressive	<b>.44</b>	<b>.39</b>	<b>.33</b>	.13	-.05	.22	.16	.21	<b>.34</b>	.19	.28	.24	<b>.33</b>	.19	.19
Great, wonderful	<b>.38</b>	<b>.34</b>	<b>.36</b>	.17	.03	.23	.05	.11	.22	.15	.22	.29	.24	.18	.15
Weird, strange	.00	-.04	-.15	-.28	-.25	-.24	<b>-.36</b>	<b>-.32</b>	-.20	-.22	-.24	-.19	.26	.25	.17
Lucky, fortunate	.21	.29	.17	.24	.16	.18	.02	.10	.26	.23	<b>.30</b>	.03	.05	.13	.12
Ordinary, average	<b>-.33</b>	-.20	-.25	.09	-.23	.03	.01	-.10	-.16	-.08	-.18	-.18	<b>-.46</b>	<b>-.32</b>	-.25
Wealthy, well-to-do	.26	.24	.14	-.05	-.14	.01	.17	.26	.10	.21	.27	.13	.07	-.02	-.05
Attractive, good-looking	<b>.39</b>	<b>.37</b>	<b>.37</b>	.15	.01	.11	.10	.14	.12	.12	.23	.14	.19	.10	.10
Beautiful, pretty	<b>.33</b>	<b>.36</b>	<b>.35</b>	.27	.08	.16	.11	.09	.14	.07	.14	.14	.14	.15	.09
Youthful, young	<b>.30</b>	<b>.35</b>	<b>.31</b>	.08	-.06	.19	.03	.06	.25	.02	.02	.12	.19	.15	.08
Feminine, masculine, <sup>a</sup>	.00	.06	.04	<b>.48</b>	<b>.37</b>	.20	.00	.00	.14	-.10	-.08	-.25	-.03	.12	.02
Healthy, well	.14	.20	.21	.21	.10	<b>.30</b>	.16	.16	<b>.32</b>	.25	.25	.23	.05	.01	.11
Tired, exhausted	-.21	-.22	-.22	-.07	-.05	-.10	-.22	-.27	-.16	<b>-.42</b>	<b>-.46</b>	-.41	-.06	-.01	-.07
Slim, slender	.02	.10	.10	.04	.04	.04	.13	.21	.06	.21	.18	.07	-.05	-.01	.02
Short, little	-.02	.01	-.09	.08	.02	.00	.02	.04	.02	-.03	-.01	-.13	-.03	.00	-.06

Note: Items have ordered by their average correlation across the three Big Five inventories examined here (AB5C, NEO, BFI). The S1 vs. S3 in parentheses indicates whether the correlations were computed in sample S1 or S3, as described in the method section, and the symbol.

<sup>a</sup> Indicates that the item differed slightly in the S3 sample, as described in Section 7.1.  $N = 653$  for analyses with the NEO, 455 for analyses with the AB5C, and 507 for analyses with the BFI. Correlations  $> |r| = .30$  are shown in bold.

agentic behavior (*bold/assertive, loud/noisy*), and sense of humor (*funny/amusing*).

6.2.3.2. *Agreeableness*. The agreeableness scales were most associated with the clusters indicating concern for others (*kind-hearted/caring, giving/generous*) vs. self-serving dispositions (*selfish/self-centered, conceited/egotistical*). Agreeableness scales were also modestly associated with clusters indicating mannerly behavior (*rude/inconsiderate, courteous/polite*), irritability (*crabby/grouchy, touchy/temperamental*), antisocial tendencies (*unfriendly/cold, cruel/abusive*) and some aspects of positive emotionality (*sad/unhappy, positive/optimistic*).

6.2.3.3. *Conscientiousness*. The conscientiousness scales were most associated with clusters indicating organization (*disorganized/messy*) and dependability (*dependable/reliable, competent/capable*). The scales were also regularly associated with clusters indicating emotional and physical poise (*stable/well-adjusted, confident/self-assured vs. awkward/clumsy*).

6.2.3.4. *Emotional stability*. The emotional stability (vs. neuroticism) scales were most associated with clusters indicating tendencies to be worrisome (*tense/anxious, afraid/scared vs. calm/relaxed*), irritable (*touchy/temperamental, angry/hostile*), to show positive or negative emotionality (*happy/joyful, positive/optimistic, vs. lonely/lonesome*), and to be well-adjusted (*confident/self-assured, stable/well-adjusted*).

6.2.3.5. *Openness/intellect*. Although the AB5C Intellect and NEO openness measures take somewhat different interpretations of the “fifth-factor” in five-factor personality structures (John & Srivastava, 1999), they showed fairly similar patterns of correlates. The measures showed their highest correlations with self-views of having talent (*creative/imaginative, intelligent/smart*) and with clusters indicating unconventionality (*radical/rebellious vs. narrow-minded/close-minded, traditional/conservative*).

6.2.3.6. *Physical, evaluative and reputational characteristics*. We also found regular associations between the AB5C and NEO scales and the physical, evaluative, and reputational trait terms that are usually excluded from lexical studies (Saucier, 1997). Extraversion scales were associated with evaluations of being *prominent/influential, great/wonderful, attractive/good-looking, beautiful/pretty* and *youthful/young*. Both extraversion and openness/intellect scales were associated with seeing oneself as more interesting (*exciting/fascinating, admirable/impressive, vs. ordinary/average*). In addition, the cluster *well-liked/likeable* showed sizable correlations with both extraversion and agreeableness scales, the cluster *weird/strange* was associated with lower conscientiousness, agreeableness, emotional stability, and higher openness/intellect, and the cluster *lucky/fortunate* showed modest associations with measures of emotional stability and agreeableness.

## 6.3. Discussion

We conducted a cluster analysis of Saucier's (1997) list of the most frequently used person-descriptor adjectives in the English language in order to form a more empirically-based and comprehensive survey of the content that can be considered to define the lower-order structure of individual differences in the lexicon. The analysis resulted in the identification of a smaller set of 61 clusters that could be used to represent the content contained within the larger set of 504 adjectives. These analyses resulted in a smaller set of clusters that represented major dimensions of content (such as extraversion or agreeableness) in approximately the proportion with which they were represented in the larger set of

**Table 2**  
Reliability estimates of IIDL measures, and associations with gender and life satisfaction.

IIDL item	Big Five  rs  > .30	Alpha	Retest <i>r</i>	Gender ( <i>d</i> )			Life satisfaction ( <i>r</i> )		
				S1	S2	S1	S3	S4	S1
Outgoing, sociable <sup>a</sup>	E	.75	.77	.28	.14	.12	.13	<b>.33</b>	<b>.28</b>
Enthusiastic, excited	E	.68	.65	.20	.33	.36	<b>.25</b>	<b>.39</b>	<b>.32</b>
Bold, assertive	E	.56	.64	-.10	-.20	-.12	.12	<b>.24</b>	.18
Happy, joyful	E, A, S	.76	.64	.16	.29	.26	<b>.45</b>	<b>.48</b>	<b>.53</b>
Loud, noisy	E	.79	.63	.04	.00	.02	.06	.13	.08
Funny, amusing	E	.84	.57	.04	.08	-.26	.14	<b>.26</b>	.07
Brave, adventurous <sup>a</sup>	E,O	.69	.61	-.06	-.16	-.28	.09	.13	.06
Bashful, shy	-E, -S	.84	.73	-.14	-.08	.02	-.12	-.33	-.21
Kind-hearted, caring	A	.76	.57	<b>.43</b>	<b>.45</b>	<b>.40</b>	.11	.19	.14
Giving, generous	A	.74	.64	.37	<b>.47</b>	<b>.43</b>	.10	.18	.12
Pleasant, agreeable	A	.65	.48	.30	.20	.18	<b>.21</b>	<b>.25</b>	.19
Thankful, grateful	A	.81	.55	.28	.33	.32	.19	<b>.33</b>	<b>.26</b>
Affectionate, loving <sup>a</sup>	A, E	.75	.59	<b>.51</b>	<b>.43</b>	.28	.11	<b>.32</b>	<b>.22</b>
Courteous, polite	A, C	.79	.52	.24	.12	.18	.10	.08	.16
Truthful, honest	A	.85	.54	.22	.14	.08	.10	.19	<b>.20</b>
Rude, inconsiderate	-A	.77	.60	-.22	-.45	-.32	-.07	-.05	-.16
Selfish, self-centered	-A	.66	.57	-.28	-.27	-.12	-.04	.04	-.10
Unfriendly, cold <sup>a</sup>	-A, -E	.74	.52	-.32	-.43	-.24	-.09	-.14	-.22
Angry, hostile	-A, -S	.67	.58	-.20	-.39	-.26	-.22	-.04	-.28
Conceited, egotistical	-A	.71	.65	-.20	-.39	-.38	.08	-.07	-.02
Cruel, abusive	-A	.76	.49	-.26	-.47	-.18	-.11	-.05	-.16
Controlling, dominant	-A	.79	.68	-.08	-.16	.06	.02	.13	.06
Dependable, reliable	C	.68	.54	.34	.35	.08	.09	.17	<b>.20</b>
Practical, sensible <sup>a</sup>	C	.77	.53	.22	.10	-.10	.11	<b>.20</b>	.15
Competent, capable	C	.76	.56	.18	.33	.02	.13	<b>.22</b>	<b>.22</b>
Disorganized, messy <sup>a</sup>	-C	.76	.77	-.06	-.16	-.22	-.09	-.16	-.16
Undependable, unreliable <sup>a</sup>	-C	.74	.56	-.24	-.37	-.12	-.05	-.19	-.21
Awkward, clumsy	-C, -S	.70	.69	.26	.35	.36	-.12	-.12	-.14
Calm, relaxed	S	.69	.60	-.10	-.37	-.43	<b>.24</b>	<b>.22</b>	<b>.20</b>
Confident, self-assured	S, E	.73	.55	-.10	-.14	-.16	<b>.35</b>	<b>.40</b>	<b>.41</b>
Stable, well-adjusted	S, C, A	.84	.64	.10	.08	-.04	<b>.35</b>	<b>.40</b>	<b>.46</b>
Positive, optimistic	S, E, A	.72	.65	.06	.16	.12	<b>.35</b>	<b>.40</b>	<b>.45</b>
Tense, anxious <sup>a</sup>	-S	.75	.64	.16	.27	.28	-.25	-.29	-.24
Afraid, scared	-S	.75	.60	.14	.12	.28	-.27	-.22	-.20
Sad, unhappy	-S, -E, -A	.82	.65	.00	-.10	-.04	-.44	-.45	-.53
Touchy, temperamental	-S	.78	.61	.08	-.10	.12	-.21	-.19	-.17
Crabby, grouchy	-S, -A	.64	.55	-.10	-.12	.00	-.22	-.15	-.26
Lonely, lonesome	-S	.85	.66	.06	-.33	-.20	-.32	-.32	-.47
Creative, imaginative <sup>a</sup>	O	.90	.77	-.02	.04	.04	.02	.09	.03
Intelligent, smart	O	.78	.66	-.02	.08	-.26	.08	.25	.13
Radical, rebellious	O, -A, -C	.81	.59	-.12	-.27	-.20	-.10	-.08	-.12
Skilled, skillful <sup>a</sup>	O	.69	.59	-.18	-.20	-.30	.14	<b>.25</b>	<b>.23</b>
Traditional, conventional <sup>b</sup>	-O	.81	.68	-.06	.00	.08	.01	.13	.18
Narrow-minded, close-minded	-O, -A	.63	.61	-.30	-.31	-.14	-.01	.00	-.03
Exciting, fascinating	E	.75	.50	.04	-.02	-.12	<b>.23</b>	<b>.36</b>	.19
Prominent, influential <sup>a</sup>	E	.79	.60	-.22	-.06	-.24	.18	<b>.30</b>	<b>.22</b>
Well-liked, likeable	E, A	.70	.58	.24	.20	.02	<b>.21</b>	.14	<b>.26</b>
Admirable, impressive	E	.75	.59	-.02	-.02	-.14	.19	<b>.37</b>	<b>.31</b>
Great, wonderful	E	.71	.63	.10	-.06	-.06	<b>.21</b>	<b>.36</b>	
Weird, strange	-C	.87	.74	-.12	-.02	-.28	-.12	-.13	-.20
Lucky, fortunate	S	.78	.46	.04	.27		<b>.37</b>	<b>.37</b>	
Ordinary, average	-O, -E	.85	.66	.22	.22	.26	-.06	-.13	.00
Wealthy, well-to-do	-	.62		-.32	-.10		<b>.28</b>	<b>.31</b>	
Attractive, good-looking	E	.75	.76	.18	-.04		<b>.23</b>	<b>.25</b>	
Beautiful, pretty	E	.83		<b>.53</b>	.39		.17	<b>.25</b>	
Youthful, young	E	.81		.14	.37		.12	<b>.21</b>	
Feminine, unmasculine <sup>b</sup>	A	.70	.90	<b>1.47</b>	<b>2.10</b>	<b>1.49</b>	.01	.02	.10
Healthy, well	A	.86		.00	.10		<b>.27</b>	<b>.33</b>	
Tired, exhausted	-S	.79		.37	.14		-.24	-.13	
Slim, slender	-	.94		-.18	-.12		<b>.21</b>	.17	
Short, little	-	.69		.26	.37		.09	-.03	

Note: All statistically significant associations ( $p < .05$ ) are italicized. Column labeled "Big Five |rs| > .30" shows any AB5C or NEO measure that correlated >.30 with cluster markers in S1, and table is in same order as Table 1. E = extraversion, A = agreeableness, C = conscientiousness, S = emotional stability, O = openness/intellect. All Cohen's *d* for gender >.40 and correlations with SWLS >.20 are shown in bold.

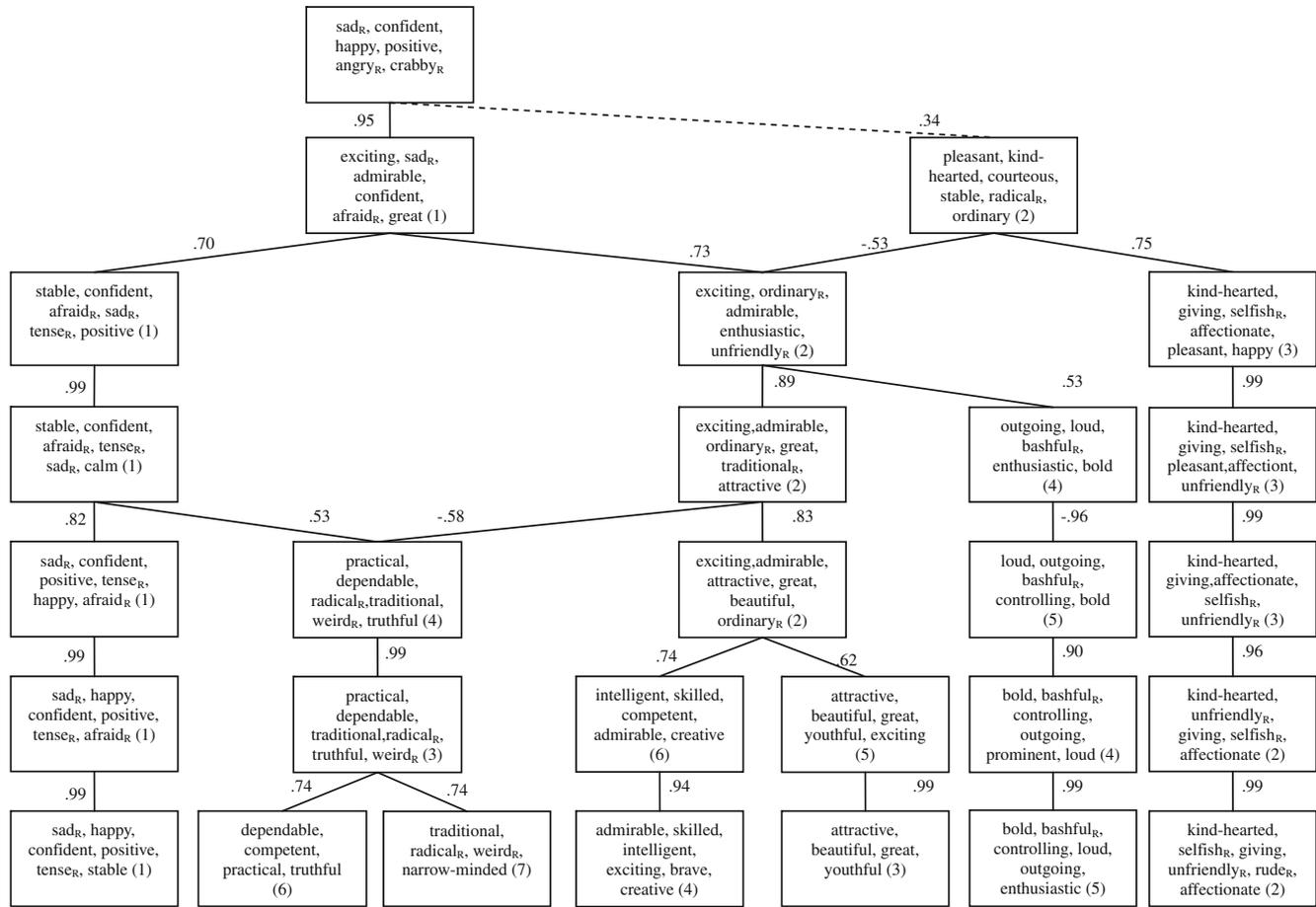
<sup>a</sup> Indicates that item was administered by a different item in S2.

<sup>b</sup> Indicates that item was administered by different cluster markers in S1; see Section 7.1 for more details. Blank cells indicate that the IIDL item was not administered in this subsample.

adjectives, indicating that the set of 61 clusters can stand as a suitable substitute for the larger lexical set without dramatically over-emphasizing or underemphasizing certain content.

As can be seen by a closer examination of Table 1, 57 of the 61 clusters showed at least one correlation of absolute magnitude

greater than .30 with either the NEO or AB5C measures of the Big Five and 41 of the 61 clusters showed at least one correlation >.40. Characteristics that were only weakly captured in the Big Five space were of particular interest because they have likely garnered less attention from personality psychologists. Some of these



**Fig. 1.** Pattern of factor emergence among 61 cluster markers. Solid lines show all correlations that exceed  $|r| = .50$  between factors from adjacent extractions. If no correlations reached this magnitude, the highest correlation is shown in dotted lines. The number within each box indicates the order in which the factor emerged within the analysis.

characteristics included adventurousness, dominance, awkwardness, honesty, rebelliousness, and cruelty. Additionally, we identified numerous clusters that reflected evaluative and reputational traits (self-perceptions of being exciting, lucky, strange, rich, impressive, and prominent) and physical traits (health, youth, weight, height, and attractiveness) that are usually excised from lexical studies of the personality trait structure (Saucier, 1997).

**7. Study 2: Construction and validation of the Inventory of Individual Differences in the Lexicon**

Whereas the primary goal of Study 1 was to generate a more comprehensive list of the content clusters that can be identified within the lexicon, the primary goal of Study 2 is to establish a suitable measure of these clusters and address some likely concerns with its use. We thus created an instrument where the cluster markers used in Study 1 were combined into single-item scales, which we termed the Inventory of Individual Differences in the Lexicon (IIDL). We next estimated the level of reliability that exists when assessing IIDL items by estimating test–retest correlations for the items over a very short retest interval, and explored the extent to which the IIDL items function similarly to the cluster markers identified in Study 1 by relating the items to the Big Five Inventory (BFI; John & Srivastava, 1999). Finally, we explored how the items of the IIDL were associated with gender and the Satisfaction with Life Scale (SWLS; Pavot & Diener, 1993) in three

samples, in order to illustrate how comprehensive instruments can reveal relationships that will likely be missed through the sole use of broad personality scales.

*7.1. Method*

*7.1.1. Development of the Inventory of Individual Differences in the Lexicon*

Similar to a number of other comprehensive instruments (e.g., Block, 1961; Westen & Shedler, 1999), we created a measure we termed the Inventory of Individual Differences in the Lexicon (IIDL) in which each cluster identified in Study 1 was represented by a single item. To make the meaning of the item clearer than might be possible with a single adjective (Goldberg & Kilkowski, 1985), we chose to form the items of the IIDL by combining the two cluster markers used in Study 1 into a single item, separated by commas. For instance, the height cluster identified in Study 1 was represented by the cluster markers *short* and *little*, and therefore was represented in the IIDL by the item “short, little.” Two exceptions were made: the cluster represented by adjectives *conservative*, *traditional*, *liberal* (rev), and *old-fashioned* was represented by the item “traditional, conventional,” after it was determined that the term *conventional*, which was not included in the original adjective set, was a better synonym of *traditional* and would better capture the core of the cluster, and the cluster represented by the terms *feminine*, *masculine* (rev), *motherly*, and *competitive* (rev) was represented by the item “feminine, unmasculine” after determin-

ing that the other cluster terms were poor synonyms of *feminine*, and to maintain the use of synonymous terms in creating all items. As each cluster identified in Study 1 was represented by a single item, this resulted in an instrument containing 61 items. The items forming the inventory can be seen in Table 2.

### 7.1.2. Participants and procedure

We report data collected from three samples, in addition to presenting some further analyses from the sample used in Study 1 (S1). We describe these additional samples below.

**7.1.2.1. Sample 2 (S2).** A total of 343 freshman students participated in a broader study concerning the interplay of personality characteristics and social relationships, particularly roommate relationships. Of these participants, 204 (60%) were females. The study was initiated near the end of the spring semester, and individuals were offered \$15 for completing the study at two time points a couple days apart. The mean lag between the first and second testing interval was 5.2 days ( $SD = 3.2$ ); the median retest interval was 4 days.

In completing the surveys, participants indicated who they would be living with the following academic year. As a major focus of this study was on roommate dyads, participants were only contacted to complete follow-up surveys if both the participant and the participant's roommate for the following academic year had completed the initial assessment ( $N = 190$ ). For the fall follow-up, these individuals were offered \$25 for completing the survey twice approximately a week apart. A total of 140 people completed the fall follow-up. In this wave, the mean lag between the first and second testing interval was 4.4 days ( $SD = 3.4$ ); and the median retest interval was 3 days. For the final spring follow-up, these same 190 individuals were offered \$30 to complete the study two times approximately a week apart, and a total of 147 people completed the spring follow-up. In this wave, the mean lag between the first and second testing interval was 6.1 days ( $SD = 3.7$ ), and the median retest interval was 5 days.

The central analyses of interest with the S2 sample concerned estimating the test–retest reliability of the items over very short time-spans. In all administrations, the materials of the survey were administered online, and participants completed the study at their convenience within the 2 weeks that the materials were made available. As part of the survey, participants completed an early version of the IIDL, using the instruction “How much do the traits describe you in general?” by rating the items on a 7-point scale ranging from 1 (*Extremely uncharacteristic*) to 7 (*Extremely characteristic*). As the cluster markers used in the IIDL were modified somewhat between the first and second waves of the study, we added IIDL items that were not surveyed in the first administration, except in cases where an item could be identified in the older version of the IIDL that contained one of the adjectives used to represent the item in the final version seen in Table 2 (e.g., the older items *outgoing/extraverted* and *prominent/well-known* were used to represent the IIDL item measured by the terms *outgoing/sociable* and *prominent/influential* in Table 2).<sup>2</sup>

<sup>2</sup> The IIDL items are discussed in the text by italicizing the adjectives and separating them by a slash mark (*outgoing/sociable*) despite the fact that participants are presented with the same items in a slightly different format (“outgoing, sociable”). This convention was adopted to avoid a confusing proliferation of commas and quotation marks in the text. The early version of the IIDL was constructed from a cluster analysis that included terms outside of the 504 “most familiar terms” identified by Saucier (1997), and thus changed when the cluster analysis was restricted to these terms. The similar items from an earlier version of the IIDL are (in the order the items are given in Table 2): *outgoing/extraverted*, *brave/fearless*, *unsympathetic/unfriendly*, *affectionate/passionate*, *messy/sloppy*, *unreliable/undependable*, *level-headed/sensible*, *tense/nervous*, *creative/artistic*, *skilled/talented*, and *prominent/well-known*.

In these cases, the old item was retained in the second wave in order to enable test–retest analyses across the waves of the study using strictly parallel items, and the old items that are not perfectly parallel with the final version of the IIDL are indicated in Table 2. Participants did not rate the IIDL items that concerned more physical or general status-related characteristics in this sample.

**7.1.2.2. Sample 3 (S3).** An additional sample consisted of 507 participants who were recruited through an introductory psychology subject pool. Of these participants, 298 were female (59%). Participants completed both the IIDL and BFI on the internet, where the items of both inventories were presented in an order that was randomized for each participant. All BFI items were rated first, and then the IIDL items were rated; all ratings were made on a scale ranging from 1 (*Extremely uncharacteristic of me; I am never like this*) to 9 (*Extremely characteristic of me; I am always like this*). Each item was presented on the screen one term at a time, and clicking on a response option caused the next item to be presented. Cronbach's alpha for the BFI scales ranged from .84 to .88. Participants also completed a range of other items unrelated to the current analyses.

**7.1.2.3. Sample 4 (S4).** Our final sample consisted of participants who completed the IIDL through a publicly available survey on the internet. The survey was called “The Celebrity Similarity Test”, and individuals were told that by completing the survey they could learn how similar they were to a range of celebrities and fictional characters. Participants were only counted within this sample if they indicated that they had never completed the survey before, if they reported that they had answered the questions truthfully, if they indicated that their answers could be used in research, and if they indicated their name and age. A total of 1479 participants met these conditions, of which 782 were female (53%). In this survey, most of the IIDL items concerning more physical traits were omitted as well as some other evaluative items.

**7.1.2.4. Satisfaction with life scale.** A subset of participants in the S3 sample ( $N = 259$ ) and most participants of the S4 sample ( $N = 1471$ ) completed the five-item Satisfaction with Life Scale (SWLS; Pavot & Diener, 1993;  $\alpha = .83$ ). The SWLS was also completed by a subset of 561 participants from the S1 sample approximately eight years after the administration of the common trait adjectives.

## 7.2. Results and discussion

### 7.2.1. Test–retest stabilities of IIDL items

The test–retest stabilities of the IIDL items over retest periods of less than a week were estimated for each of the three waves of the S2 sample. Since the rank-ordering of the most to least stable IIDL items was quite consistent across the three waves (the average column-vector correlation of stability estimates was .65), stability estimates from the three waves were averaged together ( $\alpha = .85$ ). As seen in Table 2, the short test–retest stabilities ranged from .46 to .90 (mean  $r = .62$ ) across the 54 IIDL items assessed in the sample. The IIDL items with the highest test–retest stabilities included the items *feminine/unmasculine*, *outgoing/sociable*, *disorganized/messy*, *creative/imaginative*, and *attractive/good-looking* (all  $r_s \geq .75$ ). The IIDL items with the lowest test–retest stabilities were the items *ordinary/average*, *pleasant/agreeable*, *cruel/abusive*, and *exciting/fascinating* (all  $r_s \leq .50$ ).

Although we primarily consider the IIDL items without aggregation, they can be used to create broad, multi-item Big Five scales (or other superfactor scales) if so desired. To estimate how the retest correlations of longer scales differed from those of single items, we constructed five-item scales to roughly represent the Big Five

using the five clusters that correlated most highly with each BFI trait scale as shown in Table 1, and then estimated the test–retest reliabilities of these scales in the S2 sample. The test–retest reliabilities for the five-item scales were .84, .71, .73, .82, and .77 for extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience, respectively. This corresponded to test–retest reliabilities of .69, .58, .63, .68, and .69 for the average test–retest reliabilities of the five-items of each dimension considered without aggregation. This result indicates that IIDL items can be combined to create Big Five composite scores which will have modestly higher test–retest reliabilities.

### 7.2.2. Relationship between IIDL and Big Five measures

We next examined the relationships between the IIDL items and BFI scale scores among the S3 participants (Table 1). Of central interest was the extent to which items were related to the Big Five in a similar pattern to the cluster markers used in Study 1. Since the BFI and IIDL items were collected contemporaneously, we also examined how closely single-item scales could serve as proxies for overall BFI scale scores by examining the magnitude of the correlations for the most highly correlating items for each trait.

As can be seen, 52 of the 61 IIDL items had at least one correlation with a BFI scale exceeding a magnitude of  $|r| = .30$  among the S3 participants, although the number differed dramatically by trait. A total of 21 items showed correlations of this magnitude with BFI Extraversion, 28 with BFI Agreeableness, 15 with BFI Conscientiousness, 14 with BFI Emotional Stability, and only three with BFI Openness. Since the IIDL was designed to represent content in rough proportion to its level of representation in the lexicon, we interpret the dramatically lower number of openness-related items as consistent with the finding that openness is generally less represented by common lexical terms than the remaining Big Five traits (John & Srivastava, 1999).

Within each Big Five domain, there were regularly very strong correlations between the BFI measure and certain IIDL items. For instance the item *outgoing/sociable* correlated .79 with BFI Extraversion, *rude/inconsiderate* correlated  $-.65$  with BFI Agreeableness, *disorganized/messy* correlated  $-.61$  with BFI Conscientiousness, *tense/anxious* correlated  $-.82$  with BFI Emotional Stability, and *creative/imaginative* correlated .72 with BFI Openness. The observation of several correlations exceeding .70 between single IIDL items and BFI scales indicates that single-item scales can serve as useful substitutes for superfactor scales.

### 7.2.3. Relationship between gender and individual differences

Our next analyses were conducted to illustrate how gender differences in traits could be better understood by using a comprehensive instrument than by simply reporting associations with superfactor measures. Using the S3 sample, we first estimated gender differences on the Big Five by using the BFI and five-item IIDL estimates of the Big Five described earlier; we report Cohen's  $d$  statistics. Largely replicating previous research (Feingold, 1994; Schmitt, Realo, Voracek, & Allik, 2008), we found women to be significantly higher on agreeableness ( $ds = .45/.50$  for the BFI/IIDL scales respectively) and conscientiousness (.37/.33), and less consistently found women to describe themselves as lower on emotional stability ( $-.56/-.10$ ). Women did not differ from men notably in their levels of extraversion (.04/.18) or openness ( $-.14/.04$ ). To illustrate how gender differences in personality characteristics might be better illuminated through an instrument assessing a wide range of narrow attributes, we present the correlations between gender, the cluster markers used in the S1 sample, and the IIDL items used in the S3 and S4 samples (see Table 2). Given the similarity of the associations between gender and the IIDL across the three samples (the average column-vector correlation of

gender/IIDL associations across samples was .90), we focus our discussion on relationships that were commonly observed across samples.

Despite the negligible associations between the broad, multi-item BFI and IIDL extraversion scales and gender, we found interesting heterogeneity in how more specific aspects of extraversion were associated with gender. Across samples, women generally described themselves as having higher levels of positive affect than men (*happy/joyful, enthusiastic/excited*), but lower levels of agentic aspects (*brave/adventurous, bold/assertive*).

Within the domain of agreeableness, women described themselves as being more agreeable than men for nearly every aspect assessed. However, some aspects regularly showed larger or more negligible differences than others. Women tended to report the largest differences from men on items related to affection for others (*kind-hearted/caring, giving/generous, affectionate/loving*), but about equal levels of mannerly or integrity-associated characteristics (*truthful/honest, courteous/polite*).

Within the domain of conscientiousness, women described themselves in the more conscientious direction for most traits (*dependable/reliable, competent/capable*). However, in all samples women also described themselves as significantly more *awkward/clumsy* than men.

Given the frequently documented gender differences found between men and women on broad emotional stability or neuroticism scales, there were surprisingly few associations between gender and aspects of emotional stability. In all samples, women generally described themselves as more *tense/anxious* and less *calm/relaxed* than men. However, women reported equal levels of most other aspects of emotional stability as men (*sad/unhappy, touchy/temperamental, positive/optimistic, stable/well-adjusted*), and actually reported being less *lonely/lonesome*.

There was considerable heterogeneity in the direction and magnitude of gender differences across different aspects of openness and intellect. Women consistently described themselves as less *radical/rebellious* and *skilled/skillful* than men, indicating lower openness, but also as less *narrow-minded/close-minded*, indicating higher openness. Men and women reported indistinguishable levels of the remaining aspects of openness/intellect.

We also explored how gender was associated with indicators of the more evaluative and physical characteristics typically excluded by personality scales. Unsurprisingly, women described themselves as more *beautiful/pretty, short/little, and feminine/unmasculine* than men. However, women also generally described themselves as more *well-liked/likeable* and *ordinary/average* than men, and as less *prominent/influential*.

### 7.2.4. Relationship between life satisfaction and individual differences

We next explored how the IIDL was associated with life satisfaction, again contrasting this with the findings that would be found with the sole use of Big Five scales. In the S3 sample, we first correlated the BFI and five-item IIDL scale measures of the Big Five with the SWLS, and found life satisfaction to be associated with higher levels of extraversion ( $r = .36/.42$  for the BFI/IIDL correlations respectively), agreeableness (.20/.14), conscientiousness (.26/.25), and emotional stability (.37/.37), and was unassociated with openness ( $-.03/-.09$ ). We then used the unaggregated IIDL items to examine how the SWLS was associated with more molecular traits related to and beyond the Big Five. Given the similarity of the associations across the three samples (the average column-vector correlation of SWLS/IIDL associations was .94), we discuss the general findings across studies.

Not surprisingly, we found that in the domain of extraversion the SWLS was most highly associated with the item *happy/joyful*. Beyond this, the results indicated that aspects of extraversion associated with agentic behavior (*brave/adventurous, loud/noisy, bold/*

assertive) had consistently weaker associations with life satisfaction than aspects associated with positive affectivity (*enthusiastic/excited*) and sociability (*outgoing/sociable*).

Although agreeableness aspects were generally associated with higher life satisfaction, there was substantial and regular variability in the level of these associations. Life satisfaction was most associated with interpersonal warmth (*pleasant/agreeable, affectionate/loving*), but its correlation with mannerly tendencies (*courteous/polite*) was negligible. Similarly, life satisfaction was negatively associated with outwardly antisocial tendencies (*unfriendly/cold, angry/hostile*) but was essentially unassociated with more internal self-serving tendencies (*selfish/self-centered, conceited/egotistical*).

Although there was little heterogeneity in how life satisfaction was associated with aspects of conscientiousness and emotional stability, there was interesting heterogeneity in how intellect and openness aspects were associated with life satisfaction, with different aspects of openness being positively, negatively, and not associated with life satisfaction. Life satisfaction was consistently positively associated with self-perceptions of being *intelligent/smart* and *skilled/skillful*, was consistently unassociated with being *creative/imaginative*, and was consistently negatively associated with being unconventional (*radical/rebellious* vs. *traditional/conventional*).

We also found that the SWLS was associated with self-perceptions of nearly every evaluative and physical trait assessed. Among the more physical aspects, individuals who described themselves as more *attractive/good-looking, beautiful/pretty, youthful/young, healthy/well, and slim/slender* reported higher levels of life satisfaction. Life satisfaction was regularly associated with self-perceptions of being exceptional or having high status (*exciting/fascinating, prominent/influential, admirable/impressive, great/wonderful*) and with evaluations of fitting in with others (*well-liked/likeable* vs. *weird/strange*).

## 8. General discussion

In the current investigation, we were primarily interested in delineating a more comprehensive lower-order structure of personality and showing how measures of this content can illuminate relationships between personality and other variables of interest that are not typically seen with broad measures. We modified the procedures outlined by Block (1961) and Peabody (1987) for generating inventories for the more comprehensive measurement of individual differences. We summarize here what we see as some of largest questions concerning the set of clusters identified, the inventory created to measure them, and the use of similar comprehensive approaches more generally.

### 8.1. Is the identified cluster set truly comprehensive?

The first major step in creating a comprehensive inventory concerns the identification of the content that should be measured. We believe that the procedures used in the current study represent a more clearly empirical and replicable rationale for the identification of a comprehensive list of individual difference dimensions than past approaches (Block, 1961; Peabody, 1987), and that the use of the pool developed by Saucier (1997) offers the unique strength of constructing clusters by focusing on a well-defined pool of frequently used trait terms. The methods used here should also offer some benefits relative to approaches that attempt to locate facets of superfactors (e.g., Hofstee, de Raad, & Goldberg, 1992; Roberts Bogg, Walton, Chernyshenko, & Stark, 2004; Saucier & Ostendorf, 1999), in that they should increase the likelihood of

identifying trait aspects that do not initially have high loadings on the superfactors. As hoped, our more mechanical, data-driven methods were able to identify both content considered central to major personality dimensions and also numerous dimensions that are only weakly captured by the Big Five dimensions, such as adventurousness, sense of humor, dominance, honesty, cruelty, rebelliousness, and masculinity/femininity. The analysis also identified content considered to represent major dimensions in the structure of evaluations and social effects, such as evaluations of being exciting, likeable, strange, average, and admirable (Saucier, 2010; Wood et al., 2007). The procedure also cleanly separated many of the distinguishable dimensions that could be identified within each Big Five domain. For instance, sociability, positive affectivity, and assertiveness aspects of extraversion were readily separated, as were creativity, intellect, and unconventionality aspects of openness to experience.

This said, the procedure failed to capture some content that most would consider to represent important individual differences. As a point of comparison, we can compare the 61 clusters identified to content that past researchers have pointed to as narrower aspects of broad superfactors such as the Big Five. For instance, although many of the dimensions identified by Roberts and colleagues (2004) as narrower aspects of conscientiousness were clearly identified here (orderliness, conventionality, and reliability), others were not (impulsiveness, industriousness, decisiveness, punctuality). Similarly, an exploratory analysis we conducted linking the clusters identified in Study 1 to the NEO facet measures available in the ESCS dataset used in Study 1 indicated that several NEO facet scales, including Excitement-Seeking, Achievement-Striving, Impulsivity, and the Fantasy, Aesthetic, Feeling, and Actions facets of the Openness scale did not correlate substantially (i.e., did not correlate above  $|r| = .40$ ) with any of the 61 clusters identified here. (Conversely, a large number of the clusters identified here did not correlate at this magnitude with any of the 30 NEO facets, such as the clusters *truthful/honest, loud/noisy, awkward/clumsy, creative/imaginative, radical/rebellious, weird/strange, and feminine/masculine<sub>(R)</sub>*.<sup>3</sup>)

These omissions indicate that the procedure outlined here missed some content that most would consider important dimensions of individual differences. It is important to ask whether this represents a limitation of the mechanical clustering procedure detailed here to identify content in the lexicon, or of the range of content contained in the lexical set. Although we believe our procedure for identifying content for a comprehensive inventory represents a substantial improvement over past approaches (e.g., Block, 1961; Peabody, 1987), we do not claim that the method is ideal and have described it in detail to aid investigators who wish to replicate this methodology and suggest improvements. For instance, our decision to only select facets that had at least three items was driven in part by the goal of arriving at an instrument which would have a small enough number of distinct attributes that all could be practically administrable in a short time span. However, if two-adjective clusters were allowed, an additional 23 clusters would have been identified, some of which would capture some of the omitted aspects listed above. Some of the more interesting dimensions in this supplemental set include the clusters *impulsive/spontaneous, determined/persistent, strict/firm, sarcastic/critical, trusting/suspicious<sub>R</sub>*, and *independent/self-sufficient* (see Appendix A for a full list).

The 61 content clusters identified here should thus be considered to form an open list of the content that can be identified with-

<sup>3</sup> A correlation matrix detailing how all 30 NEO facets are associated with the 61 cluster markers within the ESCS sample is available from the first author.

in the lexicon which will shift somewhat as decision rules are refined, and certainly even when applying the same extraction rules to other lexical datasets, especially in more distinct lexical datasets formed from terms in additional languages, from different types of ratings (e.g., peer ratings), or using different samples of lexical terms (e.g., trait-nouns). One interesting possibility that may emerge as these studies accumulate is that some socially important dimensions such as impulsiveness, general trust, or aspects of openness to experience will continue to be unidentified as factors that are well-represented in lexical datasets as such studies accumulate. Such findings could help to better define suspected limitations of looking to the concepts encoded in the lexicon as a primary means of identifying the important ways people differ from one another (Block, 1995; Block, *in press*).

### 8.2. How reliable and valid are single-item assessments?

Our analyses suggest that although the reliability of our assessments is decreased somewhat by the use of one or two-item scales, the cluster markers still show acceptable reliability. The single-item scales used to mark the clusters in Study 2 showed retest correlations of .65 over a couple days, and the two-item markers in Study 1 showed alphas comparable in magnitude to those found for numerous longer scales (average  $\alpha = .76$ ). As the expected reduction in a measure's correlations with other variables due to a measure's unreliability is equal to the square root of the reliability coefficient when reliability is properly estimated (Cohen et al., 2003; Nunnally & Bernstein, 1994), these levels reliability indicate that one or two-item markers of the clusters should generally show correlations that are about 81–87% of the size of the correlations that would be found with perfectly reliable estimates of a person's self-rating for the dimensions. Although we showed that longer scales can be created through combinations of IIDL items which have somewhat higher reliabilities, the reliabilities of our single-item measures are probably higher than most would suspect, and, as shown in Study 2, aggregating this content to form scales comes with the cost of losing information about the differential associations between similar trait aspects with variables of interest.

Further pointing to the reliability and validity of single-item trait markers, we were frequently able to find single IIDL items with correlations  $>.70$  with other Big Five dimensions. Correlations of this magnitude indicate that short, one-item adjective markers are better proxies of longer scales than most researchers would suspect, particularly given that the dimensions we identified and measured were not selected with any intention of measuring Big Five or other superfactor dimensions. Finally, our pattern of findings linking life satisfaction and gender to the cluster markers were highly consistent across three samples and largely paralleled the substantive findings of previous research linking narrower trait measures to gender and life satisfaction (e.g., Costa et al., 2001; Schimmack, Oishi, Furr, & Funder, 2004). This indicates that the somewhat lower reliability of one- or two-item markers does not appreciably reduce our ability to use such scales to investigate how different traits are related to variables of interest.

### 8.3. The utility of comprehensive inventories

Comprehensive inventories which make fine discriminations between related traits are particularly useful for showing how similar traits differ in their relationships with a variable of interest. For instance, across three samples in the current research, replicable gender differences were found in opposite directions for traits within the domains of extraversion (sociability vs. assertiveness), conscientiousness (dependability vs. clumsiness), emotional

stability (anxiety vs. loneliness), and openness (open-mindedness vs. rebelliousness). Although some of these discriminations have been discussed in past research (e.g., Costa et al., 2001; Feingold, 1994), others appear to be relatively novel findings due to the inclusion of a broader range and greater number of distinct trait aspects than those employed in standard superfactor measures. We also found surprisingly fine distinctions in how similar traits were associated with life satisfaction. For instance, life satisfaction was consistently associated with perceived intelligence but not creativity, and with unfriendliness but not conceitedness. The substantial heterogeneity of “within-domain” trait associations points to the limitations of glossing over these discriminations when using a single, broad factor score. Further, the impressive consistency of these differences across our three samples speaks against the argument that the aggregation of similar-but-distinguishable trait aspects is necessary in order to document replicable or stable relationships. Although the Big Five and similar superfactor frameworks are certainly useful, findings like this suggest that these frameworks may be used best as a means of organizing the multitude of distinguishable traits people use to describe one another, rather than as a recommendation that the core vectors are the only traits we need to assess.

The most immediate goal of the current project was to refine earlier procedures that have been used to create comprehensive inventories. Whereas earlier methods for the construction of comprehensive inventories have largely involved intuitive or subjective decisions to identify the content that should be measured, we show that comprehensive inventories can be created with more mechanical procedures using lexical datasets and cluster analysis. This approach also identified numerous dimensions that are not typically identified through the more common approach of identifying facets of superfactors. Nevertheless, comparisons of these two contrasting approaches to delineating the lower-order structure of personality and individual differences warrant additional attention. We also found that the single-item analyses that are frequently employed to make the broad range of content in such inventories practically assessable in research contexts (e.g., Block & Block, 2006; Funder et al., 2000; Shedler & Westen, 2007) are more valid and reliable than many researchers may suspect.

A broader goal of the project was to present a more general argument for the advantages of comprehensive inventories for addressing substantive questions. Through demonstrations linking the diverse content we identified in the lexicon to gender and life satisfaction, we argue that the use of comprehensive inventories can reveal a more nuanced picture of how personality traits are related to variables of interest than is generally possible with superfactor measures. Simple descriptive studies linking personality measures to variables of interest continue to be important in the field (Funder, 2001). However, given the frequency with which researchers address the question of “how is personality associated with X?” by correlating a Big Five measure with their variable of interest, we don't feel it is an exaggeration to say that a number of old questions may be usefully addressed anew by simply showing how the variable is associated with a personality measure that assesses more than five traits. We hope that such demonstrations will serve to encourage the wider use of comprehensive inventories in explorations of how personality traits are associated with variables of interest.

### Acknowledgments

The authors would like to thank R. Michael Furr, Peter Harms, and Brent W. Roberts for their comments on an earlier version of this paper.

## Appendix A

Clusters extracted from cluster analyses (Study 1).

- [1] Social, **social**, **outgoing**, popular, unpopular<sub>R</sub>
- [2] **Excited**, **enthusiastic**, expressive, eager
- [3] **Assertive**, aggressive, **bold**, forward
- [4] Satisfied, **happy**, **joyful**, secure, glad, pleased, cheerful, comfortable, encouraged
- [5] Humorous, **funny**, **amusing**, comical, hilarious, witty, entertaining, laughing, good-humored
- [6] **Affectionate**, **loving**, passionate
- [7] **Loud**, **noisy**, quiet<sub>R</sub>, soft-spoken<sub>R</sub>
- [8] **Brave**, **adventurous**, courageous, daring
- [9] **Bashful**, **shy**, talkative<sub>R</sub>
- [10] **Kind-hearted**, kind, compassionate, **caring**, sympathetic, understanding, helpful, thoughtful, sensitive
- [11] **Giving**, **generous**, stingy<sub>R</sub>
- [12] **Pleasant**, **agreeable**, cooperative, good-natured, nice
- [13] **Feminine**, **masculine**<sub>R</sub>, motherly, competitive<sub>R</sub>
- [14] **Polite**, **courteous**, considerate, gracious
- [15] **Thankful**, **grateful**, appreciative
- [16] **Truthful**, trusted, **honest**, decent, sincere
- [17] **Selfish**, **self-centered**, stuck-up, snobbish, greedy, arrogant
- [18] Lovable, **unfriendly**, **cold**, sweet
- [19] **Inconsiderate**, impolite, **rude**, thoughtless
- [20] **Egotistical**, **conceited**, cocky
- [21] **Cruel**, **abusive**, violent, mean, dangerous, frightening
- [22] **Dominant**, **controlling**, bossy, demanding, intimidating
- [23] **Angry**, bitter, **hostile**, furious
- [24] **Practical**, **sensible**, level-headed, realistic, rational, logical, reasonable
- [25] **Competent**, **capable**, effective, useful, resourceful
- [26] **Dependable**, **reliable**, responsible, trustworthy
- [27] **Disorganized**, neat<sub>R</sub>, **messy**, organized<sub>R</sub>, sloppy, careless
- [28] **Undependable**, **unreliable**, irresponsible, unfaithful
- [29] Self-confident, **confident**, **self-assured**, unsure<sub>R</sub>, insecure<sub>R</sub>, certain, indecisive<sub>R</sub>, wishy-washy<sub>R</sub>
- [30] **Stable**, disturbed<sub>R</sub>, **well-adjusted**, unstable<sub>R</sub>
- [31] **Relaxed**, easy-going, **calm**, laid-back, peaceful, carefree
- [32] **Positive**, negative<sub>R</sub>, **optimistic**
- [33] **Afraid**, **scared**, confused, uncomfortable
- [34] **Tense**, **anxious**, nervous, frustrated, self-conscious
- [35] **Sad**, **unhappy**, troubled, disappointed, depressed, upset, bored
- [36] Moody, **temperamental**, **touchy**, defensive
- [37] **Crabby**, grumpy, **grouchy**, cranky, irritable, complaining
- [38] **Lonely**, **lonesome**, alone, neglected, heartbroken
- [39] **Awkward**, ungraceful, graceful<sub>R</sub>, **clumsy**
- [40] **Creative**, talented, artistic, **imaginative**, gifted, clever, original
- [41] **Smart**, intellectual, **intelligent**, knowledgeable, bright, brilliant, educated, wise
- [42] **Radical**, **rebellious**, disobedient, proper<sub>R</sub>, controversial
- [43] **Skilled**, **skillful**, accomplished, successful
- [44] **Conservative**, **traditional**, liberal<sub>R</sub>, old-fashioned
- [45] **Close-minded**, **narrow-minded**, open-minded<sub>R</sub>, open<sub>R</sub>
- [46] **Exciting**, extraordinary, **fascinating**, awesome, remarkable
- [47] Well-known, **influential**, famous, **prominent**, powerful, distinguished, important
- [48] **Likeable**, **well-liked**, pleasing,
- [49] **Admirable**, excellent, outstanding, **impressive**
- [50] **Great**, **terrific**, wonderful
- [51] **Lucky**, unlucky<sub>R</sub>, **fortunate**
- [52] **Wealthy**, **well-to-do**, rich, prosperous, poor
- [53] **Weird**, **strange**, normal<sub>R</sub>, abnormal, crazy
- [54] **Ordinary**, **average**, unusual<sub>R</sub>, complicated<sub>R</sub>
- [55] **Good-looking**, **attractive**, sexy, appealing, unattractive<sub>R</sub>, desirable, handsome, seductive, adorable
- [56] Gorgeous, **beautiful**, lovely, **pretty**, glamorous, cute, charming
- [57] Old<sub>R</sub>, elderly<sub>R</sub>, **youthful**, **young**
- [58] Disabled<sub>R</sub>, **well**, **healthy**, handicapped<sub>R</sub>

- [59] **Slim**, **slender**, thin<sub>R</sub>, chubby<sub>R</sub>, skinny, fat<sub>R</sub>, big<sub>R</sub>
- [60] **Tired**, **exhausted**, sleepy, energetic<sub>R</sub>
- [61] **Short**, **little**, tall<sub>R</sub>, tiny

*Additional two-item clusters:* ashamed/humiliated, careful/cautious, casual/informal, cheap/stingy, determined/persistent, direct/straight-forward, dumb/stupid, efficient/thorough, evil/corrupt, faithful/loyal, good-for-nothing/insane, hard/rough, hard-working/productive, hot-tempered/short-tempered, impulsive/spontaneous, independent/self-sufficient, jealous/possessive, lively/playful, prompt/punctual, retarded/senile, sarcastic/critical, strict/firm, trusting/suspicious<sub>R</sub>.

*Note:* clusters are presented in the same order as given in Tables 1 and 2. Symbol “R” indicates that the item is negatively related to other cluster markers. Bold items within each cluster indicate adjectives that were used as cluster markers in Study 1; italicized items were replaced with a different item in Study 2. “Additional two item clusters” show remaining clusters defined by two items and hence not considered “well-represented” in the lexicon by the decision rules used in Study 1.

## References

- Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis*. Beverly Hills, CA: Sage.
- Ashton, M., Lee, K., & Goldberg, L. (2004). A hierarchical analysis of 1710 English personality-descriptive adjectives. *Journal of Personality and Social Psychology*, 87, 707–721.
- Bem, D., & Funder, D. (1978). Predicting more of the people more of the time: Assessing the personality of situations. *Psychological Review*, 85, 485–501.
- Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research*. Oxford, England: Thomas.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117, 187–215.
- Block, J. (in press). The five-factor framing of personality and beyond: Some ruminations. *Psychological Inquiry*.
- Block, J., & Block, J. (1980). *The California child Q-set*. Palo Alto, California: Consulting Psychologists Press.
- Block, J., & Block, J. (2006). Nursery school personality and political orientation two decades later. *Journal of Research in Personality*, 40, 734–749.
- Cattell, R. (1944). A note on correlation clusters and cluster search methods. *Psychometrika*, 9, 169–184.
- Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the sixteen personality factor questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavior sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Colvin, C., & Longueuil, D. (2001). Eliciting self-disclosure: The personality and behavioral correlates of the opener scale. *Journal of Research in Personality*, 35, 238–246.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P., Terracciano, A., & McCrae, R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81, 322–331.
- DeYoung, C., Quilty, L., & Peterson, J. (2007). Between facets and domains: 10 Aspects of the Big Five. *Journal of Personality and Social Psychology*, 93, 880–896.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116, 429–456.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52, 197–221.
- Funder, D. C., Furr, R. M., & Colvin, C. R. (2000). The riverside behavioral Q-sort: A tool for the description of social behavior. *Journal of Personality*, 68, 451–489.
- 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. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.
- Goldberg, L. R. (2006a). *Oregon Research Institute technical report* (Vol. 46).
- Goldberg, L. R. (2006b). Doing it all bass-ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality*, 40, 347–358.
- Goldberg, L., & Kilkowski, J. (1985). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs. *Journal of Personality and Social Psychology*, 48, 82–98.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504–528.

- Hofstee, W., de Raad, B., & Goldberg, L. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, 63, 146–163.
- Hofstee, W., ten Berge, J. M. F., & Hendriks, A. A. J. (1998). How to score questionnaires. *Personality and Individual Differences*, 25, 897–909.
- John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 461–494). New York, NY: Guilford.
- 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: Guilford.
- Loo, R. (2001). A caveat on using single-item versus multiple-item scales. *Journal of Managerial Psychology*, 17, 68–75.
- Lynn, R., & Martin, T. (1997). Gender differences in extraversion, neuroticism, and psychoticism in 37 nations. *Journal of Social Psychology*, 137, 369–373.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- Nunnally, J., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Ozer, D., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, 57, 401–421.
- Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the Big Five? Plenty! *Journal of Personality*, 68, 821–835.
- Paunonen, S. V., Rothstein, M. G., & Jackson, D. N. (1999). Narrow reasoning about the use of broad personality measures for personnel selection. *Journal of Organizational Behavior*, 20, 389–405.
- Pavot, W., & Diener, E. (1993). The affective and cognitive context of self-reported measures of subjective well-being. *Social Indicators Research*, 28, 1–20.
- Peabody, D. (1987). Selecting representative trait adjectives. *Journal of Personality and Social Psychology*, 52, 59–71.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385–401.
- Rammstedt, B., & John, O. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203–212.
- Roberts, B. W., Bogg, T., Walton, K., Chernyshenko, O., & Stark, S. (2004). A lexical approach to identifying the lower-order structure of conscientiousness. *Journal of Research in Personality*, 38, 164–178.
- Roberts, B. W., Kuncel, N., Shiner, R. N., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives in Psychological Science*, 2, 313–345.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg self-esteem scale. *Personality and Social Psychology Bulletin*, 27, 151–161.
- Saucier, G. (1997). Effects of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology*, 73, 1296–1312.
- Saucier, G. (2003). An alternative multi-language structure for personality attributes. *European Journal of Personality*, 17, 179–205.
- Saucier, G. (2010). *The structure of social effects: Personality as impact on others*. Unpublished manuscript. University of Oregon, Eugene.
- Saucier, G., & Goldberg, L. R. (1998). What is beyond the Big Five? *Journal of Personality*, 66, 495–524.
- Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology*, 76, 613–627.
- Saucier, G., & Skrzypinska, K. (2006). Spiritual but not religious? Evidence for two independent dispositions. *Journal of Personality*, 74, 1257–1292.
- Schimmack, U., Oishi, S., Furr, R. M., & Funder, D. C. (2004). Personality and life satisfaction: A facet-level analysis. *Personality and Social Psychology Bulletin*, 30, 1062–1075.
- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, 94, 168–182.
- Shedler, J., & Westen, D. (2007). The Shedler–Westen assessment procedure (SWAP): Making personality diagnosis clinically meaningful. *Journal of Personality Assessment*, 89, 41–55.
- Soto, C., John, O., Gosling, S., & Potter, J. (2008). The developmental psychometrics of big five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, 94, 718–737.
- Stephenson, W. (1953). *The study of behavior: Q-technique and its methodology*. Chicago: University of Chicago Press.
- Tellegen, A., & Waller, N. G. (1987). *Reexamining basic dimensions of natural language trait descriptors*. Paper presented at the 95th annual meeting of the American Psychological Association, New York, NY.
- Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality*, 38, 319–350.
- Westen, D., & Shedler, J. (1999). Revising and assessing Axis II: I. Developing a clinically and empirically valid assessment method. *American Journal of Psychiatry*, 156, 258–272.
- Westen, D., & Shedler, J. (2007). Personality diagnosis with the Shedler–Westen Assessment Procedure (SWAP): Integrating clinical and statistical measurement and prediction. *Journal of Abnormal Psychology*, 116, 810–822.
- Wood, D., Gosling, S., & Potter, J. (2007). Normality evaluations and their relation to personality traits and well-being. *Journal of Personality and Social Psychology*, 93, 861–879.