Personality and Individual Differences 70 (2014) 51–56

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

# Do personality traits conform to lists or hierarchies?

John C. Loehlin [a,*], Lewis R. Goldberg [b]

[a] University of Texas at Austin, Austin, TX, USA
[b] Oregon Research Institute, Eugene, OR, USA

ARTICLE INFO

*Article history:*
Received 20 April 2014
Received in revised form 8 June 2014
Accepted 11 June 2014
Available online 8 July 2014

*Keywords:*
Personality factors
Bass-ackwards method
Personality scales
Adjective self-ratings
Trait hierarchies

ABSTRACT

Are personality traits mostly related to one another in hierarchical fashion, or as a simple list? Does extracting an additional personality factor in a factor analysis tend to subdivide an existing factor, or does it just add a new one? Goldberg's "bass-ackwards" method was used to address this question, based on rotations of 1–12 factors. Two sets of data were employed: ratings by 320 undergraduates using 435 personality-descriptive adjectives, and 512 Oregon community members' responses to 184 scales from 8 personality inventories. In both, the view was supported that personality trait structure tends not to be strongly hierarchical: allowing an additional dimension usually resulted in a new substantive dimension rather than in the splitting of an old one, and once traits emerged they tended to persist.



## 1. Introduction

How many personality dimensions are there? If one takes all the terms that have been used in natural language to describe personality, the answer clearly is "many thousands." Allport and Odbert (1936), in their list of 17,953 person-descriptive English words from Webster's unabridged dictionary, could serve as a start, or their shorter list of 4504 personality-trait words more strictly defined. But the designers of personality inventories typically opt for assessing a good many fewer dimensions. Eysenck favored three (e.g., Eysenck & Eysenck, 1968), as did Tellegen (1985)—although slightly different; Cattell (1946) preferred sixteen; the Big Five (e.g., Digman, 1990; Goldberg, 1993) have wide current popularity; and there has been recent interest in the Big One (Musek, 2007) and the Big Six (Saucier, 2009).

It should be noted that asking "How many dimensions" presupposes that a dimensional approach to personality, as exemplified by factor analysis, is an appropriate one. It has certainly been a popular one, as evidenced by the theorists mentioned and many others, including such pioneers in the field as Guilford (1936) and Thurstone (1951). There are other legitimate ways of considering personality structure, ranging from hydraulic metaphors to brain systems, including such differing psychometric approaches as the radex (Maraun, 1997) and cluster analysis (Tryon, 1970). Our concern in the present paper, however, is with a dimensional approach.

Our general view is that a theorist or test designer can have as few or many dimensions of personality as he or she elects to measure. But how are choices of differing numbers of dimensions related? Most of the dimensional systems mentioned have been developed from the bottom up, starting with individual words, items, or item clusters. However, one can also take a "bass-ackwards" (Goldberg, 2006) or sequential factors approach to studying the relationships among different numbers of extracted factors. In examples, applications to personality began with a large number of adjectives (1710 in one case, 435 in another), from which 1, 2, 3, 4, 5, etc., factors were extracted and rotated orthogonally by varimax. The results were arranged pictorially as a series of rows of boxes, with arrows representing the correlations between the factor scores of factors at adjacent levels. An advantage of this approach is that it does not require deciding that some fixed number of dimensions describes personality, but, rather, permits comparison of the consequences of dealing with varying numbers. Moreover, the pattern of relationships between levels may help us evaluate different types of relationships among traits.

We may distinguish two extreme forms of trait organization. We will call them the hierarchy and the list. In general, in a hierarchical organization, traits at any given level in the trait hierarchy split into subtraits at the next level down. For example, in a hierarchical view of the Big Five (e.g., Musek, 2007), a general factor of personality at the top level is divided into alpha and beta factors at the next level, and these are in turn split, one into three and one into two factors, to yield the Big Five. These may then be

* Corresponding author. Address: Psychology Department, University of Texas at Austin, 1 University Station A8000, Austin, TX 78712-0187, USA. Tel.: +1 512 475 7008.
*E-mail address:* loehlin@utexas.edu (J.C. Loehlin).

http://dx.doi.org/10.1016/j.paid.2014.06.018
0191-8869/

further subdivided—e.g., for Costa and McCrea (1992), each into six facets.

If personality trait organization is strictly hierarchical, each row of a sequential-factors schema will have one trait from the preceding row split into two parts, not necessarily exactly equal, but both substantial, with the split occurring anywhere along the row. The larger part will retain its place, and the other move to the new slot in the row. And traits will only rarely persist more than a few levels downward in the structure without undergoing a split.

By contrast, in a purely list organization of traits, each step in the analysis adds to the list of traits, and major splits do not occur. A new trait emerges at each level, possibly—but not necessarily—with a few minor links to the preceding level, and the new trait as well as the traits from the preceding levels persist down through successive levels in the diagram.

These two do not exhaust the ways in which traits may be organized: for example, circumplexes (Wiggins, 1979) or other structures may occur. But a distinction between hierarchical and list-type organization for personality traits would seem worth exploring, and the sequential-factors schema represents one way of doing it.

The issue of whether personality traits do or do not form hierarchies has been controversial. The possibility of a hierarchical structure centered on the Big Five was mentioned earlier. But Revelle and Wilt (2013) have used typical levels of trait intercorrelation to compare a hierarchical structure of personality traits with that of cognitive traits, and concluded that the former has little psychometric credibility.

In the present paper, we use a sequential-factors approach to address this question, using two large data sets originally gathered for other purposes. One is the responses of 320 college students to 435 common English adjectives describing personality traits. All of the participants rated the adjectives as describing themselves, and most of them also used them to describe a person they liked of their same age and sex (Goldberg, 1990). The other starting point is 184 scales from 8 standard personality inventories that were completed by some 500 members of the Eugene–Springfield (Oregon) community sample (Grucza & Goldberg, 2007). At issue: Will the results predominantly conform to a list or to a hierarchy scheme? Will they generalize across the two data sets?

## 2. Method

### 2.1. Participants

The sample for the adjective ratings consisted of 320 undergraduates in a psychology class who rated themselves; 316 of them also rated someone of their same age and sex whom they liked (Goldberg, 1990). These 636 sets of ratings provide the first data set analysed.

The participants in the sample providing the second data set, the scale scores, were adult community residents of a wide age range who agreed to complete a number of personality questionnaires by mail over a several-year period for honoraria ranging from $10 to $25. Further details on their characteristics may be found in Grucza and Goldberg (2007). The number of participants for individual inventories ranged from 680 to 857; 514 individuals with relatively complete data over the period were used for the present analysis (details below).

### 2.2. Measures

For the student sample, 7-point rating scales were used. Four midscale response options were provided—*average or neutral*, *it depends on the situation*, *don't know* and *term unclear or ambiguous* (Goldberg, 1990). Originally, 587 adjectives were rated; they were reduced to the present 435 by eliminating less familiar ones (Saucier & Goldberg, 1996).

For the community sample, the lowest-level scales available from each inventory were taken as the starting point—these were variously labeled in the different inventories as subscales, facets, clusters, basic scales, etc.; 184 such scales from 8 inventories were used. Respondents with more than 10% missing scores (which usually meant missing one or more inventories) were eliminated from the sample; the missing scale values from the remaining participants were replaced by mean values for the scale. A number of more sophisticated methods of imputing missing data exist, but when the amount of missing data is small (1.9% at this stage for these data) simpler methods tend to give very similar results (Parent, 2013).

### 2.3. Analyses

The factor analyses involved were carried out as principal component analyses rather than strict factor analyses, for the advantages of computational economy, avoidance of Heywood cases, and the ability to calculate factor scores directly rather than having to estimate them. With large initial matrices, such as the ones involved in this study, the two methods tend to give closely similar results. (Small matrices present an entirely different story—e.g., see Loehlin, 1990). Orthogonal (varimax) rotations were used for the same reasons of simplicity and robustness as the use of principal components. In comparisons (Goldberg, 1990) involving 5 factors and 75-variable adjective-based matrices, factor scores based on five different extraction methods, including principal components, were correlated on average from .950 to .996; and factor scores from oblique and orthogonal rotations were correlated on average from .991 to .995. For the sequential-factor analyses of the present study, inter-level correlations were calculated via factor scores, either directly or via the shortcut calculation described by Waller (2007).

For practical reasons of display, the analyses in this paper will be presented only as far as 12 factors. This, however, should be adequate. A preliminary analysis using the Cudeck–Browne criterion (Cudeck & Browne, 1983) suggested that cross-replicated stability existed for 8 factors for the 435 adjectives, and 11 for the 184 scales. The Cudeck–Browne criterion involves splitting the sample into halves A and B, extracting factors from subsample A, and comparing the correlation matrix implied by them to the correlation matrix calculated directly in subsample B. Such a criterion normally improves as more and more factors are extracted, and then deteriorates as factors start to reflect merely idiosyncratic features of sample A. This procedure can then be carried out in reverse, extracting factors in sample B and testing them against sample A correlations. There is some ambiguity as to whether the criterion should be calculated over the entire matrix, or over its off-diagonal elements only. We have followed the latter procedure, to avoid dominating the criterion by the error in the diagonal. In the present case, the criterion reached a minimum at 11 factors in each direction for the scales, and 8 in each direction for the adjectives.

## 3. Results and discussion

The basic results are presented in Figs. 1 and 2, for adjectives and scales, respectively.

### 3.1. Starting with adjectives

For the adjectives, each factor is represented in Fig. 1 by the three adjectives that have the highest absolute loading on it (if that