EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT
1993, 53

INTER-ITEM CORRELATION FREQUENCY DISTRIBUTION ANALYSIS: A METHOD FOR EVALUATING SCALE DIMENSIONALITY

RALPH L. PIEDMONT
Loyola College

MICHAEL E. HYLAND
Polytechnic South West

Briggs and Cheek (1986) have suggested that the mean of inter-item correlations for a scale provides information about whether that scale is unidimensional or not. More information about the dimensionality of a scale is provided by the frequency distribution function of inter-item correlations. A method of examining the frequency distribution functions as a way of inferring dimensionality is described and illustrated. Using data from the NEO Personality Inventory, the method is employed to distinguish between uni- and multidimensional scales. This method represents a useful first step in data description that can orient a researcher to the kinds of qualities that underlie a scale.

IN their widely cited article, Briggs and Cheek (1986) reviewed the use of factor analysis and other techniques in personality questionnaire construction. One important issue that emerged in this discussion was related to the determination of the dimensionality of a scale. As part of a more general discussion of methods for determining dimensionality, Briggs and Cheek (1986) suggested that the

Request reprints from Ralph L. Piedmont, Department of Pastoral Counseling, Loyola College, 7135 Minstrel Way, Columbia, MD 21045 or from Michael E. Hyland, Department of Psychology, University of Plymouth, Drake Circus, Plymouth, Devon PL4 8AA, United Kingdom.

mean of the inter-item correlations can serve as a guide of item homogeneity in a unidimensional scale:

> We believe that the optimal level of homogeneity occurs when the mean inter-item correlation is in the .2 to .4 range. Lower than .1 and it is unlikely that a single total score could adequately represent the complexity of the items; higher than .5 and the items on a scale tend to be overly redundant and the construct measured too specific. (p. 115)

The purpose of this paper is to extend the use of mean inter-item correlations as a technique for examining homogeneity. Inter-item correlation frequency distribution analysis is based on an examination of the frequency distribution of the inter-item correlations rather than just the mean. This technique provides a descriptive tool that can orient researchers to salient aspects of their scales and that should be carried out as a first step in the data analytic paradigm of scale construction.

Theoretical Background

A matrix of the inter-item correlations from two different types of fictitious "ideal" scales is proposed. An ideal unidimensional scale is one where all the items correlate equally well with one another (e.g., where all the inter-item correlations are exactly .3). By contrast, an ideal two dimensional scale might be one where there are two sets of items such that the inter-item correlations within each set are all exactly .3 and the correlations between items in different sets are exactly zero.

Reality never corresponds to these two forms of the ideal. In reality, the pattern of correlations falls somewhere between these two extremes. In a so-called unidimensional scale there will be a tendency for subsets of items to correlate higher with each other than with other subsets. Even the best unidimensional scale has to some extent a hierarchical organization of items. By contrast, in a so-called multidimensional scale there will always be a tendency for items in one subscale to be correlated with one or more items in another. Even where subscales are orthogonal there will be, albeit low, correlations between items in different subscales.

In writing, modifying, or researching a scale, one finds it necessary at some stage to make a categorical decision about whether a scale is uni- or multidimensional in nature. This decision is based first on theoretical considerations such as whether the construct is conceived to be broad or specific. Then an empirical evaluation of the data provides a pivotal test of the hypotheses proposed.

TABLE 1
Number and Magnitude of Expected Inter-item Correlations for Scales with Varying Dimensionality

Number of Correlations at .3	at 0	Scale Type
66	0	Unidimensional, 12 items
30	36	Two dimensions of 6 items
34	32	Two dimensions, eight and four items
39	27	Two dimensions, nine and three items
18	48	Three dimensions of four items
12	54	Four dimensions of three items
6	60	Six dimensions of 2 items

Note. Number of .3 correlations computed using the formula: $(N_{\text{dim x}} \times (N_{\text{dim x}} - 1))/2$.

In practice, the most common way that data inform decisions of dimensionality is factor analysis—for example, the scree test (Cattell, 1966). Whenever the scree test suggests multiple alternatives, then post-hoc trial analyses are often run to see which number of factors makes the most psychological sense. Because interpretations of the scree test can be open to dispute, as can many other aspects of factor analysis (Comrey, 1978), it is useful to have other methods that can help determine dimensionality. Some of these other methods are described by Briggs and Cheek (1986). Inter-item correlation frequency distribution analysis is a new method which adds to the several ways in which decisions about dimensionality can be made. It provides a form of qualitative description for the degree to which the data set varies between ideal unidimensionality and ideal multidimensionality. The appeal of this methodology is that it rests on simple frequency distributions, thus it should be accessible to a broad range of researchers.

Inter-item Correlation Distributions: The Ideal Case

Inter-item correlation frequency analysis is best illustrated with an example. Let one suppose that a scale of 12 items can take one of several ideal forms. The ideal forms differ in terms of both the number of dimensions and the number of items per dimension. One may suppose further that the true correlations within dimensions are always .3 and that the true correlations between items in different dimensions are always zero. Also it may be assumed that the true correlations (e.g., .3 or 0) are always the same as the observed correlations. It is then possible to count the number of correlations at each value for each of these ideal scales, as is shown in Table 1.

Table 1 illustrates the general principle that the number of zero correlations is a function of the number of dimensions. In the unidimensional case there are no zero correlations. In the multidimensional cases the number of zero correlations increases, with increasing dimensionality creating proportionally larger percentages of null relationships in the overall inter-item distribution.

The equation for calculating the expected number of zero correlations (assuming orthogonal latent dimensions) is:

$$\left[\frac{(N_t \times (N_t - 1))}{2}\right] - \left[\sum_{i=1}^{i=k} \frac{(n_i \times (n_i - 1))}{2}\right] \qquad (1)$$

where

N_t = total number of items in overall scale

$$(N_t = n_1 + n_2 + \ldots + n_k)$$

n_i = number of items in each dimension ($i = 1, 2, \ldots k$)

Equation 1 can be readily manipulated to provide the expected percentage of zero correlations:

$$1 - \frac{\left[\sum_{i=1}^{i=k} \frac{(n_i \times (n_i - 1))}{2}\right]}{\left[\frac{(N_t \times (N_t - 1))}{2}\right]} \qquad (2)$$

For example, with two dimensions of equal length, there should be a mode at zero and a slightly smaller mode at .3. With three dimensions, the number of zeros is almost three times as great as the number of substantive correlations. With four dimensions the number of zero correlations are almost five times as frequent; six dimensions predicts 10 times as many zero correlations as .3's.

However, in the real world one does not readily find inter-item correlations all of .3, or inter-factor correlations of zero. Usually, the observed inter-item frequency distribution will be a function of several somewhat independent processes including conceptual redundancy in the items, overlap in response distributions, and sampling variability, that all coalesce to produce a spread of inter-item correlations. What is at the heart of this report is the contention that both the shape and form of this distribution can provide unambiguous insights into the underlying factor structure of a scale.

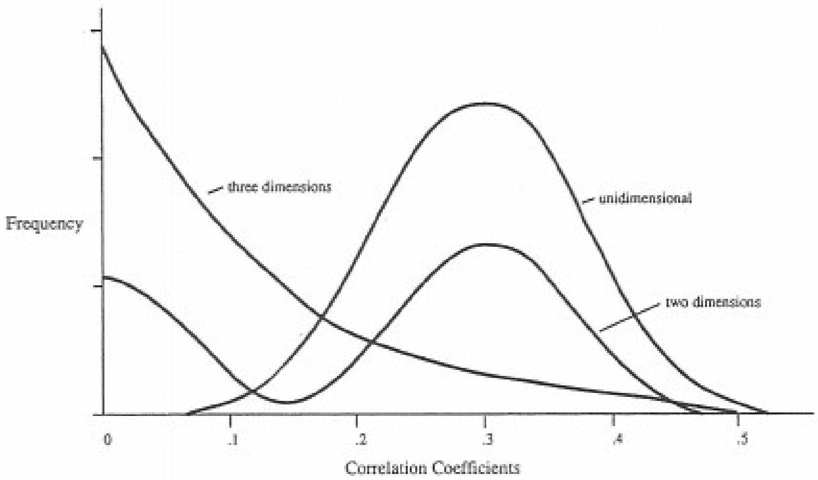

Figure 1. Theoretical distribution functions for the one-, two-, and three-dimensional scales.

Predicted Frequency Distributions

In plotting the inter-item correlation frequency distribution function of a scale, the signs of the correlations are ignored (i.e., all negatives are converted to positives) because interest is solely on the magnitude of the inter-item correlations. Restricting oneself to scales which are real but correspond clearly to either the uni- or multidimensional types, one finds that the frequency distribution function will then have one of several forms. In the case of a unidimensional scale, as noted earlier, inter-item correlations should be approximately normally distributed with a single mode falling between .2 and .4. As the number of dimensions increases, the observed inter-item correlation distribution will begin to become more positively skewed. In the case of two underlying dimensions of equal length, the resulting distribution will evidence a bimodal form, with one mode at zero and another, slightly smaller mode, between .2 and .4. With three or more dimensions, the form of the distribution becomes much more hyperbolic; it has one distinct mode at zero and then quickly tapers off. Predicted distribution functions for one-, two-, and three-dimensional scales are shown in Figure 1.

The curves in Figure 1 are hypothetical. In the current study inter-item correlation frequency distributions were investigated by using real data obtained from the NEO Personality Inventory (NEO-PI) (Costa and McCrae, 1985), a measure derived from rational and factor analytic procedures to assess the five major orthogonal dimensions of personality (Digman, 1990). Creating scales with varying factor structures would provide an opportunity both to validate empirically the previously stated predictions and to create distribution templates for evaluating the dimensionality of other scales.

Method

Subjects

Subjects consisted of 341 undergraduate students who completed the NEO-PI as part of a larger measurement study. Details are presented in Piedmont, McCrae, and Costa (1992).

Measures and Procedures

The NEO-PI (Costa and McCrae, 1985) is a 181-item questionnaire developed through rational and factor analytic methods to measure the five major domains of personality: Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C). These domains were developed to be orthogonal. Items are answered on a 5-point scale ranging from *strongly disagree* to *strongly agree*, and scale items are balanced to control for acquiescence. Internal-consistency estimates for the five domain scales ranged from .76 to .93. The domains of N, E, and O each contain six facet scales (each consisting of 8 items) designed more specifically to assess each dimension. For example, the Neuroticism domain contains the facets of Anxiety, Hostility, Depression, Self-consciousness, Impulsiveness, and Vulnerability. These facets specify in detail the kinds of dispositions that comprise this domain. Internal-consistency coefficients for all NEO-PI facets ranged from .64 to .85.

These facets were selected to serve as the basis for this study because they sample orthogonal dimensions. Using an empirical validation strategy called validimax, McCrae and Costa (1989) developed the NEO-PI so that facets within each domain showed maximal correlations, whereas interdomain facet correlations were simultaneously minimized. Thus, the NEO-PI is an ideal instrument for the purposes of this study because it provides a number of very homogeneous scales from independent dimensions.

Multidimensional prototype scales were constructed by combining facets from different domains to create scales with minimum interdimensional correlations. Intercorrelations between facets combined from different dimensions in this sample were no greater than |.30|, with half being at or below |.12|. The slight associations are most likely the product of some type of evaluative bias in subjects' responses rather than a reflection of any kind of substantive overlap (Costa and McCrae, in press).

To obtain a frequency distribution for the one-dimensional case, separately for each of the 18 facets, inter-item correlations were calculated, and a frequency distribution of the absolute values of these correlations was plotted. The results of these 18 analyses were than aggregated in an attempt to provide a more nearly stable estimate of a one-dimensional distribution. For the two-dimensional case, 18 two-dimensional prototype scales were created by combining two facet scales from different NEO-PI domains. These prototypes were formed in the following manner: N1 and EII, N2 and E2, . . . , N6 and E6; E1 and O1, . . . , E6 and O6; N1 and O1, . . . , N6 and O6. Separately for each prototype scale, inter-item correlations were calculated, and a frequency distribution of the absolute values of the correlations was plotted. Again the results were aggregated. Finally, six three-dimensional prototype scales were created by combining facet scales from the three different NEO-PI domains in the following manner: N1-E1-O1, N2-E2-O2, . . . , N6-E6-O6. Data from these prototypes were aggregated in a similar manner.

Results

Figure 2 presents the overlaid distributions for the one-, two-, and three-dimensional prototype scales. As can be seen, the one-dimensional prototype produces a unimodal distribution consistent with predictions. Also, as anticipated, this distribution is qualitatively different from the distributions generated by the two- and three-dimensional prototypes. The two-dimensional curve does not appear to be clearly bimodal. There is the expected mode at zero, with the distribution gradually tapering off. There then appears another aggregation of correlations around the .14 to .22 interval that creates a slight "step" in the curve. The three-dimensional distribution is clearly hyperbolic in form, as predicted.

The two- and three-dimensional distributions are also statistically different from each other. The percentage of zero correlations found in the two-dimension case is significantly less than the percentage of such correlations in the three-dimension case $t(22) = 2.25, p < .05$. This result indicates that the two distributions have different *y*

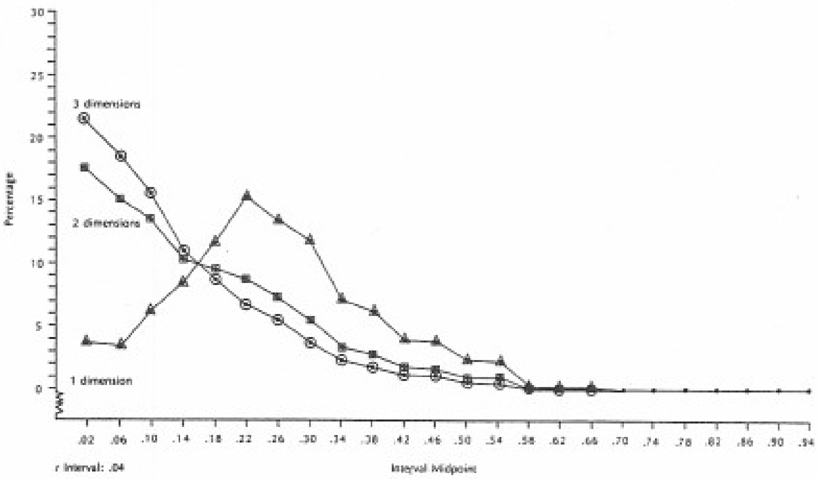

Figure 2. Inter-item correlation frequency distributions for the one-, two-, and three-dimensional prototype scales.

intercepts. Furthermore, a test for the difference of slope of the two distributions between the .00 and .36 intervals (relatively linear aspects of the curves) was also statistically significant $t(18) = 4.48$, $p < .001$—an outcome indicating that the three-dimensional scale has a steeper descent.

Discussion: The Role of Distributional Analyses

In any research endeavor it is important for an investigator to gain an intimate understanding of his or her data. Such a familiarity with the data usually begins with simple descriptions such as means, standard deviations, ranges, and scatter plots. In each of these instances the data, which are in their original form, can provide a researcher with a sense of their suitability for various statistical analyses. Particularly in the area of scale construction, the form of the original test data has important implications for the kinds of conclusions that can be made about the psychometric qualities of an instrument. For example, endorsement frequencies and the range of responses directly impact the integrity of the correlational procedures that underlie most reliability and validity analyses.

The distributional analysis technique outlined in this paper represents another tool for researchers to use in familiarizing themselves with their data. Representing a "first step" in a thoughtful analytic methodology, it has the potential to alert an investigator to issues that may arise in the research process. In a very simple and straightforward manner, a distributional analysis can provide evidence regarding the dimensionality of a scale. In a multidimensional case, the number of observed zero correlations and the form of the distribution are suggestive of the number of factors that underlie the data. These decisions are not based on the kinds of subjective criteria associated with other, more sophisticated statistical procedures such as factor analysis. Indeed, an examination of the distribution of inter-item correlations for a scale may help an investigator make a more informed choice in deciding on the number of factors to be extracted and rotated.

In the case of a unidimensional scale, the distribution of which approximates a positively skewed normal curve, the magnitude of the average $|r|$ indicates the degree of redundancy in the item content. As Briggs and Cheek (1986) noted, a high average $|r|$ indicates that the construct is narrowly defined and that any given item provides little new information over what is already afforded by the other items. A low average $|r|$ suggests a more heterogeneous item content that may impair internal consistency and predictive specificity. An average $|r|$ between .2 and .4 represents an optimal level of item specificity. The amount of variability around the average $|r|$ may reflect an overinclusion of non-optimal items in the scale, particularly as they affect the positive skewness of the distribution. Tailoring the distribution to resemble an ideal curve would help to eliminate both conceptual redundancy and excess items.

Summary

The inter-item correlation frequency distribution for the real unidimensional case was similar to that predicted. Moreover, this distribution is qualitatively and quantitatively different from those distributions of the multidimensional cases. The distribution for the real two-dimensional case, which approximated the predicted bimodal shape, was quite distinct from the three-dimensional case in terms of both y-intercept and slope. There can be no doubt that a frequency analysis can be extremely clear in discriminating between uni- and multifactorial scales. Portraying inter-item relationships in terms of a distribution provides a simple, clear, and straightforward procedure for dimensional analysis.

REFERENCES

Briggs, S. R. and Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54, 106–148.

Cattell, R. B. (1966). The scree test for number of factors. *Multivariate Behavioral Research*, 1, 245–276.

Comrey, A. L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, 46, 648–659.

Costa, P. T. and McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T. and McCrae, R. R. (in press). The NEO Personality Inventory (NEO-PI). In S. R. Briggs and J. M. Cheek (Eds.). *Personality measures (vol. 1)*. Greenwich, CT: JAI Press.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417–440.

McCrae, R. R. and Costa, P. T. (1989). Rotation to maximize the construct validity of factors in the NEO Personality Inventory. *Multivariate Behavioral Research*, 24, 107–124.

Piedmont, R. L., McCrae, R. R., and Costa, P. T. (1992). An assessment of the EPPS from the perspective of the five-factor model. *Journal of Personality Assessment*, 58, 67–78.