CHAPTER 25

Factor Analysis in Personality Research

Kibeom Lee
Michael C. Ashton

The personality researcher frequently works with a large set of variables that correlate with each other to varying extents. The complexity of such a variable set can create difficulties for data analysis and for conceptual understanding, and one way to reduce this complexity is to use *factor analysis*. Briefly, factor analysis summarizes the relations between many variables by expressing each variable as some unique combination of a few basic dimensions, known as factors. In this way, a group of correlated variables can often be treated as examples of a single, broad factor that is distinct from other factors that summarize other groups of correlated variables. By reducing many variables to a few factors, factor analysis provides a convenient method of simplifying one's variable set for the purpose of examining relations with outside criteria. In addition, factor analysis may stimulate insights into the nature of the variables themselves, by allowing the researcher to identify some common element among variables belonging to the same factor.

In this chapter, we describe the use of factor analysis in personality research and related contexts. First, we begin with a very brief and nontechnical explanation of the mathematical basis of factor analysis. Then we proceed to a discussion of the many decisions to be made by the researcher when using this technique, with special attention to some of the more complex issues that frequently arise. Next, we show an example of factor analysis using real data from a published personality study; readers may prefer to examine this example before reading the more abstract description given immediately below. Finally, we add some closing remarks about the use of this technique.

An Overview of Factor Analysis

Factor analysis attempts to reduce many correlated variables to a few broader dimensions (i.e., factors) that summarize the correlations between those variables.[1] The process of factor analysis generally begins with the calculation of a matrix of correlations among the variables that have been assessed in one's participant sample. This correlation matrix allows one to find the linear combination of variables that will produce the first and largest factor. (See Table 25.1 for an example correlation matrix, derived from a dataset introduced toward the end of this chapter.) Specifically, participants' scores on the first factor can be derived by finding the linear combination of standardized scores (i.e., *z*-scores) on the variables that has the maximum possible variance. The variance of the factor obtained in this way is known as the *eigenvalue*, and this value thus represents the size of the factor. Because the variance of any linear combination is a function of the correlations among the variables involved, the more strongly correlated the variables are, the larger the first eigenvalue will be. Consequently, those variables that correlate substantially with many other variables in the dataset are likely to be the variables that will be weighted heavily in producing the first factor. (Note that negatively correlated variables will have opposite-signed weights in the equation that produces the first factor.)

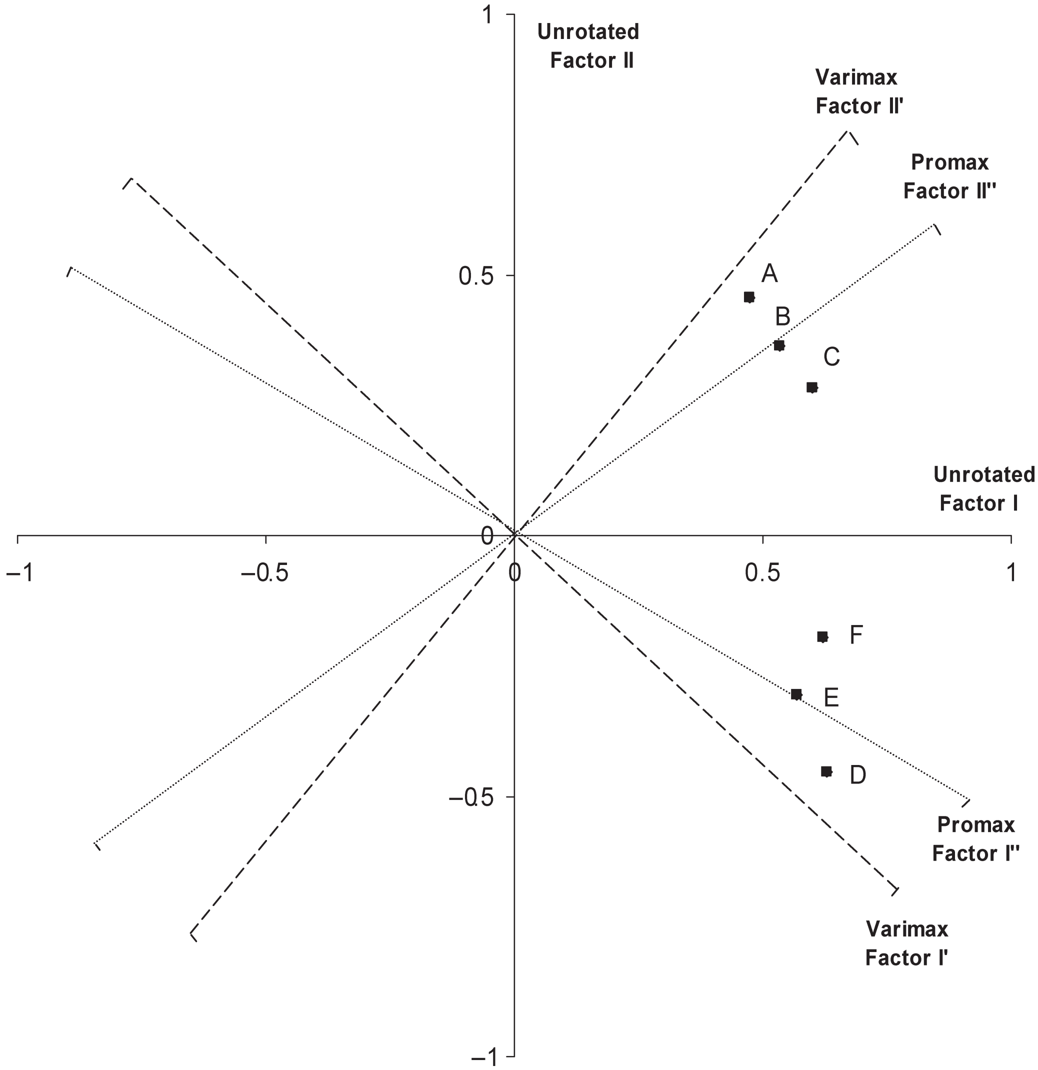
The second factor is derived in precisely the same way, except that the variance of the first factor has first been completely removed from all variables in the dataset, so that the correlation matrix being factor analyzed is actually a matrix of residual correlations after the removal of that variance. As a consequence, the second factor is uncorrelated with the first. This process of extracting additional uncorrelated factors can continue until all of the variance of every variable in the analysis is completely exhausted.

In the initial factoring stage, therefore, the variances of *k* variables are completely redistributed to produce as many as *k* orthogonal factors, each of which represents the largest possible factor after the preceding ones, and each of which is successively smaller. These characteristics are very useful in making a decision as to how many dimensions are needed to best represent the data (i.e., how many factors to extract), because the first *m* factors always provide the best summary of the variable set that could be achieved by any possible set of *m* dimensions. As we discuss later in this chapter, the task of determining the number of factors is one of the most important and difficult in factor analysis, and most decision rules involve in some way the use of eigenvalues obtained in the initial factor extraction stage.

If we extract a certain number of factors (*m*), these factors can be described in terms of their correlations with the observed variables. The correlations between variables and factors are called *factor loadings*, and a matrix showing these values is called a *factor loading matrix* (see Table 25.2 in the example at the end of this chapter).[2] A useful way to imagine the factors is to draw them as vectors of unit length drawn at right angles to all other vectors; in this way, a variable's loading on a factor is its projection on a vector, and any given variable can be located by a set of coordinates (i.e., its factor loadings) in the space that is spanned by those factors or dimensions. (See a hypothetical example in Figure 25.1, which shows the locations of six variables within the space of two dimensions, the unrotated factors I and II.)

Note that the variance of each factor (i.e., each *eigenvalue*) can be obtained from a factor loading matrix by finding the column-wise sums of squared factor loadings. For example, in the dataset introduced later in this chapter (see Table 25.2), the eigenvalue of the first unrotated factor (i.e., 4.48) can be obtained by finding $(-.83)^2 + .72^2 + .63^2 + \ldots + (-.26)^2 + .09^2 + .50^2$. On the other hand, the row-wise sums of squared factor loadings are called *communalities*, each of which corresponds to the proportion of variance in a variable that is accounted for by the retained factors. For example, in Table 25.2, the communality of Fun Seeking (.58) can be obtained by finding $.72^2 + .20^2 + .16^2$. If the communality of a given variable is particularly small, this suggests that the variable is poorly accommodated within the space spanned by the retained factors (i.e., dimensions).

Another important feature of the factor loading matrix is that a correlation between any two variables can be estimated by finding the sum of cross-products of factor loadings of the two variables on the same factors. For example, in Table 25.2, the correlation between Neuroticism and Psychoticism can be estimated by calculating the expression $(-.53)(.50) + (.59)(-.01) + (.44)(.55)$. All of the other correlations among variables can be estimated in an analogous way, and the resulting correlation matrix (i.e., often called a *reproduced correlation matrix*) is sometimes compared against the *observed correlation matrix*. Specifically, the reproduced correlation matrix can be sub-



Hypothetical Factor Pattern Matrices

Variables	Unrotated		Varimax Rotated		Promax Rotated ($r_{I''II''}$ = .43)	
	I	II	I′	II′	I″	II″
A	.47	.46	.05	.66	–.11	.70
B	.53	.36	.16	.63	.02	.64
C	.60	.28	.26	.61	.13	.60
D	.63	–.45	.77	.08	.81	–.10
E	.57	–.31	.63	.15	.64	.01
F	.62	–.20	.60	.26	.58	.14

FIGURE 25.1. A graphical illustration of factor rotation.

tracted from the observed correlation matrix to produce a *residual correlation matrix*. If all of the elements in the residual matrix are fairly small, one can conclude that the obtained factor structure provides a reasonably good approximation to the data.

The factor loading matrix obtained in the initial factoring process usually produces factors that each show substantial loadings for many variables, rather than very strong loadings for only a few variables each. As a result, inspection of this matrix usually does not allow simple interpretations of the nature of the factors. However, this result is not unexpected, given that the factors obtained in this stage are linear combinations derived exclusively by a criterion of variance maximization, with no attempt made to simplify their meaning.[3] To improve the interpretability of components, the axes of factors can be *rotated* in such a way that each factor will have high loadings for a few variables, and low loadings for the (many) others. (See again the hypothetical results shown in Figure 25.1, where the original factor axes, I and II, have been rotated to new positions, I′ and II′ and also I″ and II.″) Many methods of factor rotation are available, some of which are introduced later in this chapter. As a result of rotation, a new factor loading table is created, and on the basis of this table (see Table 25.3, as derived from the personality dataset described later in the chapter), the rotated factors are interpreted and labeled in such a way as to summarize the common elements of the variables that define each factor. As with eigenvalues, the variances of the *rotated* factors can also be obtained by summing the squared loadings within each factor. These figures are usually labeled *sums of squared loadings (SSLs)* to distinguish them from eigenvalues.

It is important to note that *orthogonally* rotated solutions are mathematically equivalent to each other. This means that the reproduced correlation matrix among variables estimated from such solutions is invariant, that the communality of a variable is invariant, and that the total of the SSLs within a given space is invariant, regardless of the rotational positions of the orthogonal axes. Rotation involves, however, the redistribution of variances in the retained factors, and hence the SSL for each of the factors taken individually will change, depending on the locations of factor axes.

Finally, after the rotated factors are interpreted (and labeled), scores on the factors can be computed for each participant for use in subsequent analyses. A widely used method of obtaining such scores is to compute a mean score across the variables that load strongly on each factor. Other, more sophisticated methods generate a set of factor scoring coefficients that are applied to individuals' scores on the original variables to produce factor scores. The most widely used such method is called the regression approach. In principal components analysis, factor scores obtained from the regression approach have a mean of 0 and a standard deviation of 1; when the factors are orthogonal, these regression-based factor scores will also be orthogonal.

Decision Steps and Complexities

Variable Set

Factor analysis is sometimes conducted simply for the purpose of data reduction, in the sense of summarizing scores on many variables in terms of scores on only a few factors. For example, if a factor defined by several highly intercorrelated personality variables were found to be correlated with some outside criterion, it would be simpler to report this single finding than to report individually the relations of each of those personality variables in turn with the criterion.[4]

Beyond serving the purpose of data reduction, factor analysis can also play a crucial role in identifying a set of basic dimensions that underlie the domain of personality itself. In other words, factor analysis may be used in the search for a few broad dimensions of personality that in combination will summarize the relations among the full array of personality characteristics. But when factor analysis is used for this purpose, the composition of the variable set is of crucial importance: If some aspects of the personality domain are under- or overrepresented, then one's factor solution is likely to omit some important factors or to include extra factors defined by trivially redundant variables. Because personality inventories assess relatively small numbers of variables selected according to the preferences of the test developer, it is highly doubtful that the variables of any inventory (or even any combination of inventories) will provide a representative sampling of the personality domain. For this reason, even the repeated recovery of a given factor solution from a personality inventory does not constitute evidence regarding the structure of personality variation more generally (cf. McCrae & Costa, 1997). It may sometimes be of interest to examine the factors underlying the scales of a given inventory (e.g., Ashton, Jackson, Helmes, & Paunonen, 1998), but again, the obtained factors are chiefly of local interest.

To find the set of major dimensions that summarize the domain of personality variation, it is necessary to analyze a set of variables that comprehensively represents the full personality domain. The only strategy that has thus far been proposed for this purpose is based on the Lexical Hypothesis (e.g., Goldberg, 1981), which is the idea that personality characteristics having importance in person description tend to become encoded in languages as single words (typically as adjectives, or as their corresponding attribute nouns). Following the logic of the Lexical Hypothesis, a complete list of the common personality-descriptive adjectives of a language would approximate a representative sampling of the domain of personality characteristics. Therefore, a factor analysis of ratings on those adjectives should produce a structure corresponding closely to the structure of personality variation itself. Interestingly, there has been some consistency in the factor structures obtained thus far from lexical studies of personality structure conducted in several diverse languages (see, e.g., Ashton, Lee, & Goldberg, 2004; Ashton, Lee, Perugini, et al., 2004; Ashton & Lee, in press).[5]

Between the extremes of simple data reduction and of identifying the structure of the personality domain, the use of factor analysis can also be applied to the exploration of some specified region of the personality domain. For example, a researcher may want to identify and to interpret a few factors that will provide a useful conceptual summary, say, of traits associated with social interactions, or with gender differences, or with cognitive styles. One such case is described later in this chapter, where an examination by Zelenski and Larsen (1999) of affect- and impulse-related traits is reviewed.

Participant Sample

When conducting psychological research, one ought to select a participant sample for which the obtained results will generalize to the population of interest. Ideally, one should have a participant sample that is not only very large but also very similar to the target population in terms of demographic variables that may conceivably influence the factor structure of the variable set in question. But it is also important to keep in mind that the factor structure of a given domain may differ, depending on how broadly one defines the population of interest. For example, large sex differences on some variables may cause those variables to correlate more strongly within a mixed-sex sample than within a sample of men only or of women only. As a result, these sex-correlated variables may define a factor of their own within the mixed-sex sample (or may attract a factor axis upon simple-structure rotation), even though those variables may instead define several different factors within a single-sex sample. Similarly, when participants are drawn from more than one language, cultural, or racial group, or when there are wide variations in the ages or in the socioeconomic statuses of participants, factor structures may also differ from those derived from samples that are homogeneous with regard to those variables.

It is also possible, of course, that the effects of demographic variables in the context of a given variable set will be trivially small. Nevertheless, the researcher should be aware of the potential for sample heterogeneity to influence factor structures. In addition, the researcher should also consider which result is of greater theoretical interest: the solutions derived from samples that are homogeneous on the demographic variable in question, or the solutions derived from heterogeneous samples. The answer to this question will likely depend on how the researcher views the demographic variable in terms of its influence on the variable set to be factor analyzed. If that demographic variable is viewed as a source of variation that is qualitatively different from (and perhaps quantitatively much stronger than) the sources of variation within demographic groups, then the results based on homogeneous samples will likely be of greater interest. But if the demographic variable is instead viewed simply as a contributor of some additional variation on certain personality variables—variation of the same kind as that occurring within demographic groups—then the results based on heterogeneous samples will likely be of greater interest.

Scaling and Distributions of Variables

When the variables to be factor analyzed are individual items rather than multi-item scales, the nature of the item response format influences the distribution of responses to the items and, as a result, may influence the magnitude of the observed relations between those items. For example, when the item response format is dichotomous (e.g., true/false), the maximum possible correlation between items having different response distributions is sharply limited (e.g., an upper bound of $r = .25$ between items having 80% true versus 20% true responses). This problem is reduced somewhat when items use a multipoint response format, such as a 5-, 7-, or 9-point scale. For this reason, we recommend the use of multipoint, rather than dichotomous, item response formats, even though the latter can still be used for factor-analytic purposes. The precise number of response options is not especially important, but participants may prefer an odd number, which allows for exactly neutral responses to some items. An item format having more response options (e.g., 9 points as opposed to 5) will allow for slightly greater precision in responses, but is also slightly more difficult for participants to use.

Depending on the nature of the variable set to be analyzed, researchers may consider the use of *ipsatized* scores (i.e., scores that have been standardized *within subjects*) instead of raw scores when calculating the correlation matrix. The purpose of using ipsatized scores is to prevent the potential distortion of factor analytic results that may result from individual differences in the overall elevation or extremity of responses to items. For example, because participants differ in their average level of endorsement of items, *independently of the content of the items*, variables will tend to correlate positively with each other. As a result, variables that are inherently opposite in nature may be roughly uncorrelated, thus producing two unipolar factors (e.g., *A* and *not A*) rather than one bipolar factor (*A versus not A*). The use of ipsatized scores eliminates this problem, but an important caution should be noted: If one's variable set tends to represent only one pole of each potential factor—for example, if one has many more items for *A* than for *not A*, and for *B* than for *not B*—then ipsatization will tend to remove variance associated with the content areas themselves, in addition to (or instead of) variance due to elevation and extremity of responses.

The decisions discussed above regarding item response options and the ipsatization of variables tend to be somewhat less important when the variables to be analyzed are multi-item scales rather than individual items.[6] First, concerns about the distributions of items are usually less serious, because the averaging or summing of responses across several items tends to produce distributions that approach normality. Second, if the multi-item *scales* each consist of several items, with each scale roughly balanced in terms of positive- and negative-keyed items, then ipsatization of those scales is unnecessary and is likely even to distort the results, unless the scales themselves are scored in such a way as to represent high and low levels of broad underlying factors about evenly.

Inspection of Data

Before conducting a factor analysis, it is important to ensure that one's data have been properly entered and coded; for example, one should check that all participants' responses to an item fall within the possible range of values. Data should also be inspected with a view to identifying the extent to which there are missing data for each participant and each item. In general, when a participant fails to respond to a substantial fraction of the items, this suggests that the data for this participant may be less than meaningful (perhaps due to inattention or to noncompliance) and ought therefore to be removed. Similarly, if a large fraction of participants fail to respond to a given item, this suggests that the data for this item may be less than meaningful (perhaps due to ambiguity or to obscurity) and ought therefore to be removed. When a participant misses only a small fraction of items, or when an item has missing data for only a small fraction of participants, methods such as mean substitution, listwise deletion, or pairwise deletion are usually adequate ways of handling the problem of missing data.

We should briefly note the problem of statistical outliers: When a participant's responses are vastly different from those of other participants—say, being 4 standard deviations (*SDs) from the mean—this may suggest that the participant's response is not due to the same sources of variation as those that underlie the distribution of the variable in general. As a re-*

sult, the researcher may decide to discard the outlier participant's data. Outliers may also be removed on the basis of bivariate relations, as, for example, when a participant has very high levels of two variables that are otherwise strongly negatively correlated. However, we recommend a very conservative approach to the removal of outlier participants, especially bivariate outliers; if a too liberal threshold is used, the researcher may distort the relations between variables by discarding participants who truly do have relatively unusual combinations of levels of the variables in question.

Stability of Results: Sample Size, Variable Intercorrelations, and Variables per Factor

Factor analyses are usually computed from a matrix of correlations between variables, and the extent to which obtained correlations will fluctuate across samples varies with the square root of sample size. Therefore, larger sample sizes will produce more stable factor analytic results, in terms of the loadings of variables on factors extracted from a correlation matrix.

In addition to sample size, there are other influences on the stability of factor analytic results. Because the sampling fluctuation of a correlation coefficient is smaller when the population value of the correlation coefficient is very large, factor loadings are more stable when the variables defining each factor tend to be strongly intercorrelated (and hence highly loaded on their respective factors). Moreover, because a variable will show less sampling fluctuation in its average correlation with many other variables than in its average correlation with few other variables, factor analytic results are more stable when each factor is defined by many variables. The above features have been demonstrated empirically by Guadagnoli and Velicer (1988), who showed in a series of simulation studies that the stability of factor analytic results depended on (1) the absolute sample size, (2) the loadings of variables on factors in the population, and to a lesser extent, (3) the number of variables per factor.

Many sources suggest that a certain minimum ratio of sample size to number of variables (e.g., 5 to 1, 10 to 1, etc.) is needed to ensure the stability of factor analytic results. These suggestions are entirely misguided. Having a larger number of variables in a factor analysis does *not* undermine the stability of factor analytic results, and neither does the ratio of sample size to number of variables influence the stability of results. Because the sampling error of a correlation coefficient depends on the sample size rather than on the number of other correlation coefficients being calculated, the factor analysis of a large variable set does not require an especially large sample size (see Guadagnoli & Velicer, 1988, for an empirical demonstration).

Component Model versus Factor Model

One decision to be made in conducting a factor analysis is whether to use a *component* model (Hotelling, 1933) or a *factor* model (Pearson, 1901; Spearman, 1904) of factor extraction. For practical purposes, the chief difference between these models is that the component model uses values of 1 in the diagonal elements of the correlation matrix that is factor analyzed, whereas the factor model uses smaller values, typically representing the variance shared between each variable and the rest of the variable set (as indexed, for example, by the squared multiple correlation of each variable with the remaining variables).

When there are many variables to be factor analyzed, the differences between the results yielded by these two methods tend to be small. This is especially so when there is a large ratio of variables to factors, and when the correlations between the variables tend to be large. In such a situation, it does not matter which model is used as a basis for extracting factors.

In the opposing situation, in which there are few variables (and, especially, few variables per factor) and relatively weak correlations between them, then the results obtained from the two methods do tend to diverge, insofar as the component model tends to produce larger loadings for the variables. On one hand, this may be viewed as a disadvantage of the component model, insofar as the loadings obtained in the component model are inflated by variance that is unique to each variable (including error variance), whereas the loadings obtained in the factor model represent more accurately the amount of variance common to the variables. On the other hand, one may view the "inflated" loadings of the component model as advantageous in the sense that they draw one's attention to potential factors that may be underrepresented in a given variable set, and particularly to the variables that are most isolated in the sense of having few strong correlates within that variable set. That is, the factor model provides a faithful representation of the variance shared among the variables, whereas the component model provides a faithful representation of the total variance of the variable set.

How Many Factors to Extract?

The question of how many factors to extract involves a tradeoff between parsimony and completeness in summarizing the relations between the variables. Ideally, one would hope to account for the relations between the variables with only one factor, but then one would need to consider whether the extraction of a second factor would provide a substantially more accurate summary of the relations between the variables. If so, then the same question would be asked regarding the extraction of a third factor, and so on until the extraction of an additional factor would provide such a small increment in accuracy of one's summary that it would not justify the loss of parsimony.

This logic is embodied in the use of several rules for deciding the number of factors to extract. One such rule is simply to extract only those factors whose eigenvalues exceed a certain minimum size. Kaiser (1960) suggested that only factors with eigenvalues of at least 1 should be extracted, because further factors would account for less variance than that of any one variable, and therefore would not help in reducing the variable set to a smaller number of dimensions. However, a fundamental flaw of the eigenvalue-1 criterion is that the number of factors extracted will depend heavily on the number of variables in the analysis, even though the number of variables representing a domain does not change the dimensionality of that domain. More generally, any criterion based solely on eigenvalue size is also highly problematic insofar as one will often extract some factors whose eigenvalues exceed only very slightly those of other factors that are not extracted.

A more defensible strategy for deciding on the number of factors to extract is to consider the *differences* between eigenvalue sizes for successive factors. To use this method, known as the "scree test" (Cattell, 1966), one begins by inspecting the sizes of the eigenvalues of all k possible factors in an analysis of k variables, or by inspecting a graph of those eigenvalues (i.e., a scree plot). One then looks, moving from the kth factor to the k-1th, to the k-2th, and so forth, at the increases in the sizes of the eigenvalues. In most datasets, these increases are very small until one is within the ten (or fewer) largest factors. But when there appears the first noteworthy increase or "jump" in the sizes between the eigenvalues of two successive factors, say, the fourth and the third (see the example in Figure 25.2), this suggests that (in this case) three factors should be extracted. According to the logic of the scree test, one would be justified in ceasing the extraction of factors at three, because the extraction of the third factor provides a notably greater increase in accuracy than does the extraction of the fourth factor. In contrast, if one were to extract a fourth factor, there would be little reason not to extract a fifth factor also, or a sixth, and so on.

Although we agree with the reasoning behind the scree test, we should point out that it is sometimes difficult to judge exactly where the first real "jump" in eigenvalues does occur. (For example, even in the example in Figure 25.2, a case may be made for the extraction of four factors, based on the small jump between the fifth and fourth factors.) Sometimes the scree test suggests two or more alternative numbers of factors to extract; for example, a fairly small jump between the eighth and seventh eigenvalues and a much larger jump between the fifth and fourth would suggest the extraction of either four or seven factors. But rather than being a shortcoming of the scree test, this is likely to be a reflection of reality, as a given variable set may be meaningfully summarized, for example, by a set of four larger factors or by a set of seven smaller factors.

Alternative methods for deciding on the number of factors to extract include the parallel test (Horn, 1965), the minimum average partial method (Velicer, 1976), and the chi-square test (Bartlett, 1950, 1951). The parallel test involves the comparison of eigenvalues obtained from the observed correlation matrix with the average eigenvalues obtained from correlation matrices generated at random from a population in which all correlations equal zero and in which the numbers of variables and participants are the same as those of the observed data. One then retains only those factors that have eigenvalues exceeding those of their corresponding random eigenvalues. The minimum average partial method (Velicer, 1976) is

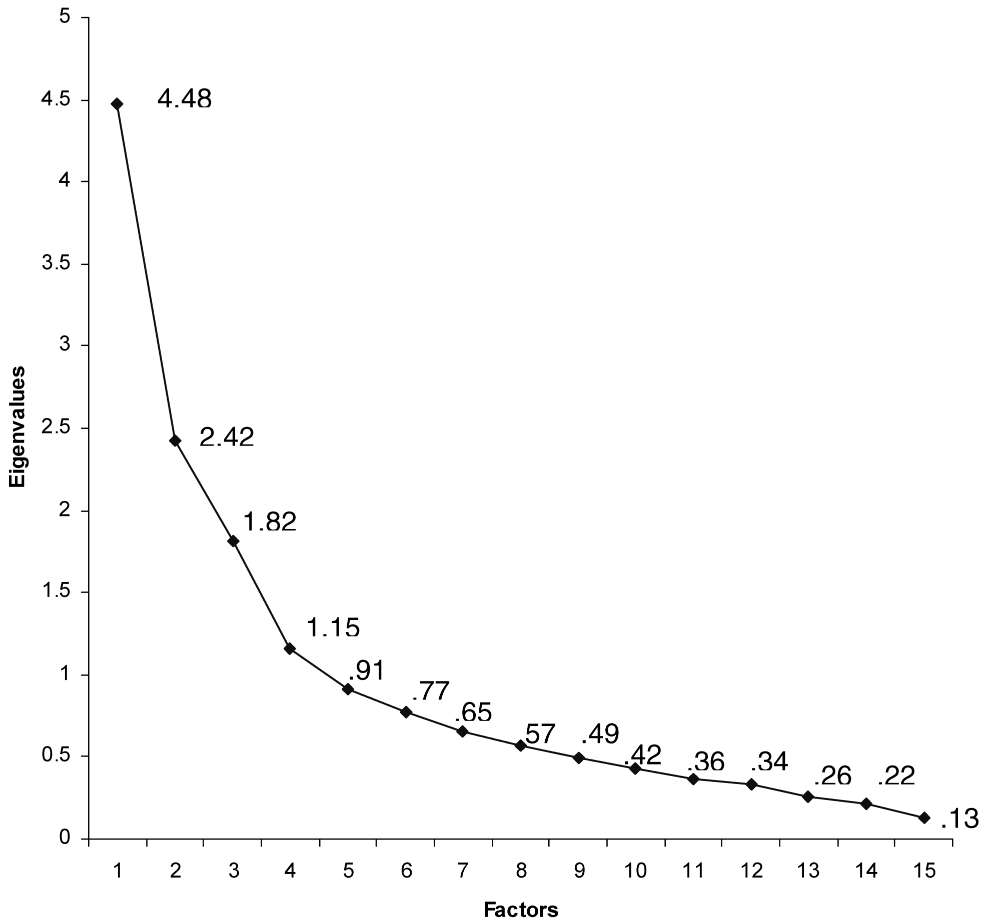


FIGURE 25.2. Scree plots of eigenvalues of factors from the Zelenski and Larsen (1999) dataset.

based on the changes in partial correlations between variables after successive factors are extracted. In this method, components are successively partialled out from the original correlation matrix, and for each resulting matrix of partial correlations, the average of the squared partial correlations is calculated. The optimal number of factors is reached at the point when this average value reaches a minimum.[7]

Finally, a chi-square significance test can be used to determine the number of factors if one adopts the maximum likelihood (or generalized least squares) factor extraction method (cf. the Bartlett test for principal components analysis). In this method, a goodness-of-fit test (i.e., a chi-square test) for a given number of factors is provided by comparing the reproduced correlations between the variables (as derived from the obtained factors) with the observed correlations. Beginning with a one-factor solution, the researcher continues to extract factors until a statistical test first produces a nonsignificant result, because this indicates that the reproduced and observed correlations are not statistically significant. This method, however, has been known to have two crucial problems. First, it requires the assumption of multivariate normality. Second, it is sensitive to sample size, because of its reliance on statistical significance testing. For example, when the sample size is very large, the first solution to produce a nonsignificant result may include a large number of factors, and such a solution would be unparsimonious and unlikely to be replicated in other participant samples or other variable sets.

Once the number of factors to be extracted is determined, it is informative to report the percentage of the total variance that is accounted for by the retained factors. However, we do not suggest any wholesale guideline regarding the minimum proportion of total variance that should be accounted for by a factor solution, because this proportion depends on the reliabilities of the variables being analyzed and on whether the variables have been selected as "markers" that strongly define a given set of factors. For example, the largest replicated factor solutions found in lexical studies of personality structure, based on relatively unreliable ratings on single adjectives representing the entire personality domain, typically account for 20–30% of total variance. In contrast, the largest replicated factor solutions in investigations of omnibus personality inventories, based on relatively reliable scores on multi-item scales developed specifically as factor markers, typically account for 50–75% of total variance.

Factor Replicability and the Number of Factors to Extract

The aforementioned methods of deciding on the number of factors to extract are all based on attempts to identify a point at which further extraction of factors would provide little improvement in the summary of the relations between the variables. An alternative approach is to identify the largest number of factors that can be replicated across subsamples of participants (see Everett, 1983, for a widely used test of factor replicability). We recommend against the use of such methods as applied to *random split-halves* of a given sample, for the simple reason that the replicability of any factor structure derived from a given variable set is a function of sample size: With a small enough sample, any structure will sometimes fail to be recovered from random split-halves of that sample; conversely, with a large enough sample, any structure will be consistently recovered from random split-halves of that sample.

However, we do recommend the use of tests of factor replicability for the comparison of factor structures obtained from participant samples that differ in some substantive way, such as in their demographic composition or the source of the data. For example, if one wishes to compare factor structures from men and from women, or from Australians and from Americans, or from self-ratings and from peer ratings, then methods such as Everett's test will allow one to determine whether or not a given structure replicates across different types of samples. When a given factor structure is found to replicate across demographic groups or across rating sources, this provides much more impressive evidence of replicability than does any result based on random split-halves of a given sample. (Moreover, if a given solution replicates across demographic groups or across rating sources, then that solution is obviously a meaningful one, regardless of any failure to replicate that solution across the much smaller subsamples created by random split-half divisions of a given sample.) Note that the use of tests to find the number of replicable factors requires the use of very large samples; otherwise, a factor structure may fail to replicate across demographic groups or across rating sources simply because of sampling error.

As noted above, tests of factor replicability can be useful for deciding on the number of factors to extract from a given variable set, at least when replicability is evaluated across very large samples consisting of *different types of participants*. But if one has the more ambitious aim of discovering the dimensionality of a given domain—not simply the dimensionality of a particular variable set taken from that domain—then there is a much more important issue than that of the replicability of a solution across different samples. The important issue, instead, is whether a solution is replicable *across different variable sets* that are each selected to be roughly representative of the domain of interest. If, for example, a given factor structure of personality characteristics is found to be very similar across various sets of indigenous personality-descriptive adjectives taken from different languages, then this provides very impressive evidence of the replicability of the solution, regardless of any results derived from comparisons of samples on the basis of a single variable set. When investigating the dimensionality of a given domain, it is important to remember that one is sampling not only from a population of participants, but also from a population of variables.

A few additional remarks are warranted regarding the decision of the number of factors to extract. First, although it is obviously of much interest to find the "ideal" factor solution underlying one's variable set, it is also useful to examine the nature of the dimensions obtained in solutions involving different numbers of factors (Goldberg, 2006). The researcher can proceed systematically by first examining the first unrotated factor, then the rotated two-factor solution, three-factor solution, and so on, until a solution is reached in which some of the factors become too small and too weakly defined to be interpretable. This exercise is likely to provide the researcher with a deeper understanding of the relations among his or her variables and of the plausibility of various solutions; furthermore, by reporting these results at least briefly, the investigator allows future researchers to compare obtained solutions involving a given number of factors with corresponding solutions derived from other variable sets. And finally, if a given factor solution is of particular interest because of its relevance to the researcher's theoretical expectations, then he or she may well decide to examine that solution in detail, regardless of the results of more algorithmic procedures described earlier in this section.

Rotation of Factors and Simple Structure

After extracting a given number of factors, the researcher will want to interpret those factors. However, when more than one factor is extracted, the patterns of factor loadings of the variables are often quite complex, with many factors being defined modestly by many variables, and with many variables showing moderate loadings on many factors. To better understand the factor space that summarizes the relations among a set of variables, it is convenient to reorient the factor loading matrix to a mathematically equivalent position in which each factor is defined strongly by a few variables and not at all by most variables, and in which each variable defines one factor strongly and the other factors not at all. That is, the researcher rotates the obtained factor axes (or vectors) so that the new positions of those vectors will produce a "simple structure" that allows easier interpretation of the factors, by categorizing variables more neatly within the various factors.

The first decision that must be made when rotating factors is whether to use an orthogonal rotation, in which all vectors will remain at right angles to each other after rotation, or an oblique rotation, in which the angles between the vectors will be allowed to deviate from orthogonality. That is, the factors remain mutually uncorrelated in the case of orthogonal rotations, but are permitted to correlate with each other in the case of oblique rotations. The choice of an orthogonal or an oblique rotation depends mainly on the expected pattern of relations among the variables. If one expects that the domain of variables being analyzed contains at least two roughly independent dimensions (e.g., when analyzing a wide array of personality variables), then an orthogonal rotation is preferred, in order to obtain a simple representation of the location of each variable within the space of those dimensions. If one expects instead that the domain of variables being analyzed is dominated by a single large general factor (e.g., when analyzing cognitive ability variables, or personality variables thought to be facets of one personality dimension), then an oblique rotation is preferred, in order to identify clusters of variables that are particularly closely linked and to indicate the extent to which those clusters are intercorrelated.[8]

When an orthogonal rotation is used, the researcher obtains a single set of factor loadings for each variable on each of the uncorrelated factors. When an oblique rotation is used, there are two matrices representing the relations between the factors and the variables (see Table 25.4, based on data introduced later in this chapter). One matrix, known as the pattern matrix, shows regression coefficients associated with the factors on each variable; therefore, these values represent the unique contributions of the factors to the variance in each variable. The other matrix, known as the structure matrix, shows the correlations between the variables and the factors. (Note that when all factors are uncorrelated, regression coefficients and correlations are identical, which is why orthogonal rotation methods generate only one factor-loading matrix.) It is the pattern matrix that most clearly expresses the simple structure achieved by an oblique rotation; therefore, this matrix should be reported whenever oblique rotations are performed. However, the structure matrix is also informative, by showing how strongly each variable is correlated with each factor; sometimes, a variable having low regression coefficients on a factor will nevertheless be strongly correlated with that factor, if factors are correlated substantially and if the variable has high regression weights on the other factors. In addition to the pattern and structure matrices, a matrix showing correlations between the factors is also generated when one performs an oblique rotation. When space is limited, a researcher may report only the pattern matrix and the factor correlation matrix, because the structure matrix is obtained by finding the product of the pattern matrix and the factor correlation matrix.

For both orthogonal and oblique rotations, there are several alternative algorithms that can be used, depending on one's strategy for achieving simple structure. For orthogonal rotations, the most widely used algorithm is varimax (Kaiser, 1958), which rotates the factors so that the variances of the squared factor loadings on *each factor* are maximized. In other words, varimax simplifies *each factor* by forcing the variables to show either strong loadings or near-zero loadings on a given factor. Another orthogonal rotation algorithm is quartimax (see, e.g., Neuhaus & Wrigley, 1954), which rotates the factors so that the variances of the squared factor loadings of *each variable* are maximized. In other words, quartimax simplifies *each variable* by forcing it to show a strong loading on one factor and near-zero loadings on the other factor(s). A third orthogonal rotation algorithm is equamax (Saunders, 1962), which represents a compromise between the varimax and quartimax algorithms. For most variable sets, the solutions yielded by these algorithms are very similar. However, varimax is usually preferred over quartimax (or the equamax hybrid), because the researcher is usually more interested in simplifying the interpretation of the factors than in simplifying the location of the variables. Moreover, varimax rotations produce rotated factors that are more nearly equal in size (as indexed by the sums of squared loadings) than are those produced by quartimax rotations. Figure 25.1 shows the varimax rotation (see axes I′ and II′) as applied to a hypothetical two-factor space.

For oblique rotations, two of the more widely used algorithms are promax (Hendrickson & White, 1964) and direct oblimin (Jennrich & Sampson, 1966). Promax proceeds by taking a varimax factor-loading matrix and then creating a new matrix by raising the factor loadings to some exponent (but without changing the sign of the loadings). The exponent, called *kappa*, is typically assigned a value of 4; when loadings are transformed in this way, they all become much smaller, but the *ratios* between the (originally) higher and lower loadings become much greater, with the latter becoming vanishingly small and thus simplifying the structure. The factors are then rescaled to return them to their original length, and the original (varimax) factor axes are rerotated in such a way as to be as close as possible to the factor axes of the new matrix. Note that when variables originally showed substantial secondary loadings, and hence were located in the interstitial factor space rather than along any one factor axis, the newly rotated factor axes will be "pulled" from the original orthogonal positions toward those interstitial regions in which variables are clustered, thereby reducing secondary loadings. As a result, those axes become oblique, as seen in the hypothetical example of axes I″ and II″ in Figure 25.1.

Direct oblimin proceeds by finding a rotation of the initial extracted factors that will minimize the cross products of the factor loadings; this generates a simple-structured solution because those cross products are small when many of the loadings are close to zero. The extent of the correlation among factors in a direct oblimin solution is influenced by a parameter, usually called *delta*, which is typically assigned a value of 0 but can range from large negative values (producing near-orthogonal solutions) to positive values (producing more oblique solutions) that hypothetically range as high as 4/3, but cannot exceed 0.8 in some computer statistical packages. In our experience, the differences between promax and direct oblimin solutions tend to be fairly small, but we tend to prefer promax because of its (relative) conceptual and computational simplicity.

As explained above, the purpose of factor rotation is usually to produce a "simple structure" solution in which each factor is defined by only a few variables, and each variable defines only one factor. For many variable sets, rotational algorithms can approach the ideal of simple structure quite closely, so that there are few variables that divide their loadings across two or more factors. But for many other variable sets, it is inevitable that many variables will divide their loadings across factors, despite the best efforts of rotation algorithms. This reflects the fact that many individual difference variables tend to be distributed throughout the space defined by two or more dimensions, not grouped neatly along a single set of axes within that space. For example, personality characteristics involving interpersonal behaviors and emotional reactions tend to be spread almost evenly throughout a *space* of at least three dimensions (e.g., Saucier, 1992), rather than being isolated along orthogonal axes. As a result, rotational algorithms may yield quite different results depending on the particular set of variables being analyzed, or even depending on the sample on which the variables are measured. It is not entirely clear whether the space mentioned above would be summarized more elegantly by a set of three orthogonal axes representing bossiness, friendliness, and anxiety, or by another set of three orthogonal axes representing shyness, irritability, and sentimentality. In such cases, one may ultimately decide on the preferred factor axis locations by finding the solution that seems to recur most frequently across variable sets and subject samples, or by finding the solution that one finds simplest to interpret theoretically.

In addition to the strategies mentioned above, there is another strategy for factor rotation to be noted, which does not involve a search for simple structure within a given dataset, but rather an attempt to match a previously obtained (or theoretically expected) factor structure as closely as possible. This approach involves the use of a targeted rotation (also called a Procrustes rotation), in which the researcher specifies a "target" matrix of loadings expected on the basis of prior data or theory, and then rotates the obtained factors in such a way as to maximize similarity to the specified target matrix. When applied to oblique factors, this method is highly problematic, because the obtained matrix can be rotated to almost any target solution, no matter how implausible, if the correlations between the factors are allowed to be extreme. When applied to orthogonal factors, however, a close approximation to the target matrix can be achieved only when the obtained matrix is indeed nearly a rotational variant of the target matrix. Formulas for calculating the similarity, or congruence, between the target matrix and the matrix calculated by targeted (Procrustes) rotation of the obtained matrix are provided by Paunonen (1997). The use of orthogonal Procrustes rotation is especially useful when a variable set is not especially simple-structured, and tends to produce rotated factor axes whose locations vary from one sample to another.

Factor Scores or Scale Scores?

After conducting a factor analysis, the researcher often wants to calculate scores for each participant on the obtained factors. When one performs a principal components analysis followed by an orthogonal rotation, perfectly uncorrelated *factor scores* can be calculated (and can be provided by statistical computing packages). When one performs a common factor analysis followed by an orthogonal rotation, the factor scores can only be estimated, usually by multiple regression, so the obtained scores will usually not be precisely uncorrelated. An alternative to calculating (or estimating) factor scores is to calculate *scale scores* corresponding to the obtained factors, by finding the unit-weighted sums of the variables that define each factor most strongly. Even when applied to orthogonal factors, however, this method will usually produce somewhat correlated factors. This is because some of the variables that define a given factor will have nontrivial secondary loadings on another factor, and rarely will these secondary loadings be perfectly balanced between positive and negative values. The fact that one can obtain orthogonal scores by calculating factor scores, but not by calculating scale scores, may be seen as an advantage of the former approach over the latter. However, a disadvantage of calculating factor scores is that the values of the factor scoring matrix are sample dependent and in small samples may fluctuate widely; this problem obviously does not afflict the calculation of scale scores via unit-weighted combinations of items.

Confirmatory Factor Analysis?

Researchers interested in testing hypotheses regarding the factor structure of a given variable set frequently employ confirmatory factor analysis (CFA). In CFA, one specifies a priori one's hypotheses regarding the number of factors, the loadings of variables on factors, and the correlations (if non-zero) between the factors. Using the covariance matrix calculated on the basis of the data obtained from one's subject sample, the CFA algorithm evaluates the extent to which the hypothesized structure matches the observed relations between the variables, reporting one or more "goodness-of-fit" statistics. Since the 1980s, CFA has been widely used in psychological research, including personality research. Despite the widespread popularity of CFA, however, we believe that researchers should be cautious in their use of this method in personality research, for the reasons that we explain here.

First, if the researcher has a large variable set—for example, the several hundred adjectives typically investigated in lexical studies of personality structure—it is implausible that the researcher would be able to specify accurately, on theoretical grounds, the behavior of so many variables. Instead, it is likely that for many of the variables the researcher would either have no real hypothesis or would misspecify their locations.

Alternatively, if the researcher has a smaller variable set—for example, the 10 to 50 scales of a typical omnibus personality inventory—the specification of the theoretically expected loadings of variables is much more feasible. Unfortunately, CFA frequently fails in these cases, by rejecting even those factor structure models that clearly replicate across different types of participant samples and clearly include all of the large factors underlying the variable set. In the case of CFA models that assume perfectly simple-structured, perfectly orthogonal factors, the obtained levels of fit are usually very poor, because most personality variables are not associated univocally with only one factor; however, even when substantial secondary loadings and/or factor intercorrelations are incorporated, the obtained levels of fit still tend to be somewhat poor (see, e.g., McCrae, Zonderman, Costa, Bond, & Paunonen, 1996). One can obtain models having somewhat better levels of fit by incorporating *all* nontrivial secondary loadings, but this undermines the purpose of testing a theoretically driven structure.

As an example of these problems, consider the NEO Personality Inventory—Revised (NEO-PI-R), whose 30 facet-level scales consistently produce the same five large factors in exploratory factor analyses, albeit with some variation in the varimax-rotated locations of two of the factor axes. When CFA is applied to even large-sample data for the NEO-PI-R scales, the result is a rejection of the five-factor model that is hypothesized on a priori grounds, even when the factors are allowed to correlate and even when substantial secondary loadings are specified (McCrae et al., 1996). In contrast, the use of targeted orthogonal Procrustes rotation (described above; see Paunonen, 1997) does support the structural model underlying the NEO-PI-R (McCrae et al., 1996).

Case Example: Zelenski and Larsen (1999)

As an example of the use of factor analysis, consider the dataset of Zelenski and Larsen (1999), who extracted and rotated three factors from a variable set consisting of 15 self-report questionnaire scales, based on responses of 86 persons. Those variables were selected to represent the constructs of the personality theories of Eysenck, Gray, and Cloninger, and together span a variety of traits related to positive and negative emotions and to impulse expression versus control. Note that, because this variable set was selected with the express aim of defining a specified (three-dimensional) factor space, the results of this analysis would not be expected to reveal the structure of the personality domain more generally. However, the analysis would be very useful for the purposes of locating each variable within this theoretically specified subspace of the personality domain, and of providing factor scores that would allow these variables' common relations with external criteria to be summarized concisely.

The correlations between the variables are shown in Table 25.1. Although Zelenski and Larsen reported results based on common factor analysis (specifically, principal-axis factoring), we show here the results based on principal components analysis. As can be seen in the list of eigenvalues and the scree plot in Figure 25.2, the differences in size between adjacent eigenvalues increase noticeably between the fourth and third factors, thus suggesting a three-factor solution. (Actually, as noted earlier, a weaker case may also be made for a four-factor solution, but here we report the three-factor solution, which allows comparison with the theoretically guided analysis by Zelenski and Larsen (1999).)

Table 25.2 shows the loadings and communalities of the variables in the unrotated factor solution. Notice that it is difficult to find a simple interpretation of these factors, as each factor shows moderately high loadings for many diverse variables. For example, the first unrotated factor is defined by variables related to positive affect, to impulse expression, and to (low) negative affect. The second unrotated factor is defined in part by variables related to negative affect, but also by other variables, such as Reward Responsiveness, Reward De-

TABLE 25.1. Correlations between the Personality Variables of Zelenski and Larsen (1999)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Drive	1.00														
2. Behavior Inhibition	–.09	1.00													
3. Reward Responsiveness	.46	.27	1.00												
4. Fun Seeking	.35	–.28	.40	1.00											
5. Extraversion	.30	–.13	.25	.39	1.00										
6. Neuroticism	–.06	.61	.21	–.23	–.24	1.00									
7. Psychoticism	.28	–.27	.12	.33	.10	–.01	1.00								
8. Punishment Expectancy	.03	.23	.02	–.02	–.08	.50	.00	1.00							
9. Reward Expectancy	.46	–.23	.30	.30	.46	–.40	.10	–.20	1.00						
10. Impulsiveness	.33	–.19	.28	.50	.29	–.07	.57	–.06	.17	1.00					
11. Venturesomeness	.23	–.33	.12	.46	.32	–.33	.32	–.09	.35	.35	1.00				
12. Novelty Seeking	.18	–.28	.21	.43	.35	–.19	.37	–.14	.18	.59	.25	1.00			
13. Harm Avoidance	–.32	.53	–.17	–.52	–.54	.60	–.32	.22	–.54	–.32	–.40	–.52	1.00		
14. Persistence	.30	.22	.29	.07	.17	.07	–.14	.10	.33	–.13	.16	–.33	–.14	1.00	
15. Reward Dependence	.00	.45	.30	.13	.19	.24	–.05	.12	.16	–.06	–.01	.03	.08	.12	1.00

TABLE 25.2. Loadings and Communalities of the Variables in the Unrotated Factor Solution

	Factors			
Variables	1	2	3	h^2
Harm Avoidance	–.83	.15	.20	*.75*
Fun Seeking	.72	.20	.16	*.58*
Novelty Seeking	.63	–.08	.47	*.63*
Impulsiveness	.63	.09	.56	*.72*
Reward Expectancy	.63	.20	–.49	*.68*
Venturesomeness	.63	–.01	–.06	*.40*
Extraversion	.62	.23	–.20	*.47*
Drive	.53	.41	–.08	*.46*
Reward Responsiveness	.35	.70	.04	*.62*
Behavior Inhibition	–.53	.67	.05	*.73*
Reward Dependence	–.02	.61	–.03	*.37*
Neuroticism	–.53	.59	.44	*.82*
Punishment Expectancy	–.26	.40	.28	*.31*
Persistence	.09	.52	–.59	*.62*
Psychoticism	.50	–.01	.55	*.55*
Eigenvalues	4.48	2.42	1.82	

pendence, and Persistence. The third unrotated factor is defined by variables related to impulse expression, and also by (low) Persistence and (low) Reward Expectancy.

When these three factors are rotated to a varimax solution (see Table 25.3), the interpretation of factors becomes much clearer. The factors of the rotated solution are defined by variables assessing (1) positive affect and sensitivity to rewards, (2) impulsivity and thrill-seeking, and (3) negative affect and sensitivity to punishments, respectively. Comparison of the varimax-rotated principal components of Table 25.2 with the varimax-rotated common (principal axis) factors reported by Zelenski and Larsen (1999, Table 2) reveals a very similar pattern of results, apart from an (unimportant) change in the order of the first and second *rotated* factors. Factor loadings are somewhat higher in the principal components analysis than in the common factor analysis (see the earlier section on the factor and component models), but even in the current example the differences are rather minor, in spite of the small number of variables per factor and the somewhat low proportions of common variance for some of those variables. Note that although we have not reported the communalities of the variables in Table 25.3, these values are identical to those of Table 25.2, because the process of rotating factors does not change the communalities of variables.

Alternatively, an oblique rotation of the factors may be performed. Table 25.4 shows the results of a promax rotation of the three extracted factors. For this oblique rotation, we have reported three matrices: first, the pattern matrix shows the regression weights of each factor on each variable; second, the structure matrix shows the correlations of each factor with each variable; third, the factor correlation matrix shows the correlations between the three factors. Note that the correlations between the three factors are rather small, with the highest correlations having absolute values of approximately .30; the pattern and structure matrices are thus relatively similar to each other and to the rotated matrix from the varimax solution of Table 25.4. In this case, the use of an orthogonal rotation (rather than oblique) seems justified: Given that there is little indication of a general higher-order factor, and given that most variables show a reasonably simple structure with respect to the varimax factor axes, there is little need to abandon the conceptual simplicity of the solution defined by mutually orthogonal factors.

The factor analysis reported by Zelenski and Larsen (1999) illustrates both the conceptual and the practical usefulness of the technique. From a conceptual standpoint, the three factors represent the major constructs that are shared among several theories of the causal basis of individual differences in emotion- and impulse-

TABLE 25.3. Loadings of the Variables in the Varimax-Rotated Factor Solution

	Factors		
Variables	1	2	3
Reward Expectancy	.76	.03	–.30
Persistence	.66	–.40	.16
Extraversion	.62	.24	–.16
Drive	.61	.27	.07
Reward Responsiveness	.61	.25	.44
Venturesomeness	.42	.36	–.30
Impulsiveness	.16	.84	.01
Novelty Seeking	.11	.77	–.16
Psychoticism	.03	.74	–.02
Fun Seeking	.48	.59	–.09
Neuroticism	–.22	–.02	.88
Behavior Inhibition	.03	–.32	.79
Harm Avoidance	–.54	–.39	.56
Punishment Expectancy	–.08	.04	.55
Reward Dependence	.35	–.04	.49
Sum of squared loadings (SSL)	3.07	2.93	2.72

TABLE 25.4. Pattern and Structure Coefficients of the Variables in the Promax-Rotated Factor Solution, and Correlations between the Factors

	Pattern coefficients			Structure coefficients (factor loadings)		
Variables	1	2	3	1	2	3
Reward Expectancy	.78	–.15	–.28	.77	.18	–.36
Persistence	.76	–.52	.14	.57	–.32	.17
Reward Responsiveness	.62	.23	.51	.61	.27	.35
Extraversion	.61	.13	–.11	.66	.35	–.24
Drive	.60	.20	.13	.65	.35	–.02
Venturesomeness	.37	.26	–.25	.49	.45	–.39
Impulsiveness	.04	.87	.13	.30	.84	–.13
Psychoticism	–.08	.79	.08	.15	.74	–.14
Novelty Seeking	–.01	.78	–.06	.24	.79	–.29
Fun Seeking	.41	.53	.00	.58	.66	–.22
Neuroticism	–.18	.16	.90	–.27	–.17	.88
Behavior Inhibition	.12	–.21	.78	–.07	–.41	.83
Punishment Expectancy	–.06	.14	.58	–.10	–.05	.54
Reward Dependence	.40	–.03	.52	.31	–.06	.47
Harm Avoidance	–.48	–.23	.51	–.63	–.53	.65
Sum of squared loadings (SSL)				3.48	3.45	3.05

	Factor correlations		
	1	2	3
Factor 1	1.00		
Factor 2	.32	1.00	
Factor 3	–.15	–.30	1.00

related personality traits. From a practical standpoint, the calculation of factor scores for the three rotated factors allowed Zelenski and Larsen an efficient means of examining the relations between this variable set and various dependent variables, including participants' emotional responses to a laboratory mood induction.

Summary

Factor analysis is a useful tool in the study of personality and of individual differences more generally. The purpose of this technique is to reduce a large set of correlated variables to a much smaller set of dimensions. Each of those dimensions, or factors, can be calculated as a linear combination of the variables, and each variable can be located within the space that is defined by those dimensions. This method is useful for the purposes of simplifying one's variable set and of understanding the common nature of the variables that jointly define a given factor.

When conducting a factor analysis, it is best to use a large participant sample that is representative of the population being examined, as well as a variable set that is representative of the domain being studied. The process of factor analysis begins with the computation of a matrix of correlations between those variables. One then extracts factors, each of which accounts in turn for the maximum amount of variance between the variables, and all of which are mutually uncorrelated. Several criteria exist for deciding how many factors to extract; these methods generally aim to find the optimal tradeoff between parsimony and completeness in explaining the covariances between the variables (in the case of the common-factor model) or the variance of the entire variable set (in the case of the component model). After extracting the desired number of factors, one may then rotate those factors in such a way as to produce a "simple structure," in which each factor is to be defined strongly by a few variables only, thus facilitating the interpretation of factors. In rotating factors, one may choose an orthogonal solution (in which all factors remain uncorrelated) or an oblique solution (in which factors will be correlated). Finally, one may choose among various methods of computing or estimating participants' scores on the factors.

Acknowledgments

This research was supported by Social Sciences and Humanities Research Council of Canada Grant Nos. 410-2003-0946 and 410-2003-1835. We thank John Zelenski for providing the data from the Zelenski and Larsen (1999) study. We also thank Richard Robins and John Zelenski for helpful comments on an earlier version of this chapter.

Notes

1. At this point, we do not yet draw any distinction between common factor analysis and principal components analysis. The fundamental processes involved in these two forms of analysis are similar except insofar as their treatment of the unique aspects of the variables' variances are concerned, as we discuss below. For the sake of simplicity, we use the term *factor* to apply both to common factors and to principal components.
2. Note that these factor loadings are not the same as the weights that are applied to the variables in calculating scores on these factors. Instead, the factor loadings can be understood as weights that are applied to factor scores to predict standardized scores on each variable.
3. However, the first unrotated factor obtained in the initial factoring process—the largest factor—usually carries important and meaningful information. For example, in factor analyses of cognitive ability tests, the first unrotated factor is usually a vector representing general intelligence (Spearman, 1904); moreover, in factor analyses of personality variables, the first unrotated factor is usually a vector contrasting socially desirable and socially undesirable characteristics.
4. However, there will be cases in which a researcher would predict, on rational grounds, that one or more personality variables defining a factor should be *especially* strongly correlated with a given criterion. In these cases, the use of factor analysis as a data reduction tool should not preclude the examination of the individual variable(s) having special a priori interest.
5. For related domains of individual differences, the problem of obtaining a representative sampling of variables within the domain also applies. Attempts to produce a complete structure of cognitive abilities (e.g., Carroll, 1993) have paid far too little attention to the problem of delineating the domain and sampling it in a representative fashion, but admittedly this problem may be difficult to solve. Researchers attempting to recover the structure of other domains may have an easier task. For example, it is conceivable that a nearly complete list of the occupations (or hobbies) practiced in a given time and place could be obtained and sampled, thus allowing a systematic

exploration of the structure of vocational (or recreational) interests. Similarly, a nearly complete list of the salient political issues in a given time and place may also be identifiable, thus allowing meaningful study of the structure of political attitudes.

6. As described here, a multi-item scale includes not only a predefined personality scale but also an ad hoc item parcel. See Bandalos and Finney (2001) and Kishton and Widaman (1994) for issues related to item parceling procedures commonly used in structural equation modeling.
7. One difficulty in the implementation of both of the above methods is that commonly used statistical computing packages do not provide algorithms for their use.
8. An example of the choice between orthogonal and oblique rotations involves analyses of the structure of hierarchically organized personality inventories, in which several broad factor scales each consist of several narrower trait (or facet) scales, each of which in turn consists of several items. If one wishes to examine the structure of the variables within any one of the presumed factor-level scales within the inventory, then one should apply an oblique rotation to factors derived from an analysis of all *items* within all facet scales belonging to that factor scale. If one wishes to examine the structure of the entire inventory, then one should apply an orthogonal rotation to factors derived from an analysis of all *facet scales* of the inventory. Note also that a factor analysis of all items within a well-constructed inventory would likely produce a few broad factors similar to those defined by the facet-level scales. However, a factor analysis of the same set of items would be extremely unlikely to generate a very large number of narrow factors that would correspond cleanly to the entire array of facet scales included in the inventory, no matter how well constructed that inventory might be. This fact argues against the suggestion (Carroll, 2002) that hierarchical factor analysis be applied to item-level personality variables.

Recommended Readings

Goldberg and Digman (1994) and Goldberg and Velicer (2006) give a very clear and thorough introduction to the use of exploratory factor analysis.

Nunnally (1978) also provides a reader-friendly introduction to factor analysis, including a pedagogically useful example of how to perform factor analysis by using the centroid method, a technique similar to principal components analysis. Nunnally and Bernstein (1994) give an expanded discussion of several aspects of factor analysis.

McCrae and colleagues (1996) demonstrate the difficulties associated with the application of confirmatory factor analysis to personality research, and also give a nice example of targeted orthogonal Procrustes rotation, the use of which is explained by Paunonen (1997).

For readers interested in the mathematical basis of factor analysis, textbooks on this topic include Harman (1976), Gorsuch (1983), and Cureton and D'Agostino (1983).

References

Ashton, M. C., Jackson, D. N., Helmes, E., & Paunonen, S. V. (1998). Joint factor analysis of the Personality Research Form and the Jackson Personality Inventory: Comparisons with the Big Five. *Journal of Research in Personality*, *32*, 243–250.

Ashton, M. C., & Lee, K. (in press). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review.*

Ashton, M. C., Lee, K., & Goldberg, L. R. (2004). A hierarchical analysis of 1,710 English personality-descriptive adjectives. *Journal of Personality and Social Psychology*, *87*, 707–721.

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

Bandalos, D. L., & Finney, S. J. (2001). Item parceling issues in structural equation modeling. In G. A. Marcoulides & R. E. Schumaker (Eds.), *Advanced structural equation modeling: New developments and techniques* (pp. 269–296). Mahwah, NJ: Erlbaum.

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology*, *3*, 77–85.

Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. *British Journal of Psychology*, *4*, 1–2.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Carroll, J. B. (2002). The five-factor personality model: How complete and satisfactory is it? In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 97–126). Mahwah, NJ: Erlbaum.

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

Cureton, E. E., & D'Agostino, R. B. (1983). *Factor analysis: An applied approach*. Hillsdale, NJ: Erlbaum.

Everett, J. E. (1983). Factor comparability as a means of determining the number of factors and their rotation. *Multivariate Behavioral Research*, *18*, 197–218.

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

Goldberg, L. R. (2006). 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. R., & Digman, J. M. (1994). Revealing structure in the data: Principles of exploratory factor analysis. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 216–242). New York: Springer.

Goldberg, L. R., & Velicer, W. F. (2006). Principles of exploratory factor analysis. In S. Strack (Ed.), *Differentiating normal and abnormal personality* (2nd ed., pp. 209–237). New York: Springer.

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

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103*, 265–275.

Harman, H. H. (1976). *Modern factor analysis* (3rd ed., rev.). Chicago: University of Chicago Press.

Hendrickson, A. E., & White, P. O. (1964). Promax: A quick method of rotation to oblique simple structure. *British Journal of Statistical Psychology*, *17*, 65–70.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*, 417–441, 498–520.

Jennrich, R. I., & Sampson, P. J. (1966). Rotation for simple loadings. *Psychometrika*, *31*, 313–323.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*, 187–200.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.

Kishton, J. M., & Widaman, K. F. (1994). Unidimensional versus domain representative parceling of questionnaire items: An empirical example. *Educational and Psychological Measurement*, *54*, 757–765.

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

McCrae, R. R., Zonderman, A. B., Costa, P. T., Jr., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors on the revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, *70*, 552–566.

Neuhaus, J. O., & Wrigley, C. (1954). The quartimax method: An analytical approach to orthogonal simple structure. *British Journal of Statistical Psychology*, *7*, 81–91.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Paunonen, S. V. (1997). On chance and factor congruence following orthogonal Procrustes rotation. *Educational and Psychological Measurement*, *57*, 33–59.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine, Series 6*, *2*, 559–572.

Saucier, G. (1992). Benchmarks: Integrating affective and interpersonal circles with the Big-Five personality factors. *Journal of Personality and Social Psychology*, *62*, 1025–1035.

Saunders, D. R. (1962). Trans-varimax. *American Psychologist*, *17*, 395.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*, 321–327.

Zelenski, J. M., & Larsen, R. J. (1999). Susceptibility to affect: A comparison of three personality taxonomies. *Journal of Personality*, *67*, 761–791.

CHAPTER 30

Studying Personality Processes

Explaining Change in between-Persons Longitudinal and within-Person Multilevel Models

William Fleeson

The purpose of this chapter is to guide the use of analytic models for studying the processes underlying personality variables. It is important that personality psychologists have available powerful and widely understood methods for studying processes, for at least three reasons (Cervone & Mischel, 2002; Fleeson, 2004; Funder, 2001; Pervin, 1994). First, interest in process has always been strong among personality psychologists (Allport, 1937; Cattell, 1966; Larsen, 1989), but has grown recently as personality psychologists become increasingly interested not only in what personality is but also in how it works (Funder, 2001; McCrae & Costa, 1999; Mischel, 2004). Second, such advancement in understanding process is the hallmark of a mature science and reflects the field's growing self-confidence. Finally, advancement in describing processes will reveal the potential for personality change as well as promising opportunities to effect such change. Meeting this growing interest in studying processes have been assessment, computing, and conceptual advances that make studying processes more practical. This chapter provides initial guidance in implementing these advances.

Process refers to a combination of actions, changes, or events that bring about an outcome. The outcome can be the creation of something new or a change in the existing level of a variable. Personality psychologists typically investigate the latter, with research directed at explaining change in existing variables such as traits, well-being, behavior, or life events. Processes can be simple, such as when one change brings about another change, or they can be complex, such as when multiple events cascade in a temporal sequence to bring about an ultimate outcome. Processes can also be ongoing, in cases when the outcome fluctuates and its value at any given time is determined by other concurrent events in a mathematically describable manner. By identifying the steps or changes that bring about an outcome, processes describe how things work (the mechanism) and identify points at which interventions can efficiently produce changes in the outcome.

Structure is the complement to process; a structure consists of the variables or parts of personality and their typical or fixed relationship to each other. Processes operate on and within these structures. Although the term *structure* is often used as a shorthand for *covariance structure*, referring to the correlations between individual differences in traits such as the Big Five, it is also used more generally to refer to any set of variables and their relationships. A complete description of personality requires knowledge of both the structures being operated on and the processes operating on them. Personality psychology primarily has focused on identifying variables and their covariance structures, but it is also necessary to enhance research on processes (Cervone, 2004; Fleeson, in press).

Most psychological processes happen within a person, but processes can be studied by a between-persons or by a within-person approach (Epstein, 1983; Lamiell, 1997).[1] The traditional between-persons approach usually investigates simpler processes, attempting to explain change in one variable by changes in or actions of only one or a few other variables. Between-persons approaches do so by testing whether differences between people in an outcome variable are related to corresponding differences between people in an explanatory variable. For example, researchers may investigate whether extraversion leads to happiness by testing whether happier people are that way because they are more extraverted than other people (Costa & McCrae, 1980). Once such a relationship is identified between persons, it is inferred that a process occurred within each individual to lead to their respective levels on the outcome variable. However, the correspondence between between-persons relationships and within-person processes is complex. Sometimes the directly analogous process is inferred to operate within persons. For example, because individuals who use problem-focused coping methods have less stress, it is inferred that when each individual uses a problem-focused coping method, his or her stress will be reduced. At other times the between-persons relationship indicates where to look for a within-person process. For example, extraverts' higher level of happiness has directed some researchers to investigate what it is that extraverts do that leads to happiness (e.g., Rusting & Larsen, 1998). At still other times, the process can be studied only between persons, because the explanatory variable changes so rarely or so slowly that within-person approaches are impractical. For example, extraverts may be happier because of enhanced functioning of the dopamine system, as determined early in development (Depue & Collins, 1999).

Within-person approaches investigate processes by assessing each individual on multiple occasions and comparing those occasions to each other, one individual at a time. The degree to which the outcome and explanatory variable covary across occasions is calculated for each individual, and the average of these covariances describes the within-person relationship (averaging is necessary because individuals show slightly different relationships owing to chance alone). Within-person approaches may reveal different answers than between-persons approaches reveal, because variables may vary across persons for different reasons than they vary within people across occasions (Robinson, 1950). For example, across persons the relationship between extraversion and positive affect is strong and positive, but this may not mean that the relationship within persons between extraversion and positive affect is positive. In fact, changing the amount anyone acts extraverted may have no impact on that person's happiness, because extraversion and positive affect may vary within people for different reasons than they vary between people (Fleeson, Malanos, & Achille, 2002). Often the within-person processes are the processes of primary interest to psychologists, because psychologists are interested in how the mind works. However, personality psychologists tend to use the more practical between-persons approaches and then make inferences about the processes happening within a person.

This chapter describes both between-persons approaches and within-person approaches, with the goal of promoting the study of process. Although process is studied with many types of psychological variables, this chapter is directed at researchers interested in personality concepts, such as traits, goals, attachment styles, well-being, personality disorders, self-concepts, and coping. I hope to provide enough detail to guide the actual use of these methods. However, rather than focus on technical details, I focus on how statistical techniques can be used to answer interesting questions about how personality works.

Between-Persons Approaches to Studying Change

All research investigating process shares the goals of identifying the causes and consequences of personality variables and of discovering how personality variables work. When studied between persons, the effort is usually known as studying "change." This is mainly because the common longitudinal method is to measure the outcome twice and test whether an explanatory variable predicts changes in the outcome variable from the first occasion to the second. It is a between-persons approach because it tests whether differences between individuals in the explanatory variable are aligned with differences between individuals in change. For example, Carver and Scheier (1994) investigated whether individual differences in coping are correlated with individual differences in stress change from a semester beginning to a semester end. An important advantage of studying change is that, when done correctly, it enhances the ability to make conclusions about causality. Thus, researchers can not only identify the variables that are involved in the process of change, but can also have confidence that the direction of causality flows from the explanatory variables to the outcome, rather than the reverse.

This section of the chapter describes methods that allow conclusions about causality and change and diverts the reader away from some tempting but ultimately less effective methods. Although researchers often use regression to study change, explications of the appropriate methodological and analytic procedure are not widely published and researchers are hard-pressed to find guidance on these issues (Darlington & Smulders, 2001). Because change is so critical to personality theory and because suboptimal analytic techniques are occasionally used, it is important for the appropriate analyses to be available to personality researchers.

Step-by-Step Guide

Longitudinal designs are the designs of choice for at least two reasons: (1) They provide a more rigorous test of causal hypotheses than is possible using concurrent designs, and (2) they allow time to pass so that the processes in question have an opportunity to produce change. Longitudinal designs entail at least two times of measurement of the same participants. For example, a researcher may measure coping and stress at both the beginning of a semester (time 1) and the end of the semester (time 2), or measure extraversion and happiness each at the beginning of the study (time 1) and then again 10 years later (time 2). The hypothesis tested in such designs is that the explanatory variable is responsible for change in the outcome—more precisely, that differences in the explanatory variable will produce changes in the outcome.

At least four variables are needed to test this hypothesis: a participant identification variable, the explanatory variable (EV; also known as the independent variable) measured at time 1, and the outcome (also known as the dependent variable) measured both at time 1 and at time 2. Figure 30.1 shows this design and three possible regressions for investigating the effect of the explanatory variable on the outcome. It is well known that the first regression, a standard cross-sectional regression in which the time 1 outcome is predicted from the time 1 explanatory variable, is not sufficient for determining causal direction, because any relationship between the two variables could be due to the explanatory variable causing the outcome, the outcome causing the explanatory variable, or a third variable causing both.

A second possibility, improved but still not sufficient, is to predict the outcome at time 2 rather than at time 1. This analysis, part of the cross-lagged approach, is an improvement because the future cannot affect the past—the outcome at time 2 cannot have caused the explanatory variable at time 1, so any revealed relationship between them cannot be due to the outcome from time 2 causing the explanatory variable. However, the explanatory variable in such an analysis is not independent of influence from the outcome, because it is not independent of the outcome at time 1. If the outcome is stable at all from time 1 and to time 2, the joint influence of the time-1 outcome on the explanatory variable and the stability of the outcome is more than enough to produce a spurious relationship between the explanatory variable and the time-2 outcome, even if the explanatory variable does not cause the outcome. Thus, the causal direction between the explanatory variable and the outcome remains ambig-

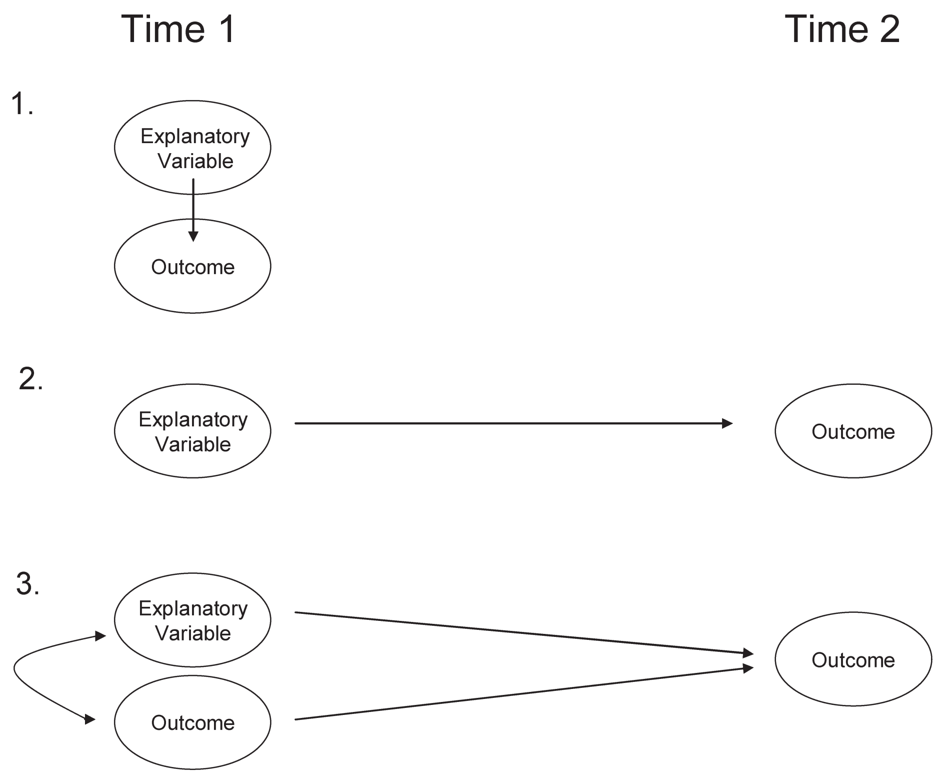


FIGURE 30.1. Three possible regressions for investigating the effect of the explanatory variable on the outcome. The bottom represents the recommended analysis.

uous in such an analysis. To fix this problem, the explanatory variable must be made independent of the outcome from time 1.

The final and recommended analysis is a multiple regression, predicting the time-2 outcome simultaneously from two variables: the time-1 explanatory variable and the time-1 outcome. Whenever multiple independent variables (IVs) are included in a simultaneous multiple regression, the relationships are calculated for only part of each IV, namely, the part of each IV that is independent of the other IVs. In the recommended analysis, this means that the relationship for the explanatory variable is calculated for the part of the explanatory variable that is independent of the time-1 outcome (allowing for imperfection in measurement of the time-1 outcome). If this relationship exists, it cannot be due to the time-1 outcome causing the explanatory variable (because the regression made the explanatory variable independent of the time-1 outcome) and it cannot be due to the time-2 outcome causing the explanatory variable (because the future cannot affect the past), so it must go in the other causal direction (or be due to a third variable).

The magnitude and direction of the relationship are described by the unstandardized beta for the explanatory variable (the slope). In particular, an unstandardized beta communicates the number of points the time-2 outcome, changes for every point the part of the explanatory variable that is independent of the time-1 outcome changes (these points are the points on the scales of the explanatory variable and the outcome). In this case, the multiple regression makes the explanatory variable independent of the time-1 outcome and then predicts the time-2 outcome from that independent part. Because the explanatory variable is independent of the time-1 outcome, any relationship between the explanatory variable and the time-2 outcome must have arisen between times 1 and 2, and thus cannot be caused by the outcome.

This analysis supports conclusions both about change and about causal direction. Because the analysis "equates" the participants statistically on the outcome at time 1, it is in fact predicting change in the outcome. That is, the analysis reveals what happens to the time-2 outcome among individuals who at time 1 have the same level on the outcome but different levels on the explanatory variable. If the outcome changes from time 1 to time 2, and these changes are caused by the explanatory variable, then people who started with the same standing on the outcome will change to a standing on the outcome as a consequence of their time-1 explanatory variable standing. By conducting a multiple regression in this fashion, the researcher has directly tested whether the process creating change in the outcome involves this particular explanatory variable.

Case Examples

Extraversion and Positive Affect

An important personality finding of the last 20 years is that positive affect is associated more strongly to extraversion than to any other variable, including objective circumstances (Costa & McCrae, 1980; Lucas & Baird, 2004; Staudinger, Fleeson, & Baltes, 1999). This finding was first discovered by Costa and McCrae in 1980, when they proposed that internal personality factors might be one component in the process that results in happiness. Even though the process was presumed to work within the one individual, Costa and McCrae used a between-persons method. They used a longitudinal design, measuring extraversion at time 1 and positive affect 10 years later at time 2. Unfortunately, they did not have positive affect measures available at time 1, so they could use only the second of the analysis techniques illustrated in Figure 30.1. They found that extraversion predicted positive affect even 10 years later ($r = 0.23$). That is, individuals who differed from each other in extraversion at time 1 had corresponding differences in positive affect 10 years later. Because they could not control for time-1 positive affect, it is possible that this correlation was due to positive affect causing extraversion rather than extraversion causing positive affect. Even though Costa and McCrae were unable to determine unambiguously the direction of causality, the suggestion that extraversion might be involved in the process determining happiness was so important that it inspired decades of research to further investigate the process. Only recently has an experimental approach been used to conclude that extraversion in fact causes positive affect (McNiel & Fleeson, 2006).

Coping and Mental Health Symptomology

Stressors impact mental health, but do so differently for different people. The way an individual copes with stressors has been proposed as an important part of the process affecting change in mental health symptomology (Lazarus, 2000). In particular, trying to avoid the problem has been proposed to be generally less effective at reducing symptomology than is acknowledging or trying to solve the problem that led to the stressor. Aldwin and Revenson (1987) tested whether coping affected change in symptomology with a between-persons longitudinal approach. They measured symptoms (on a 0–44 scale) two times about a year apart, along with several ways of coping (on 1–4 scales) at the first time point.

They predicted time-2 symptoms (the outcome) in a multiple regression with the several time-1 ways of coping (the explanatory variables) and time-1 symptoms as the independent variables. Three of the coping styles had significant relationships to time-2 symptoms. Because time-1 symptoms were controlled in the analysis and time-2 symptoms cannot affect the past, these significant relationships cannot be due to symptoms causing ways of coping, but rather must be due to ways of coping (or a third variable) causing symptoms. In particular, escapism (0.45) and social support mobilization (0.20) predicted increased symptoms at time 2, whereas instrumental action (–0.18) predicted reduced symptoms. These betas describe the changes in symptoms for each point change in the respective ways of coping, independent of all other ways of coping and of time-1 symptoms (it was unspecified in the article whether these betas were unstandardized betas). Thus, the results show that coping is likely to be an important variable in the process affecting changes in mental health symptomology.

Complexities

The basic analyses as described above are fairly straightforward, but there are several complications that may arise in the course of any particular study. Here I discuss two important and frequent complications that arise from researchers attempting other less optimal ways to measure change and two important caveats to the conclusions based on these methods. For additional complexities and possibilities, see Darlington (1990).

Difference Scores as Outcomes

An intuitively appealing but less effective technique for predicting change is to compute the difference in the outcome variable between time 1 and time 2 and then use the difference score as the dependent variable. Such a method appears to measure change, because time-1 scores on the outcome are subtracted from time-2 scores on the outcome so that positive numbers indicate increases in the outcome and negative numbers indicate declines in the outcome.

What undermines this technique is that difference scores are usually correlated with time-1 scores on the outcome. Participants who scored higher on the outcome at time 1 are more likely to have a low or negative difference score (more likely to decline), and those who scored higher on the outcome at time 1 are more likely to have a high or positive difference score (more likely to increase on the outcome).

This correlation between time-1 scores and difference scores creates two problems. The first problem is an ambiguity about what the outcome is. When the difference scores are correlated with time-1 scores, the difference scores reflect two variables (i.e., change over time and the time-1 score). Thus, it will not be clear which of these two concepts is responsible for any observed association to the explanatory variable. In particular, the time-1 outcome may be associated with the explanatory variable and so produce a spurious association of the explanatory variable to the difference score when it actually has no relationship to change on the outcome. The way to disambiguate this is to partial the time-1 outcome out of the explanatory variable, as explained above. The second problem is that the causal direction of the effect cannot be specified. Because the difference score contains variance of the time-1 outcome, any relationship between the difference score and the explanatory variable may be due to the time-1 outcome rather than to the explanatory variable. What is desired by the researcher is to predict change that is independent of the time-1 outcome—the analysis for doing so is the recommended analysis described above (replacing the time-2 outcome by difference scores in the recommended analysis will not change the unstandardized beta for the explanatory variable).

In contrast to this position, Rogosa (1995) makes a spirited defense of the raw difference score as the best measure of change, partly because it straightforwardly assesses whether each individual changed on the outcome. However, correlations between an explanatory variable and the difference score are almost always ambiguous as to whether they are due to the value of the time-1 outcome or to change. Specifically, if the explanatory variable is related to the outcome at time 1, and the outcome at time 1 in turn is negatively related to change, this will introduce a spurious negative component to the association between the explanatory variable and change. The spurious negative component may mask an underlying positive effect of the explanatory variable on the time-2 scores. The way to check for this is to compare those individuals with the same value of the outcome at time 1 but different values of the explanatory variable, and to observe whether they end up with different resulting scores of the time-2 outcome (i.e., to conduct the recommended analysis). However, I do agree with Rogosa that analyzing change at the individual level has many advantages, as is explicated in the subsequent section "Within-Person Multilevel Models for Studying Process."

Residualized Change Scores as Outcomes

Recognition of the problems with difference scores has led to occasional use of residualized change scores. Also derived from longitudinal designs, residualized change scores are created before the main analysis by conducting a preliminary multiple regression in which the time-2 outcome is predicted from the time-1 outcome. The residuals are saved from this analysis and are similar to difference scores in that they describe the direction and magnitude of change from time 1 to time 2. However, residuals are improved over difference scores in that they are independent of the time-1 outcome (the residualizing removes the dependency). Thus, participants are no more likely to have a positive direction of change only because of starting low and no more likely to have a negative direction of change only because of starting high.

However, residualized change scores still are not recommended for studying change for at least two reasons. First, removing the time-1 outcome from the time-2 outcome (by residualizing) but not removing it from the explanatory variable will leave irrelevant and obscuring variance in the explanatory variable (namely, variance related to the outcome at time 1). This irrelevant variance may seriously reduce the explanatory variable's apparent effect (Darlington & Smulders, 2001; Furr, 2005) and underestimate its role in producing change in almost all cases. Because the recommended analysis does not have this problem, and already reveals the explanatory variable's effect specifically on change in the outcome, it is almost always superior. (Removing the time-1 outcome from both the explanatory variable and the time-2 outcome produces the same unstandardized beta as does the recommended analysis above.)

Second, no modification to the outcome is necessary for testing hypotheses about change in the outcome. This is because it is very difficult for an explanatory variable to have a causal impact on a time-2 outcome score without changing that score. Thus, the recommended analysis already predicts change in the outcome, even without directly calculating a change score.

It is possible, however, as some researchers propose, that the variables that cause change in an outcome are different from those that cause the creation or level of the outcome. Such researchers are proposing a moderator variable, because they are proposing that the effect of the explanatory variable on the outcome depends on some other factor, such as time. Testing a moderator hypothesis requires identifying that moderator variable, assessing it, and using an interactive term to test the theory (see Chaplin, Chapter 34, this volume), not partialling the outcome on an earlier time. For example, a researcher may propose that the outcome is created at a young age by one set of variables and then changed at older ages by another set of variables (e.g., Sroufe, Egeland, Carlson, & Collins, 2005); here age is proposed as a moderator of the effect of those variables on the outcome. Linear changes to the outcome do not represent moderation and do not adequately test such theories (Cronbach & Furby, 1970; Darlington, 1990).

In sum, residualized change scores may result in inaccurate results and gain little conceptually in regard to predicting change rather than level (Cronbach & Furby, 1970; Darlington, 1990). The only potential gain is a reduction in total variance in the time-2 outcome, which may increase percentages of explained variance under some conditions.

Measurement Error in the Outcome

The final two complexities present limitations to the conclusions that can be drawn from these analyses. First, inevitable measurement error in the outcome prevents absolute certainty about causal direction. The ability to conclude causality depends on removing the time-1 outcome from the explanatory variable, so that any remaining relationship to the later outcome is entirely independent of the earlier standing on the outcome. However, this is accomplished only imperfectly. Researchers must use a measurement of the time-1 outcome rather than the true standing. All measurements have error, so this procedure does not entirely remove the true standing on the earlier outcome from the explanatory variable—it removes most but not all of it. Thus, the explanatory variable may remain slightly contaminated by the time-1 outcome. This slight contamination is unlikely to account for observed relationships between the explanatory variable and the time-2 outcome and is not usually a serious problem, but it is important to be aware of this problem, particularly in the case of outcomes that have poor reliability.

The Ecological Fallacy

The second caution concerns making conclusions about processes within individuals on the basis of differences between individuals (what Robinson, 1950, called the ecological fallacy). Between-persons research investigates psychological processes by relating differences (variances) across individuals and determining whether individuals who differ on the explanatory variable also differ systematically on an outcome variable. Part of the value in identifying between-persons relationships comes from the inference that the same relationship holds up within people. For example, the finding that individuals who use instrumental action to cope with problems accrue fewer symptoms down the road obtains value partly from the inference that individuals will reduce their symptoms when they use instrumental action. However, this inference changes *what* the differences are across, and this change may weaken the inference. Rather than relating differences across individuals, within-person relationships relate differences across occasions, within each person taken one at a time. Differences across occasions may occur for different reasons than those producing differences across people; thus, the relationships among those differences are not likely to be the same as the relationships among differences across people.

The recommended between-subjects analysis goes a long way in rectifying this problem; this is one reason the recommended analysis is promoted in this chapter. However, direct inferences to a within-person process are still vulnerable to the ecological fallacy. Part of the reason for this is that between-persons and within-person designs typically differ in time frame, and different processes may operate at different speeds. Additionally, between-persons designs are sensitive to both long- and short-term processes, even when the occasions differ only by a short time. Another reason is that between-persons designs are vulnerable to between-person third variables such as response styles (within-person designs are also vulnerable to some third variables).

For example, within individuals, amount of exercise is likely positively associated with fatigue, because an individual will be more fatigued on occasions he or she exercises than on occasions he or she does not. However, between individuals, amount of exercise may be negatively associated with fatigue, because individuals who exercise more than others will likely reduce their overall fatigue (Puetz, O'Connor, & Disham, 2006). Thus a between-persons design may lead to the conclusion that exercise reduces or does not affect fatigue whereas a within-person design may lead to the conclusion that exercise increases fatigue. This may be true even if the recommended between-persons design is employed and even if the occasions in the between-persons design are the same as the occasions in the within-person design.

Another example is the between-persons relationship between extraversion and positive affect: Individuals who are higher than others on extraversion tend to be higher than others on positive affect. This relationship partly obtains value from the inference that individuals may have a route to enhanced life quality, in that variation within each person across occasions in his or her extraversion may predict variation in positive affect. However, extraversion variation within individuals across hours may not be associated with variation within individuals in positive affect, because extraversion variation within a person across hours may occur for different reasons than does extraversion variation between individuals. If extraverts' greater happiness than introverts' is due to a greater number of dopamine neurons (Depue & Collins, 1999), within-person variations in how extraverted a person is acting at the moment would not have an impact on positive affect. In this case, a between-person design would lead to the conclusion that extraversion predicts increases in happiness, whereas the within-person design would lead to the conclusion that extraversion has no effect on happiness.

Finally, Marco, Neale, Schwartz, Shiffman, and Stone (1999) tested whether the times the average individual used problem-focused coping showed reduced stress as compared with the times the same individual did not use problem-focused coping. Every hour for 2 days, participants described whether a stressor occurred in the previous hour, how they coped with it, and their current mood. A within-person analysis revealed no significant predictions of later mood from the coping strategies used. Thus, the between-persons approach led to the conclusion that problem-focused coping lowers negative affect, whereas the within-person approach led to the conclusion that problem-focused coping does not lower negative affect. It is unknown whether this was due to the time frame used (the positive effects of problem-focused coping may take a while to manifest), to the ineffectiveness of coping in reducing negative affect, or to some other cause.

In many cases, the inference from a between-persons relationship to the within-person relationship will be accurate at least in direction. Correspondence increases in likelihood if the occasions in the two designs are more similar. However, the inference is rarely likely to be accurate in the precise magnitude and occasionally will be inaccurate even in direction (Robinson, 1950). For these reasons, researchers need to be cautious when making this inference.

Within-Person Multilevel Models for Studying Process

The goal of within-person approaches to studying process is to identify the causes and consequences of personality variables and to discover how personality works. By studying within-person processes directly, observing changes within one individual unfolding over time, within-person approaches avoid the ecological fallacy.[2] The basic plan is to observe a (usually small) number of individuals each over a period of time, measuring a few variables repeatedly. Then the variations and covariations of those variables are analyzed to characterize each individual in terms of his or her particular patterns of variability and his or her own particular functional processes (Allport, 1937; Epstein, 1983; Lamiell, 1997; Nesselroade, 1991; Tennen, Affleck, Armeli, & Carney, 2000). Patterns and processes can also be compared across individuals to identify common processes that operate within all individuals. This section describes multilevel modeling of experience-sampling data as an ideal method for studying process within persons. However, processes can be studied with other within-person approaches as well, such as dynamic *p*-factor analysis, spectral analysis, and triple-typology analysis (see Singer & Willett, 2002; Larsen, 1989; Nesselroade, 1991; Vansteelandt & VanMechelen, 2004).

Within-person approaches to process can address the same kinds of questions as are addressed by between-persons approaches. A few examples are whether extraversion predicts positive affect (Fleeson et al., 2002), whether stressors increase stress and whether coping reduces stress (Marco et al., 1999; Mroczek & Almeida, 2004), and whether situations influence trait-manifesting behavior (Fleeson, in press). There are several important ways in which within-person approaches differ from between-persons approaches. First, the outcome variable both increases and decreases over time within the same individual. The intention accordingly shifts from trying to predict whether and in which direction the outcome changes, to trying to predict what level the outcome will have at any moment, given the concurrent values of the explanatory variables. Second, within-person approaches are most effective at studying rapid processes and rapidly varying outcomes. The research goal often is to predict changes that occur over the course of days, hours, or even seconds rather than over years. Third, it is more difficult to assess direction of causality because the explanatory variables and outcomes are measured simultaneously, and the lags of causal effect are often shorter than the intervals between measurements. Fourth, the within-person approach starts with the individual: It characterizes one individual at a time by his or her own particular processes and functions. Then, it also allows generalization across individuals and quantification of differences between individuals in their processes. Fifth, as a result of these characteristics, within-person approaches focus on assessing the ongoing psychological functioning of individuals as they live their lives in naturalistic settings. They aim to obtain mathematical functions to describe and predict what state an individual will be in at each and every given moment.

Multilevel modeling facilitates addressing at least two types of theoretical question. The first type of question concerns what the within-person process is. The within-person process is the relationship between variables across occasions, taking each individual one at a time and then averaging across them. Averaging across individuals is necessary because chance and other factors will lead to each individual revealing a slightly different relationship between the variables in any particular study. Averaging cancels out these chance variations. Multilevel modeling produces one answer to this question, which describes the within-person relationship between the variables for the average or typical individual. This one answer may be different from the answer provided by a between-persons approach because different psychological principles may determine variance and covariance across occasions than the principles that determine variance and covariance across persons.

The second type of question starts with taking the differences between individuals in the within-person relationship seriously and investigates whether those differences are due to more than just chance. Such differences in the covariance relationship may in fact represent differences between people in how they function psychologically. This question marks a radical distinction from between-persons approaches. Between-persons approaches assume that the same psychological process operates within all individuals; it may not be evident in a given individual in a given study, but that is due to chance or to other processes masking the operation of the one principle. In contrast, multilevel modeling addresses the possibility that the psychological process operates differently for different people and that these differences describe enduring characteristics of the individuals' personalities. Multilevel modeling tests this possibility by testing whether differences between individuals in the relationship between the variables are more than would be expected from chance.

Step-by-Step Guide

Multilevel modeling (MLM) is actually quite simple, just like the ordinary least squares (OLS) regression that is routine for most personality psychologists. However, MLM was developed for a wide diversity of purposes, and the computer programs were written to accommodate the many possible purposes. This diversity of options and capabilities in applying MLM to a specific case can be overwhelming. The purpose of this section is to provide a guide specifically focused on applying MLM to the particular case of studying personality processes. There are other excellent guides to MLM (e.g., Bolger, Davis, & Rafaeli, 2003; Kreft & de Leeuw, 1998; Nezlek, Chapter 29, this volume) that are more general and written for the multiple diverse applications that MLM is capable of.

The data for this technique typically consist of a few explanatory variables and a few outcomes measured for only a few participants (20–50 participants are usually enough) on a large number of occasions for each participant (e.g., 25–200 per participant). For example, 50 participants may report how extraverted they acted and how much positive affect they experienced every 3 hours for 2 weeks. The power of these studies comes from the large number of occasions per participant rather than from the large number of participants.

I provide specific instructions for conducting MLM in SPSS, because this is a widely used program and because it embeds MLM into a wide array of data manipulation and analytic capabilities. Some familiarity with SPSS is assumed; I am basing instructions on SPSS 14, so details may vary in other versions (commands appear to be identical in SPSS 15.0). However, there are other excellent programs available, such as HLM 6 (Raudenbush, Bryk, & Congdon, 2004) and SAS. MLM requires some preparatory work. The first step is to make sure the data are in the correct form. Each row must correspond to one occasion, with occasions sorted by participant. Thus, each participant will have many rows. There should be one column for each variable, plus one additional column indicating the participant who generated the occasion (e.g., the participant's ID number). If, alternatively, the data are organized such that each participant has only one row, with different variables for each occasion, the "restructure" dialogue box in the SPSS data menu can restructure the data into the correct order.

Centering is an important consideration, and how it should be done depends on the particular interest of the researcher (Kreft & de Leeuw, 1998). However, when trying to discover within-person processes, it is almost always best to center the explanatory variables within each person, to be sure that between-persons variance does not contaminate the result and that it is a pure description of within-person processes. To center variables within people, use the aggregate command in the "data" menu (after putting the data in the correct form). Put the participant ID into the "break variable" box and the explanatory variable into the "summaries of variables" box, and make sure "add aggregated variables to active dataset" is checked. If this has successfully created a new variable, with each participant's mean on the explanatory variable indicated for each of his or her occasions, then subtract this new variable from the original explanatory variable to make the centered explanatory variable that will be used in subsequent analyses. Open the compute dialogue under the transform menu, type the name of the explanatory variable (followed by "_centered") in the "target variable" box, and put the explanatory variable, a minus sign, and the newly aggregated mean variable in the "numeric expression" box.

The analysis program in SPSS is called Mixed Models–Linear and was designed very broadly to be able to accomplish a wide variety of purposes. Getting it to conduct a within-person MLM is a bit awkward. After opening Mixed Models—Linear in the Analyze drop-down menu, put the participant ID in the "subjects" box and click "continue"; put the outcome in the "dependent variable" box and the centered explanatory variable in the "covariate" box (factors are used when the explanatory variable is a categorical variable). In the "fixed . . . " dialogue, add the explanatory variable to the "model" box, and make sure that "include intercept" is checked. In the "random . . . " box, choose unstructured for the covariance type, add the explanatory variable to the "model" box, and check "include intercept." Also move the participant ID to the combinations box (this step is easy to overlook because a similar step was completed earlier). In the “statistics . . . ” dialog check “parameter estimates” and “tests for covariance parameters.”

“Fixed” and “random” are used in multiple ways in statistics and in MLM. In this case, a *fixed effect* refers to an effect that is the same for all participants and a *random effect* refers to an effect that is allowed to vary randomly across participants. In most cases, in investigating personality processes, the effect of the explanatory variable will be assumed to have both a fixed and a random part. The fixed part refers to the average or universal effect of the predictor on the outcome for people; the random effect refers to the fact that the effect of the predictor may differ across individuals for reasons particular to each individual. In terms of the intercept, the fixed part refers to the overall average level of the outcome for people in general; the random part refers to individual aspects of each participant that give him or her a different average level of the outcome across occasions.

The “covariance type” refers to the constraints put on the matrix describing the variances and covariances among the coefficients (the intercept and one for each beta) across participants. This matrix indicates how much participants differed in their intercepts or effects of the explanatory variable, and how much the effects of the explanatory variable were related to each other and to the intercept. Choosing “unstructured” for the covariance type means that there are no constraints and that the matrix will be determined by the variances and covariances in the data.

Before interpreting the results, it is important to look for warnings or errors. If there are warnings, it is best to try to resolve them because, otherwise, they may mean that the output is incorrect. For example, if the analysis fails to converge, try increasing the number of iterations in the estimation dialog box. Other reasons for nonconvergence include too many parameter estimates or setting up the analysis incorrectly. Compare your syntax to that in Table 30.1 to identify possible errors.

MLM provides a coefficient to address the first type of question, whether the explanatory variable predicts changes in the outcome variable for the average or typical person. The coefficient is just like an unstandardized coefficient from regression, with magnitude and direction to indicate whether and how strongly the explanatory variable predicts changes in the outcome. This coefficient is the first entry in the row labeled by the explanatory variable, in the “estimates of fixed effects” table. Because the coefficient is unstandardized, it describes the number of points the outcome changes for the average person when the explanatory variable changes one point (these points are the points on the scales of the predic-

TABLE 30.1. Example Syntax Generated by Pasting from a Successful Multilevel Model Run in SPSS

```
sort cases by ID.

AGGREGATE
 /OUTFILE=*
 MODE=ADDVARIABLES
 /BREAK=ID
 /extra_mean = MEAN(extra).

COMPUTE extra_centered = extra - extra_mean .
EXECUTE .

MIXED
 positiveaffect WITH extra_centered
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
 SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE)
 PCONVERGE(0.000001, ABSOLUTE)
 /FIXED = extra_centered | SSTYPE(3)
 /METHOD = REML
 /PRINT = SOLUTION TESTCOV
 /RANDOM INTERCEPT extra_centered | SUBJECT(ID) COVTYPE(UN) .
```

tor and outcome). The rest of the row for the explanatory variable provides a significance test on this coefficient. A significant coefficient means that an ongoing process has been identified that describes and predicts the variation in the outcome.

The intercept row describes the average person's score on the outcome on those occasions when the explanatory variable has a score of 0. If the explanatory variable was centered as recommended, this describes the average person's score on the outcome on the average occasion.

MLM provides a variance to address the second type of question, the extent to which the explanatory variable predicts changes in the outcome variable differently for different people. This variance is just like any other variance that describes how much a quantity varies across people. The quantity in this case is the coefficient that describes the relationship between the explanatory variable and the outcome. The novel, exciting, and sometimes daunting concept is that a finding (a relationship or a coefficient) can vary across people, just like any other quantity can. Normally, personality psychologists conceive of personality as consisting of differences in levels on variables; this quantity describes personality as consisting of differences in relationships between variables. That is, the psychological process relating the outcome to the explanatory variable can be different for different people. This variance is the statistic that describes whether the general principle relating two variables does not apply to some people, does apply to most people but in different strengths, or is even completely reversed for some people.

This variance is listed in the "UN (2,2)" row of the "estimates of covariance parameters" table (if there are more predictors, each subsequent predictor's variance is found in a row that has a number higher than 2 repeated in the parentheses, for example, "UN (3,3)"). The estimate is the variance—it is usually best to take the square root of this number (which will make it much larger if the variance is less than 1) and turn it into the more interpretable standard deviation (*SD*). The larger the *SD*, the more individuals differ in the process (i.e., in the direction, presence, or magnitude of the process). These numbers often appear to be very small even when they represent large differences between individuals in the strength or direction of the process. The way to evaluate their magnitude is to compare them to the value of the beta they modify. Most such betas are also less than 1, so *SD*'s on them will be smaller than that. If the distribution is relatively normal, approximately two-thirds of the participants will have a process coefficient within plus or minus 1 *SD* of the typical individual's process coefficient. Thus, an *SD* of even 0.20 or so on a typical beta of 0.30 describes substantial variation in the strength of the process across individuals.

The rest of the row provides a significance test on the variance; if the variance is significant, it means that individuals differ from each more than would be expected by chance sampling of occasions. This implies at least three overlapping interpretations. First, the process differs in its weight in the psychological functioning of individuals, and perhaps does not constitute the psychological functioning of some individuals at all. Second, the strength and relevance of this particular explanatory variable in this process differs across individuals. Third, differences between individuals in processes are reliable individual difference variables that describe part of the individual's personality (Fleeson, 2007; Mischel & Shoda, 1995). Note that a nonsignificant variance does not mean that individuals do not differ from each other—it means only that this particular study did not provide evidence to allow concluding that they differ reliably. Note also that the accuracy of this significance test is still being improved (Raudenbush & Bryk, 2002).

I suggest the following guidelines for reporting the results of MLM. Because MLM is a relatively new technique, researchers have often included equations in their articles. However, as the technique becomes more common, this practice should fade and become no more common than writing out equations for regressions is now. Second, tables are recommended for presenting the key results, but it is usually best to include only the important information; including only the important information enhances communication and interpretation of the results. Again because MLM is new, there is a temptation to overreport information, cluttering the tables and reducing the reader's comprehension. Rather, report the process coefficient and its significance for the average individual, and the *SD* across individuals in the process coefficient and the *SD*'s significance, for each explanatory variable. If the intercept is reported, it should follow the same format, with its value for the average individual and the standard deviation (*SD*) across individuals (the standard deviation is the square root of the variance estimate in the "UN(1,1)" row). Finally, when reporting results, use similar terms to those used in this chapter rather than the Greek letters or technical names for the coefficients. For example, report the "typical individual's relationship," rather than γ_{10}.

Case Example: Extraversion and Positive Affect

Most work on the relationship between extraversion and positive affect has been between persons, leaving it unclear whether this work translated into a within-persons process, such that individuals can become happier by acting more extraverted. In fact, there is reason to believe it may not translate; for example, introverted individuals may not derive happiness from the same activities as do extraverts, and even extraverts' happiness may not come from how they are acting. Fleeson and colleagues (2002) conducted a within-person process study to address this question directly. In an experience-sampling study, participants reported how extraverted they were acting (the explanatory variable) and how much positive affect they were experiencing (the outcome) every few hours.

This section provides a walk-through of the steps in MLM for these data. In order to facilitate the reader's following along, artificial data based on a real dataset (10 participants and 8–20 occasions per participant) are provided in Appendix 30.1 (the data can also be e-mailed upon request to Fleesonw@wfu.edu). Note that each row corresponds to one occasion and shows the extraversion and positive affect reported on that occasion, as well as an arbitrary identification number of the participant to which the occasion belongs. Each participant has multiple rows, one per occasion (the data are shown in multiple columns to save space, but in SPSS each row should have three entries). Note that a typical study would have more occasions and more participants.

Because state extraversion varied both between and within participants, and because the interest is in the correlates of within-person variation in state extraversion, the first step is to center extraversion on each person's mean. This allows the results to be interpreted as describing within-person processes only. Open the "aggregate . . . " dialog in the data menu, move ID to the "break variable" box, extra to the "summaries of variables" box, and make sure the "add aggregated variables to active dataset" box is checked. After clicking "ok," check that extra_mean is included in the data file and that all occasions for a given participant have the same value. Finally, subtract extra_mean from extra to create extra_centered.

In the MLM analysis, the effects of extraversion were assumed to be both fixed (part of the effect was common to all participants) and random (part of the effect of acting extraverted was allowed to differ across individuals). The same was assumed for the intercept (it was assumed that there was a common average level of positive affect for all participants and that each participant had unique factors that adjusted his or her own particular average level of positive affect). To run the MLM analysis, open the mixed models-linear dialog in the analyze menu. Move ID to the "subjects" box and click "continue"; move positiveaffect to the "dependent variable" box and extra_centered to the "covariate" box. In the "fixed . . . " dialog, add extra_centered to the "model" box and make sure that "include intercept" is checked. This step instructs the analysis to generate both the average level of positive affect (PA) for the typical participant (by clicking "include intercept") and the association between extraversion and positive affect for the typical participant (by adding extra_centered to the model). In the "random . . . " dialog, choose unstructured for the covariance type, add extra_centered to the "model" box, and check "include intercept"; this step allows both the average level of positive affect and the association between extraversion and positive affect to differ by participant. In addition, move ID to the combinations box to instruct the analysis to group the cases by participant. In the "statistics . . . " dialog, check "parameter estimates" and "tests for covariance parameters." Table 30.1 shows the syntax generated by clicking "paste" for a correctly entered model, including centering. If problems are encountered, generate syntax by clicking the "paste" button and compare it in detail with this example.

Running the MLM analysis revealed the output shown in Table 30.2 and the results shown in Table 30.3. The fixed effects table in the output shows the results for the average participant. The intercept row describes the average person's positive affect on those occasions when extra_centered had a score of 0. Because

TABLE 30.2. Selected Output Generated by Running the Syntax in Table 30.1 on the Sample Data

Estimates of Fixed Effects(a)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	4.037936	.328432	8.961	12.295	.000	3.294480	4.781391
extra_centered	.603568	.096111	7.192	6.280	.000	.377527	.829609

a Dependent Variable: positiveaffect.

Estimates of Covariance Parameters(a)

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		.614374	.079300	7.747	.000	.477050	.791227
Intercept + extra_centered [subject = ID]	UN (1,1)	1.030863	.509549	2.023	.043	.391253	2.716091
	UN (2,1)	–.101249	.123736	–.818	.413	–.343767	.141269
	UN (2,2)	.038032	.044154	.861	.389	.003908	.370137

a Dependent Variable: positiveaffect.

extraversion was centered, this means that the average person's score on positive affect on an average day was 4.04. The row for extra_centered addresses the first type of question, the degree to which extraversion predicts changes in positive affect for the average or typical person. The coefficient, 0.60, is just like an unstandardized coefficient from a regression, with magnitude and direction to indicate whether and how strongly extraversion predicted changes in positive affect. Because the coefficient is unstandardized, it means that, for the typical or average individual, each point he or she increased his or her momentary level of extraversion was associated with a little more than a half point increase in his or her level of positive affect (both extraversion and positive affect were on 1–7 scales). This result was significant, meaning that the within-person association between extraversion and positive affect is greater than expected by chance and that an ongoing process describing variation in positive affect has been identified.

Thus, this is a case in which the between-persons approach and the within-person approach produced answers in the same direction but different magnitude. Specifically, extraversion and positive affect are positively related both across people and across occasions within each person. However, the relationship is stronger within people than it is across people. This result is surprising, because this is a case in which the inference from the between-persons relationship to the within-persons relationship is not as plausible, and it is easily imaginable that the relationship would not hold up within people (it is easy to imagine that acting extraverted would not have the same impact on happiness that actually being an extravert would have). However, not only does this relationship describe within-person variation, it does so even more strongly than it does between people. Thus, this result demonstrates the need to examine within-person processes directly, because it shows that the result may be different

TABLE 30.3. The Within-Person Process Relating Extraversion to Positive Affect: Results from a Multilevel Model on a Small Subset of Data

Process	Typical individual	Standard deviation across individuals
Association of extraversion to PA	.60**	.19
Average PA	4.04**	1.02*

Note. The outcome is positive affect (PA), and the association of extraversion to PA for the typical individual means that every point the typical individual changes in extraversion across occasions is associated with a .60 increase in positive affect. In the full sample, the *SD* across individuals in this association was significant, $p < .01$, meaning that individuals differed reliability in the association of extraversion to PA.
* $p < .05$; ** $p < .001$.

than expected on the basis of the between-subjects research.

The next set of output, in the "covariance parameters" table, concerns how applicable these average effects are to each of the individuals and how much individuals differed from each other in these effects. The "estimate" column provides variances (and covariances) across participants in each of the fixed effects described above. The variance across individuals in average positive affect is printed in the row for UN(1,1)—the intercept is parameter number 1, and 1,1 means that the estimate is its covariance with itself, that is, its variance. The variance of 1.03 means that individuals differ quite a bit in how much positive affect they experience on an average day. Taking the square root of the variance to create a standard deviation and assuming a relatively normal distribution leads to the conclusion that about two-thirds of the participants have an average positive affect between 3.02 and 5.06 and about one-third have an average positive affect outside that range. This variance is significant, meaning that these participants differ from each other more than would be expected from chance.

The variance listed in the UN(2,2) row indicates how much individuals differed from each other in the coefficient relating positive affect to extraversion. This variance is exciting because it indicates whether and to what extent different individuals are described by different psychological processes. In this case, the question is whether the psychological process that relates positive affect to extraversion is different for different people. Because this estimate is a variance, it is usually best to take its square root and turn it into the more interpretable *SD*, 0.19 in this case. Assuming a relatively normal distribution, approximately two-thirds of the participants have an association between extraversion and positive affect between 0.41 and 0.79. The significance test on this variance shows that this difference was not greater than expected by chance, and so this reduced dataset did not produce evidence that individuals differ in this process more than would be expected by chance (e.g., due to more than the particular occasions sampled). However, in the actual data (Fleeson et al., 2002) the *SD* was 0.14 and significant, $p < 0.01$, meaning that individuals differed from each other in the process relating positive affect to extraversion. Some individuals experience a very strong positive affect boost at those times when they act extraverted, and some individuals experience only a moderate boost. These differences are not large, but because they are significant they reflect reliable aspects of these individuals' personalities. That is, it is possible to generalize from the particular occasions sampled for each individual to enduring differences between these individuals in their psychological processes regarding extraversion and positive affect.

Table 30.3 shows a way to report these results intended to maximize clarity, communication, and interpretation without cost to accuracy. The column for the typical individual shows the fixed effects that describe the process on average and whether the fixed effects differ significantly from zero. The column for the SD shows the random effects that describe how much individuals differed from each other and whether those differences are greater than expected from chance.

Complexities

In the interest of clarity, the preceding discussion described a relatively straightforward use of MLM. This section briefly extends MLM in ways that are more complex.

Using OLS Rather than MLM

It is possible to address almost the same hypotheses with the ordinary least squares (OLS) regression that most personality researchers use regularly and routinely. This would be accomplished by conducting a regression separately for each participant, predicting the outcome from the explanatory variable. Then, each participant's beta would be entered into a new data file. The average of these betas provides the average or typical individual's process coefficient (which can be tested against zero with a one-sample *t*-test). The *SD* across these coefficients indicates how much individuals differ from each other in the process. Using OLS in this method is familiar to researchers and provides an intuitive grasp of what MLM does. However, MLM has at least three advantages over OLS. First, MLM is more convenient, because the OLS procedure requires entering the betas from the regressions in a new data file. Second, MLM weights participants by the reliability of their coefficients, so more reliable data contribute more to the conclusion. OLS weights participants equally. Third and most important, MLM provides a significance test on the variability across participants' process coefficients. This is the key issue for personality psychologists, because it tests whether individual differences in the processes are greater than expected from chance and, rather, due to different psychological processes operating for different individuals (Fleeson, in press; Magnusson, 1999; Mischel & Shoda, 1995).

Predicting Individual Differences in the Process

A very exciting possibility of MLM is the opportunity to explain and predict individual differences in processes. Recall that MLM provides a coefficient describing the typical individual's process and a variance describing how much individuals differ from each other in that coefficient (in that process). A significant variance in the coefficient, establishing that individuals differ reliably in the process, is a very important first step because it suggests that the nature of personality includes individual differences in how processes operate. The next step is to explain why there are differences in those processes. One way to do this is to invoke other personality variables. For example, extraversion might predict the strength of the within-person process relating acting extraverted to happiness.

To test this kind of hypothesis, in which a personality variable is proposed to be related to individual differences in the process, the researcher adds the personality variable as a column in the dataset, with each participant's value on the personality variable repeated in each row for that participant. The personality variable should be centered across participants (this can be accomplished by subtracting the mean of the personality variable from the personality variable). This variable is called a person-level or level-2 variable because it does not vary within a person, only across persons. In the mixed models command, first add the personality variable as another covariate. Then, in the "fixed . . . " dialog, highlight both predictors, make sure "factorial" is selected, click "add," and run the analysis.

A significant interaction term in the "estimates of fixed effects" table means that the personality variable predicted individual differences in the process coefficient: A positive interaction means that individuals higher on the personality variable had a more positive relationship between the explanatory variable and the outcome, whereas a negative interaction means that the individuals lower on the personality variable had a more positive relationship between the explanatory variable and the outcome. The estimate for the explanatory variable in the fixed effects table has a new interpretation in this type of analysis—it is now the effect of the explanatory variable for participants with a personality variable score of 0 (the mean if it was centered). Finally, the estimate for the personality variable is its effect on individuals' average levels of the outcome. The significance for the variance in the process coefficient indicates whether there is evidence for remaining individual differences in the process beyond the variance explained by the personality variable.

In following this procedure, Fleeson and colleagues (2002) found in fact that trait extraversion significantly predicted the coefficient relating acting extraverted to positive affect, but that the prediction was negative. That is, introverts experienced a larger positive affect boost when acting extraverted than did extraverts.

Additional Explanatory Variables

More than one explanatory variable can be tested simultaneously. This works similarly to adding additional explanatory variables in multiple regression. Add the additional variables to the covariates box, to the fixed effects box, and to the random effects box. To recreate a typical multiple regression, do not add them as factorials but only as main effects. The resulting coefficients can be interpreted as in a multiple regression, revealing the effect of each explanatory variable while controlling for the other explanatory variables. However, because MLM estimates variances and covariances among all coefficients and intercepts, these models become unstable with only a few variables. Caution is urged.

Causal Direction

Finally, it is possible to address causal direction between the explanatory variable and the outcome. This invokes a similar procedure to that used in between-persons analyses: The explanatory variable and the outcome at time 1 are used to predict the outcome at time 2. However, times 1 and 2 are usually successive occasions only hours apart. If there is a significant relationship between the explanatory variable and the time-2 outcome, this cannot be due to the outcome causing the explanatory variable, because the future cannot affect the past and because the explanatory variable was made independent of the earlier time-1 outcome in the model. However, two difficulties arise in using this method. First, because of the rapid variations in the explanatory variable and in the outcome, having the correct time lag between time 1 and time 2 is crucial and may be very difficult. For example, if the effect of the explanatory variable on the outcome is immediate and short-lived, a lag of a few hours between occasions may not pick up that effect. Second, studies of this sort typically have missing occasions. Missing occasions break the chain of lags between occasions, and so reduce power. Not only does this create missing data, the analysis will erroneously skip ahead to the next occasion to fill in the lag unless the researcher prepares the data ahead of time.

Future Directions

There are at least five future directions for this method. First, software improvements are continually improving the convenience of this analytic tool. What is needed most is software simultaneously integrated with other software so that data transformations are easy to do within the same program, and targeted at studying within-person processes (such as HLM) rather than written broadly to include all possible uses of mixed models (such as SPSS). Also needed are labels and instructions that are intuitively compelling to researchers familiar with regression and other common statistical techniques.

Second, it is critical to be able to include multiple predictors and their interactions in these models. Personality psychologists typically need to control many possible third variables and desire to test complex process models. Current packages become unstable quickly under such conditions. Future software and statistical advances should be directed at solving these problems.

Third, a high priority is to improve the reliability of or replace the significance test on the variances across individuals (i.e., determining whether individuals differ reliably in the direction, presence, or magnitude of the process). The standard errors as currently calculated may be inaccurate with low *N*'s (although most experience sampling methods (ESM) studies may have sufficient *N*'s). More work is needed to establish a universally accepted and accurate test of the significance of individual differences in the process (Raudenbush & Bryk, 2002).

Fourth, factor analyses may benefit from a similar treatment. Currently, factor analyses primarily are applied to all individuals equally. What would be valuable would be the ability to conduct factor analyses for each individual uniquely at the same time that the analyses are compared across individuals. This concept, called *p*-technique factor analysis by Cattell (1952), is currently possible but is also less convenient than it could be. *P*-technique is very promising for discovering factor structures that differ from individual to individual, and may possibly be integrated into the MLM framework. Personality psychology may make significant advances with *p*-technique if it could be more convenient to use.

A final future direction is for process to become normal personality research. The *Journal of Personality and Social Psychology* covers two personality topics: personality processes and individual differences. In the past, attention to individual differences has been predominant, but recently attention to process has grown. Understanding process is critical to explaining how personality works and to identifying the mechanisms underlying personality. In addition, detailing the steps in a process often locates opportunities, in the form of modifiable actions or events, for altering or improving processes to result in improved mental health. Both between-persons and within-person approaches are critical in the study of process. Recent developments in software and equipment make within-person approaches more accessible, adding them to the set of personality psychologists' techniques and thereby reducing the need for inferences from between-person designs to within-person processes. These recent developments in software and equipment, combined with texts to guide their use and interpretation, may make for a process revolution in personality psychology.

Acknowledgments

Preparation of this chapter was supported by National Institute of Mental Health Grant No. R01 MH70571. I would like to thank Mike Furr and Eric Stone for discussions about change, Sarah Ross and

Martin Lynch for comments on earlier drafts, and Faye Reece for assistance in preparing the manuscript.

Notes

1. Many important processes happen between individuals; those processes should be studied between individuals. The point being made here is that the processes that are theorized to occur within individuals are often studied by comparing between individuals rather than by comparing within individuals over time.
2. However, within-person approaches are vulnerable to a kind of reverse ecological fallacy, namely the inference that a process demonstrated to operate in the short term operates similarly in the long term. This inference is not always valid.

Recommended Readings

Aldwin, C. M., & Revenson, T. A. (1987). Does coping help? A reexamination of the relation between coping and mental health. *Journal of Personality and Social Psychology*, *53*, 337–348.

Cronbach, L. J., & Furby, L. (1970). How should we measure change—Or "should" we? *Psychological Bulletin*, *74*, 68–80.

Darlington, R. B. (1990). *Regression and linear models*. New York: McGraw-Hill.

Fleeson, W., Malanos, A., & Achille, N. (2002). An intra-individual process approach to the relationship between extraversion and positive affect: Is acting extraverted as "good" as being extraverted? *Journal of Personality and Social Psychology*, *83*, 1409–1422.

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Lamiell, J. T. (1997). Individuals and the differences between them. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 117–141). San Diego, CA: Academic Press.

Singer, J. D., & Willett, J. B. (2002). *Applied longitudinal data analysis*. New York: Oxford University Press.

References

Aldwin, C. M., & Revenson, T. A. (1987). Does coping help? A reexamination of the relation between coping and mental health. *Journal of Personality and Social Psychology*, *53*, 337–348.

Allport, G. W. (1937). *Personality: A psychological interpretation*. Oxford, UK: Holt.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*, 579–616.

Carver, C. S., & Scheier, M. F. (1994). Situational coping and coping dispositions in a stressful transaction. *Journal of Personality and Social Psychology*, *66*, 184–195.

Cattell, R. B. (1952). The three basic factor-analytic research designs—their interrelations and derivatives. *Psychological Bulletin*, *49*, 499–520.

Cattell, R. B. (1966). *Handbook of multivariate experimental psychology*. Chicago: Rand McNally.

Cervone, D. (2004). The architecture of personality. *Psychological Review*, *111*, 183–204.

Cervone, D., & Mischel, W. (2002). Personality science. In D. Cervone & W. Mischel (Eds.), *Advances in personality science* (pp. 1–26). New York: Guilford Press.

Costa, P. T., Jr., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, *38*, 668–678.

Cronbach, L. J., & Furby, L. (1970). How should we measure change—Or "should" we? *Psychological Bulletin*, *74*, 68–80.

Darlington, R. B. (1990). *Regression and linear models*. New York: McGraw-Hill.

Darlington, R. B., & Smulders, T. V. (2001). Problems with residual analysis. *Animal Behaviour*, *62*, 599–602.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, *22*, 491–569.

Epstein, S. (1983). A research paradigm for the study of personality and emotions. In R. Dienstbier (Series Ed.) & M. M. Page (Vol. Ed.), *Nebraska Symposium on Motivation: Vol. 30. Personality: Current theory and research* (pp. 91–154). Lincoln: University of Nebraska Press.

Fleeson, W. (2004). Moving personality beyond the person–situation debate: The challenge and the opportunity of within-person variability. *Current Directions*, *13*, 83–87.

Fleeson, W. (2007). Using experience sampling and multilevel linear modeling to study person–situation interactionist approaches to positive psychology. In A. D. Ong & M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 501–514). New York: Oxford University Press.

Fleeson, W. (in press). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality*.

Fleeson, W., Malanos, A., & Achille, N. (2002). An intra-individual process approach to the relationship between extraversion and positive affect: Is acting extraverted as "good" as being extraverted? *Journal of Personality and Social Psychology*, *83*, 1409–1422.

Funder, D. C. (2001). Personality. *Annual Review of Psychology*, *52*, 197–221.

Furr, R. M. (2005). *The analysis of change across two time points: A comparison of difference scores,* *residualized change scores, and multiple regression*. Winston-Salem, NC: Wake Forest University.

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Lamiell, J. T. (1997). Individuals and the differences between them. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 117–141). San Diego, CA: Academic Press.

Larsen, R. J. (1989). A process approach to personality psychology: Utilizing time as a facet of data. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 177–193). New York: Springer-Verlag.

Lazarus, R. S. (2000). Toward better research on stress and coping. *American Psychologiset, 55*, 665–673.

Lucas, R. E., & Baird, B. M. (2004). Extraversion and emotional reactivity. *Journal of Personality and Social Psychology, 86*, 473–485.

Magnusson, D. (1999). Holistic interactionism: A perspective for research on personality development. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 219–248). New York: Guilford Press.

Marco, C. A., Neale, J. M., Schwartz, J. E., Shiffman, S., & Stone, A. A. (1999). Coping with daily events and short-term mood changes: An unexpected failure to observe effects of coping. *Journal of Consulting and Clinical Psychology, 67*, 755–764.

McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In O. P. John & L. A. Pervin (Eds.), *Handbook of personality: Theory and research*. (2nd ed., pp. 139–153). New York: Guilford Press.

McNiel, J. M., & Fleeson, W. (2006). The causal effects of extraversion on positive affect and neuroticism on negative affect: Manipulating state extraversion and state neuroticism in an experimental approach. *Journal of Research in Personality, 40*, 529–550.

Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology, 55*, 1–22.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246–268.

Mroczek, D. K., & Almeida, D. M. (2004). The effects of daily stress, personality, and age on daily negative affect. *Journal of Personality, 72*, 355–378.

Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 92–105). Washington, DC: American Psychological Association.

Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry, 4*, 122–130.

Puetz, T. W., O'Connor, P. J., & Dishman, R. K. (2006). Effects of chronic exercise on feelings of energy and fatigue: A quantitative synthesis. *Psychological Bulletin, 132*, 866–876.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed., Vol. 1). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). Hierarchical linear and non-linear modeling (Version 6.0) [Computer software]. Lincolnwood, IL: Scientific Software International.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*, 351–357.

Rogosa, D. (1995). Myths and methods: "Myths about longitudinal research" plus supplemental questions. In J. M. Gottman (Ed.), *The analysis of change* (pp. 3–66). Mahwah, NJ: LEA.

Rusting, C. L., & Larsen, R. J. (1998). Personality and cognitive processing of affective information. *Personality and Social Psychology Bulletin, 24*, 200–213.

Singer, J. D., & Willett, J. B. (2002). *Applied longitudinal data analysis*. New York: Oxford University Press.

Sroufe, L. A., Egeland, B., Carlson, E. A., & Collins, W. A. (2005). *The development of the person: The Minnesota Study of Risk and Adaptation from Birth to Adulthood*. New York: Guilford Press.

Staudinger, U. M., Fleeson, W., & Baltes, P. B. (1999). Predictors of subjective physical health and global well-being: Similarities and differences between the United States and Germany. *Journal of Personality and Social Psychology, 76*, 305–319.

Tennan, H., Affleck, G., Armeli, S., & Carney, M. A. (2000). A daily process approach to coping: Linking theory, research, and practice. *American Psychologist, 55*, 626–636.

Vansteelandt, K., & Van Mechelen, I. V. (2004). The personality triad in balance: Multidimensional individual differences in situation–behavior profiles. *Journal of Research in Personality, 38*, 367–393.

APPENDIX 30.1. SAMPLE DATA

ID	Extra	Positive Affect
1.00	2.00	3.25
1.00	3.50	3.25
1.00	2.00	2.25
1.00	3.25	3.00
1.00	2.50	2.50
1.00	2.00	2.25
1.00	3.50	3.25
1.00	3.50	3.25
2.00	4.50	5.25
2.00	2.00	2.50
2.00	3.50	4.50
2.00	3.75	4.00
2.00	2.50	2.75
2.00	2.00	1.75
2.00	4.25	3.25
2.00	4.75	4.50
2.00	4.50	4.50
2.00	5.75	5.00
2.00	2.75	3.50
2.00	3.00	2.00
2.00	2.00	2.50
2.00	5.75	5.75
2.00	4.25	2.75
2.00	2.25	2.25
2.00	4.50	5.50
9.00	2.75	3.75
9.00	5.75	3.00
9.00	5.00	5.75
9.00	4.50	4.75
9.00	4.00	4.50
9.00	3.25	3.00
9.00	4.25	2.50
9.00	3.50	3.25
9.00	3.00	3.50
9.00	1.75	4.50
9.00	3.00	2.50
9.00	3.50	3.75
18.00	3.25	4.50
18.00	2.75	4.75
18.00	3.75	5.50
18.00	2.75	3.75
18.00	3.00	4.00
18.00	6.00	5.50
18.00	4.50	4.50
18.00	3.00	4.50
18.00	3.25	5.50
18.00	3.25	4.50
18.00	2.75	4.75
18.00	4.00	5.75
18.00	3.25	5.50
18.00	3.50	4.75
18.00	2.75	5.00
22.00	5.25	6.75
22.00	4.75	6.50
22.00	5.50	6.50
22.00	5.50	7.00
22.00	5.00	5.75
22.00	4.50	6.00
22.00	4.75	6.75
22.00	4.75	6.50
22.00	6.00	6.50
22.00	4.50	5.00
22.00	3.75	5.75
22.00	6.25	6.50
26.00	3.00	2.25
26.00	3.00	3.50
26.00	2.25	3.25
26.00	4.75	3.50
26.00	5.00	3.50
26.00	4.00	2.50
26.00	1.25	1.25
26.00	2.75	2.75
26.00	1.50	1.25
26.00	2.00	1.75
26.00	3.00	1.75
26.00	2.50	3.00
26.00	2.75	3.00
26.00	4.50	4.50
26.00	3.00	2.50
26.00	2.00	4.50
26.00	2.50	4.25
26.00	4.00	5.00
26.00	3.25	2.75
26.00	1.25	1.00
31.00	6.00	4.00
31.00	3.00	3.25
31.00	5.00	4.75
31.00	1.00	1.75
31.00	5.50	4.75
31.00	5.75	5.00
31.00	5.75	5.75
31.00	4.75	5.00
31.00	5.00	5.50
31.00	4.75	4.75
33.00	4.50	4.50
33.00	3.00	2.75
33.00	5.50	5.25
33.00	2.50	2.00
33.00	4.00	1.25
33.00	5.00	6.50
33.00	5.00	3.00
33.00	3.25	2.50
33.00	4.25	3.00
33.00	5.50	4.50
35.00	4.75	4.25
35.00	4.75	4.25
35.00	4.00	4.25
35.00	5.25	4.25
35.00	5.00	4.25
35.00	7.00	6.25
35.00	6.75	5.75
35.00	7.00	5.50
35.00	6.00	3.25
35.00	5.50	5.00
35.00	6.00	5.00
35.00	6.00	4.25
35.00	5.50	5.25
35.00	5.00	5.50
35.00	5.25	5.33
35.00	6.00	5.25
35.00	5.25	3.00
35.00	5.50	2.75
37.00	3.00	3.25
37.00	3.75	4.00
37.00	3.25	3.50
37.00	2.75	3.00
37.00	2.33	2.50
37.00	3.00	3.50
37.00	4.00	4.00
37.00	2.50	2.50
37.00	3.00	2.75
37.00	3.50	3.00
37.00	2.75	3.50
37.00	3.50	2.75
37.00	3.25	3.50
37.00	3.75	4.00
37.00	4.00	4.75
37.00	4.00	4.00
37.00	4.00	4.50

Guilford Publications
72 Spring Street
New York, NY 10012
212-431-9800
800-365-7006
www.guilford.com