MEASUREMENT, *5*(4), 217–235

FOCUS ARTICLE

Idiographic Filters for Psychological Constructs*

John R. Nesselroade
University of Virginia, Charlottesville

Denis Gerstorf
Pennsylvania State University, University Park, Pennsylvania

Sam A. Hardy
Brigham Young University, Provo, Utah

Nilam Ram
Pennsylvania State University, University Park, Pennsylvania

Ideally, the unit of analysis in psychology is the individual. However, many psychological methods do not cope well, either at the level of construct definition or at the level of measurement, with individuality in behavior. There is little leeway for constructs to be both idiosyncratically tailored to the individual, and still identified as having the same core meaning between individuals. Generally accepted ways to aggregate data in the course of empirically studying relationships simply ignore idiosyncrasy—but do so at some as yet undetermined costs. We identify, and illustrate with empirical data, the main features of an individually-oriented approach to general construct definition and measurement that seem to challenge the established

Correspondence should be addressed to John R. Nesselroade, Psychology Department, University of Virginia, PO Box 400400, Charlottesville, VA 22904-4400. E-mail: jrn8z@virginia.edu

*Look for additional commentaries and a rejoinder to this article in an upcoming issue of *Measurement*.

concepts and norms of "good" measurement practice in behavioral research yet, at the same time, can be reconciled with them.

Key words: nomothetic, p-technique factor analysis, factorial invariance, idiosyncratic measurement

INTRODUCTION

A fundamentally important question for behavioral science is what to do about *idiosyncrasy* while attempting to establish *general* lawfulness. Simply diluting idiosyncrasy by "averaging it out" over many cases is neither satisfying nor ultimately productive. We propose an alternative—explicitly filtering idiosyncrasy out of nomothetic relationships, present an approach for doing so, and illustrate it empirically.

Our proposal, which emphasizes the individual as the unit of analysis (see e.g., Molenaar, 2004), involves identifying and filtering out irrelevant[1] interindividual differences to focus on interindividual similarities. After having seemed to spurn the traditional measurement concept of invariance we re-employ it, but at one step removed in abstraction from current measurement practice. Several matters that set the stage for our proposal include (a) the historical idiographic/nomothetic debate, (b) the scientific concept of invariance, (c) data aggregation practices, (d) concepts of intra-individual variation, and (e) the respective roles of inter-individual differences and similarities. We examine each in turn.

The Idiographic/Nomothetic Debate

Idiographic concerns emphasize the uniqueness of the individual, whereas nomothetic concerns emphasize generality in behavioral lawfulness (e.g., Allport, 1937; Bem & Allen, 1974; Lamiell, 1981, 1988; Molenaar, 2004; Rosenzweig, 1958, 1986; van Kampen, 2000; Zevon & Tellegen, 1982). These two long-established, very general perspectives imply quite different methods and emphases. Some proposed attempts at rapprochement between the idiographic and nomothetic traditions have implicated a key role for p-technique factor analysis (Nesselroade & Ford, 1985; Zevon & Tellegen, 1982), which involves applying the factor model to one person's multivariate time-series to identify patterns of covariation within the individual over time (Cattell, 1952, 1963). Bereiter (1963) labeled p-technique the logical way to study relationships among variables. With regard to idiographic/nomothetic rapprochement, the key idea

[1]What is irrelevant depends on one's purpose and purposes obviously differ from instance to instance. Some of the variation we will propose filtering out as irrelevant might well be the substance of another's investigation.

is that a more precise characterization of individual level behaviors (what p-technique does) can better inform the articulation of lawful relationships characterizing groups.

Some of the ways in which individuals differ from one another are better classified as differences in *kind* rather than in *amount* (Tucker, 1966). The dominant methodologies of differential psychology are more appropriate for the latter than for the former (Nesselroade, 2002, 2006). Qualitative differences mistaken for quantitative differences can seriously distort relationships and are a prescription for diluted nomothetic relationships. We argue for the merits of recognizing the idiosyncratic and then creating a "firewall," to use today's vernacular, between it and lawful, nomothetic relationships.

The Concept of Invariance

Invariant relationships are a primary focus of science (Keyser, 1956). Ascertaining which properties remain invariant under which transformations both enables the formulation of key relationships and establishes their generality. For example, although it may seem somewhat counter-intuitive, the relationship between distance fallen (s) by an object and time (t), $s = 1/2\ g \cdot t^{2,}$, is invariant for objects having different mass.

In psychological measurement, factorial invariance concepts have evolved from an early emphasis on the factor-loading pattern to include intercepts, uniquenesses, and sometimes the factor intercorrelation matrices (see e.g., Meredith, 1993; Millsap & Meredith, 2006). Interest in the factor loadings—the relationships between manifest variables and latent variables—has been fundamental, particularly in research on abilities, personality, and other individual differences domains and, more currently, in the measurement models of structural equation modeling (SEM) (Horn, McArdle, & Mason, 1983; Meredith, 1993, Millsap & Meredith, 2006; Thurstone, 1947). Factor loadings are key to testing abstract relationships against empirical data.[2] When loading patterns are invariant across comparisons (e.g., between genders or over time), the nature of the variables is typically interpreted to be the same across those comparisons, thus legitimizing the comparisons.

Substantive interests often center on the relationships among the theoretical constructs (elements of a theory) and perhaps testing whether or not these structural relationships are invariant over different subpopulations, at different ages, etc. However, the usual basis for assurance that we have indeed "captured" the same theoretical constructs in different subgroups of data and are thus making a meaningful comparison of invariant structural relationships, is a demonstration that the factor loadings that link the constructs to the observables are invariant over the different subgroups.

[2]For a somewhat contrary view, see Nunnally (1973).

The loci of invariant relationships are clear in the example just given and in SEM where the invariance concerns the linkage between observable and unobservable variables (e.g., the measurement model, factor loadings). A second level of invariance has to do with the linkages among unobservable variables or hypothetical constructs (e.g., the structural model, relationships among factors). The first level of invariance, because it is linked to observables, has become the tool by which we examine and test the second level. This essential distinction is illustrated in Figure 1.

Consider two different situations that can come into play. First, the measures may not explicitly be the same. For example, suppose one is studying the relationships between parental rearing styles and later personality characteristics. The personality measures appropriate for age 6 and those for age 15 may be very different even though one is focusing on ostensibly the same personality constructs e.g., dominance and sociability. Second, the measured or observed variables may ostensibly be the same but do not actually represent the same qualities in different situations. Kagan (1980) has addressed these and related issues in his discussion of continuity and discontinuity. Such situations render a simple interpretation of the concept of invariance inapplicable and jeopardize data aggregation and meaningful comparisons.

Researchers have sought ways of usefully turning qualitative differences at one level into quantitative ones at a different level of abstraction. For example, instead of enumerating the presence or absence of specific environmental risk factors as predictors of later development, Sameroff, Peck, and Eccles (2004) simply counted the number of such risk factors characterizing each participant.

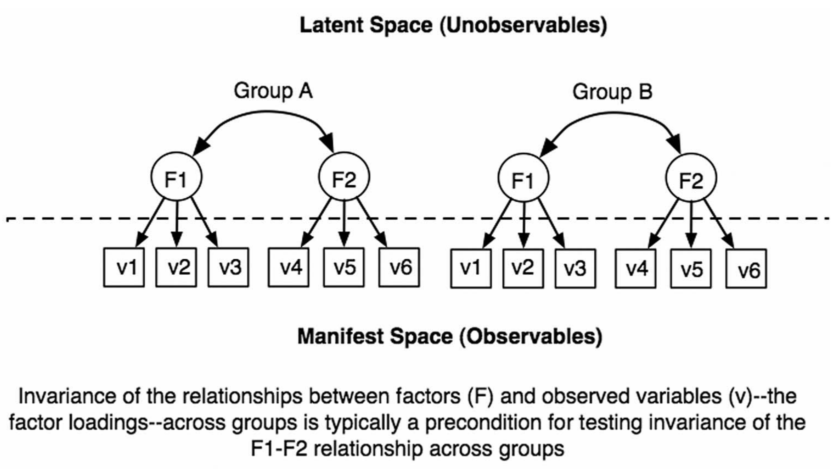

FIGURE 1 Two levels of invariance that are involved in testing relationships among constructs.

McArdle (2006) proposed handling the problem in this way. Instead of asking, "How well do you play tennis?" ask "How well do you play your favorite sport?" This allows the idiosyncrasy of some people liking to play tennis, others preferring to play golf, etc., to be recognized and set aside while getting answers that are commensurate with each other in terms of a more abstract quality, namely self-reported athletic performance. One can think of such operations less as mixing apples and oranges and more as inspecting fruit.

Summing up, invariance concepts are crucial in science and important strides in defining them in psychometric terms have been made (see e.g., Meredith, 1993; Millsap & Meredith, 2006). But just how the concepts can be most effectively utilized is a question that seems not yet fully answered. Our objective is to identify a basis for arguing that abstract constructs are general even when the observable manifestations are different from one individual to another. But we wish to formalize such thinking using legitimate concepts of invariance. Doing so would accommodate both idiographic and nomothetic perspectives but appears to negate the literal view of factorial invariance that has been fundamental to justifying the interpretation that the abstract construct is indeed the same. We provide an alternative not by abandoning the concept of invariance but by supplying more flexibility to the way it is applied.

Data Aggregation

Researchers have questioned the blind aggregation of data across persons to identify individual level processes (e.g., Lamiell, 1981, 1988; Nesselroade & Molenaar, 1999). Aggregates such as the mean are criticized because they may not correspond to a single instance in a given sample (e.g., no family has 2.3 children). Similar objections can be voiced regarding variances and covariance structures (e.g., Daly, Bath, & Nesselroade, 1974).

We believe that individual level factor analysis (e.g., p-technique) is a promising way to establish a meaningful basis for aggregating information across multiple individuals (Nesselroade & Ford, 1985; Zevon & Tellegen, 1982) but question the traditional role of factorial invariance in guiding the aggregation of multiple p-technique analyses. Below, we will present an easing of the traditional notion of factorial invariance that facilitates aggregation of information across participants and enhances the detection of nomothetic relationships while recognizing and isolating idiosyncrasy.

Intra-Individual Variability and Change

The past three decades have seen a great deal of interest develop in the study of intra-individual variability, a number of new analytic techniques by which to accomplish it, and the emergence of valuable conceptions of relatively short-term change such as states versus traits (Cattell & Scheier, 1961; Hamaker, Molenaar, & Nesselroade, 2007). Fundamentally, the tools of individual differences research have been applied to variation within the individual over time (e.g., multivariate time-series) in lieu of variation over individuals at one occasion of measurement (Nesselroade, 2002). Yet, it remains the case that the bulk of behavioral research tends to feature between-individual variation. Borsboom, Mellenbergh, and van Heerden (2003), for example, discussed relationships between within-individual and between-individual variation in relation to using psychological constructs and noted that although our theories are populated largely by within-individual concepts the related empirical work is usually based on between-individual information. They, too, point out that attempts to link within-individual processes to between-individual latent variables are increasing and show some promise of success.

Also stemming from concerns with individual-level process has been the development of advanced modeling techniques for analyzing intra-individual variability. Among these are dynamic factor analysis models that extend the power and versatility of p-technique factor analysis while maintaining its original purpose of studying how an individual varies over time (Browne & Nesselroade, 2005; Molenaar, 1985; Nesselroade, 2002). It is on this general line of approach that we rely to accommodate both idiographic and nomothetic concerns (see e.g., Molenaar, 2004; Nesselroade & Molenaar, 1999).

Inter-Individual Differences and Similarities

From a traditional, individual differences perspective, relationships of certain kinds (e.g., factor loadings, factor interrelationships, etc.) established on group data are assumed to hold for all persons, thus rendering individual differences in amounts of attributes meaningful. One of our concerns is the establishment of explicit invariant relationships at the individual level that will support a claim that the relationships hold equally well at the group level. We seek to minimize qualitative differences between individuals in order to aggregate quantitative information more justifiably. We propose to identify idiosyncrasy and set it aside (filter it out) in order to enhance the similarities of behavior patterns, thereby promoting relationships that have greater generality. Thus, our primary concern is not to evaluate quantitative differences among participants; we are trying neither to minimize nor maximize them. Rather, our purpose is to create the circumstances that ensure that statements regarding such quantitative differences are accurate.

Without some version of invariance—some fixed reference frame—there is no meaningful way to aggregate information from different participants into a nomothetic framework. Our proposal is straightforward. We are going to define invariance at the more abstract level of factor intercorrelations, allowing the factors to be defined by somewhat different patterns of observables from person to person, as necessary or desirable, thus integrating the idiographic and nomothetic.

Before spelling out the proposal in more detail, we ask: Are there precedents for defining the "same" abstract constructions on different observed variable substrates? We illustrate a rather simple example in Figure 2, which shows three subpopulations of geometric figures: P_1 = rectangles, P_2 = circles, and P_3 = triangles. Consider the abstraction we call Area (A). In the case of rectangles, A rests on the observable (measurable) manifestations of length (l) and width (w) and is computed as their product, $A = l \cdot w$. For circles, A is a function of the (measurable) radius and is computed as πr^2. For triangles, $A = 1/2\ bh$, where b is the base and h the altitude of the triangle. Now, suppose we introduce a second abstraction, Volume (V), by extending the members of the subpopulations along a third spatial dimension by some measurable amount s. Despite the subpopulation differences in the way A is computed, the abstractions A and V bear exactly the same relationship to each other across subpopulations i.e., $V = sA$ in all three subpopulations. Thus, A and V have the same meaning from one subpopulation to another but they do not rest on relationships to observables that are invariant across subpopulations.

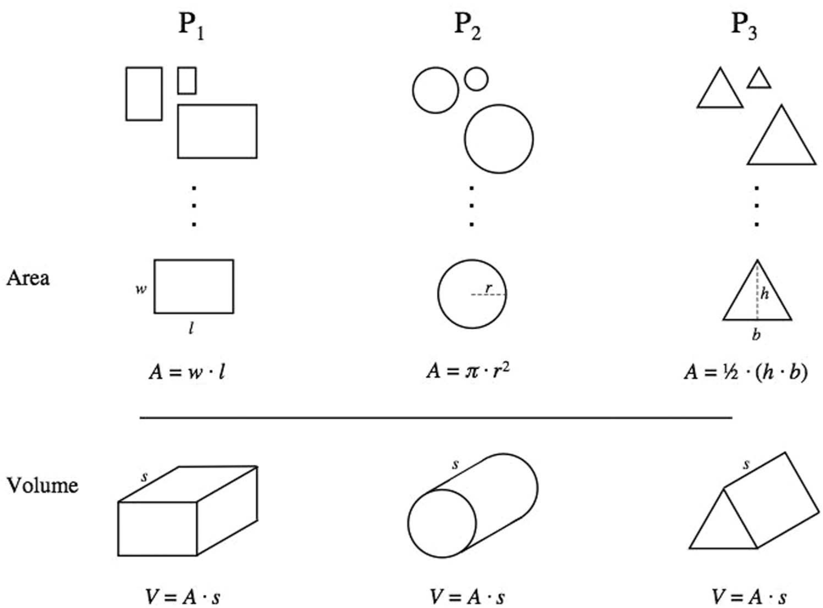

FIGURE 2 Three subpopulations illustrating how the "same" abstractions (e.g., area and volume) can rest on different observable variables.

Another example comes from the field of visual perception.[3] In rendering estimates of distance perception, judges can employ different mechanisms, depending upon what information is available in a given context, yet arrive at similarly accurate distance estimates. Thus, commensurability of abstract products arrived at by quite different mechanisms is a fundamental idea, some key implications of which will be explored in the next section.

P-TECHNIQUE FACTORS AS IDIOGRAPHIC FILTERS

Our proposal is as follows. Conduct individual level analyses, such as p-technique factor analyses, using theory-guided conceptions to the extent possible. Because of possibly idiosyncratic features of behavior, do not insist on invariant factor loadings across cases. Instead, allow for any necessary individuality in the manifest expression of underlying constructs. Individuality might be due to uniqueness of expression of the construct, peculiarities of the measurement process, or both. One's experiential history may uniquely "stamp" his or her pattern of overt behavior associated with constructs such as anxiety or extraversion or, even though the same physical measurement instruments are used, all individuals may not show variability on all variables, thus rendering those variables not factorable for that person. Either way, it would not be possible to demonstrate classical factorial invariance from one participant to another.

We propose to capitalize on the flexibility of p-technique analysis (or its extensions such as dynamic factor analysis) to define invariance in a different way that allows idiosyncrasy to creep into the factor loadings. We define invariance on the relationships among the constructs (the factors) and conduct the analyses accordingly. Factor A and Factor B may be identified in Person 1's and in Person 2's p-technique solutions but Person 1's Factor A and Person 2's Factor A may not have the same pattern of loadings on the observed variables. (Remember, we may not even be factoring the same subsets of variables in different cases if some variables are lost to a given person's data because of lack of intra-individual variability.) Similarly, Factor B in the two cases may evince somewhat different loading patterns. Presumably, however, there will be sufficient resemblance that the substantive labels given the factors seem reasonable.[4] But, and here is the key, we specify the

[3]We are grateful to our colleague, Dennis R. Proffitt, for this example. See (Proffitt, 2006) for detail.

[4]An anonymous reviewer pointed out in an earlier version of this manuscript that this statement is imprecise and ambiguous. We agree, but do not yet have a full answer to the criticism. The approach we are advocating has to rely on a strong enough theoretical basis concerning the phenomenon that there can at least be intersubjective agreement that the same phenomenon is being captured. One is reminded of the Supreme Court's discussion of obscenity in which Justice Stewart could not define pornography but claimed that he "knew it when he saw it." The situation is another firm reminder of the interdependence of theory and method.

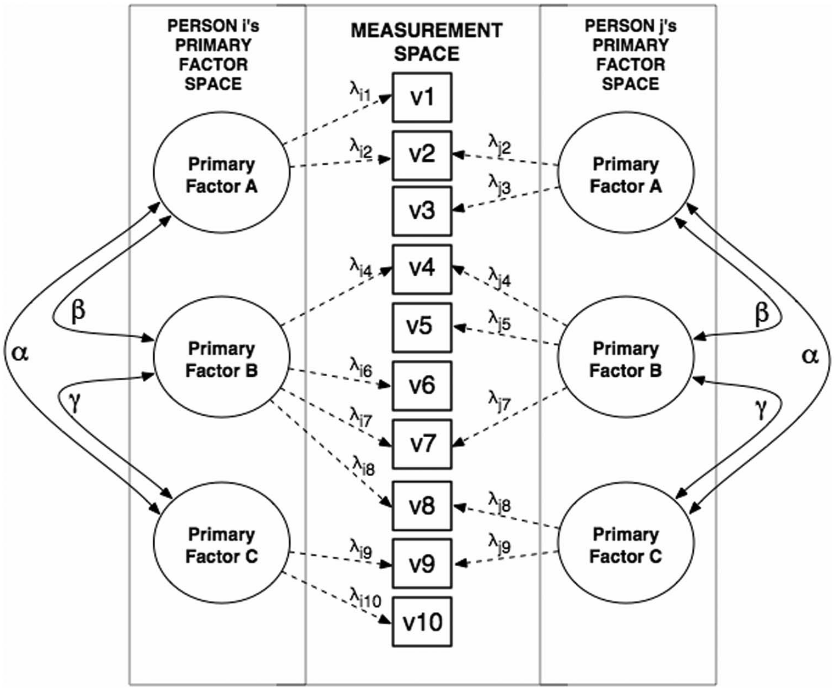

FIGURE 3 Graphical representation of factor solutions for two individuals with invariance at the level of factor intercorrelations rather than factor loadings.

correlation between Factor A and Factor B to be precisely the same for Person 1 as it is for Person 2. The general idea is illustrated graphically in Figure 3. The figure portrays hypothetical p-technique factor solutions for two individuals. The loading patterns exhibit idiosyncrasies in that individuals differ in which items (v1 to v10) define the respective primary factors A, B, and C, but the three factor intercorrelations, represented here as α, β, and γ are invariant across individuals.

There is some precedent in the literature for resting the interpretations of constructs on their intercorrelations as compared to their relationships to observables (Corballis & Traub, 1970; Livson, 1973). Although these examples represent the group-analysis tradition, they clearly signified the kind of non-invariant construct-to-manifest variable relationships we are considering.

AN EMPIRICAL DEMONSTRATION

To illustrate the concepts and ideas that we have been discussing and, as an empirical test of their plausibility, we report a reanalysis of data originally published by Lebo and Nesselroade (1978). The data consist of five multivariate time-series (p-technique data) representing the daily self-reports of women ($N = 5$) on 75 adjective rating scales of affect over more than 100 days. The adjective rating scales were culled from the affect literature as it existed in the early 1970s and represented Depression, Fatigue, Energy, Concentration, Social Affection, and Well-Being.

First, we eliminated no- and low-variance items from the individual data sets. This left between 24 and 33 items per participant to be analyzed further. Confirmatory p-technique factor analyses were conducted separately for each participant using the dynamic factor analysis program, DyFA (Browne & Zhang, 2005). The individual participant's factor loading patterns were specified using the affect literature as a guide (see Lebo & Nesselroade, 1978) but, of course, the target loading pattern for each individual was adjusted to the particular subset of variables comprising her data. The primary reason for these analyses was to get an initial grasp of the nature of both the factor loading patterns and the factor intercorrelations. These analyses indicated that three factors (Energy, Well-Being, and Social Affection) could reliably be identified for all five participants, and two more factors (Fatigue and Concentration) could be identified for four of the five participants. Of course, each loading pattern displayed a certain level of idiosyncrasy because of the differing (but overlapping) subsets of variables.

To provide more direct between-person comparisons, in a second set of confirmatory analyses (using Mplus, version 3; Muthén & Muthén, 1998), the five data sets were analyzed in one framework—a multi-group specification with each participant's multiple occasions of measurement constituting an independent group. This five-group specification contained the same variables and same factors for all five data sets. This was accomplished by inserting, where needed, variables with variances of near zero values, and factors with unit variance to populate each participant's data with the same numbers of variables and factors.

To accommodate the varying degrees of incompleteness under the missing-at-random assumption (McArdle, 1994), we used full-information maximum likelihood estimation. We specified the expected zeros in the loading matrices according to hypotheses based on the literature and our sampling of measures, but left the expected salient loadings free to be estimated independently for each participant. The key innovation was constraining the factor intercorrelation matrices to be estimated to be precisely the same for all participants.[5]

[5]The literature on factorial invariance, especially the papers by Meredith (1964, 1993), suggests that factor covariance, rather than correlation matrices should be the focus of our model specification. We needed to impose a metric for each solution, but did not want to do that by fixing a factor loading for each factor for each participant. The meaning and possible equality (or lack thereof) of the measurement variable metrics across participants is a difficult and complicated question which we are not addressing in this article. Rather, we elected to work with the correlational metric of setting all factor variances to unity.

Finally, we allowed for idiosyncrasy in the loading patterns by "tailoring" them, in a limited way, to the participant. We fixed non-significant loadings to zeroes and freed elements implicated by large modification indices, individual-by-individual, so that they could load on other factors. The "tailoring" can be seen in Tables 2–5 where the hypothesized "core" of marker variables for a given factor appear first, followed by the more idiosyncratic loadings that were allowed to be estimated.

RESULTS

Final model properties are presented in Tables 1–6. The factor intercorrelations are given in lower triangular form in Table 1. This matrix is exactly the same for all five participants. In it, Well-Being and Social Affection correlate highly positively with each other ($r = .92$; $SE = .01$), Energy correlates positively with Well-Being ($r = .61$; $SE = .03$), Social Affection ($r = .62$; $SE = .03$), and Concentration ($r = .77$; $SE = .04$), and negatively with Fatigue ($r = -.66$; $SE = .03$).

Tables 2–6 show a substantial degree of congruence in that many items load the same factors across participants (e.g., cheerful as an indicator of the Well-Being factor in Table 2). But, idiosyncrasy is also obvious from (a) the only partially overlapping subsets of indicators per factor, (b) non-invariant loadings across participants, and (c) insufficient samplings of variables to define some factors (e.g., the factors Fatigue and Concentration are not identified in Participant 3).[6]

TABLE 1
Factor InterCorrelations Constrained to be Equal across All Five Participants

Factor	*1*	*2*	*3*	*4*	*5*
1. Well-Being	1.00	–	–	–	–
2. Social Affection	.92(.01)	1.00	–	–	–
3. Energy	.61(.03)	.62(.03)	1.00	–	–
4. Fatigue	−.34(.05)	−.37(.05)	−.66(.03)	1.00	–
5. Concentration	.50(.05)	.48(.06)	.77(.04)	−.53(.05)	1.00

Note: Values in parentheses are standard errors.

[6]Some readers may be curious about the preponderance of Participant 5's loadings in excess of 1.00. We factored covariance matrices and constrained the factors to have variances of unity. Participant 5's variables had an average variance of 1.82, which was much higher than the average for Participants 1–4 (.37, 1.29, 1.05, and .86, respectively.)

TABLE 2
Factor Loading Patterns for Five Participants on the Well-Being Factor

	Participant				
Item	*1*	*2*	*3*	*4*	*5*
Cheerful	.57(.06)	.91(.09)	.97(.06)	.71(.06)	1.32(.10)
Relaxed	.49(.07)	.88(.15)	.73(.07)	.68(.10)	2.00(.21)
Contented	.53(.07)	.79(.09)	.81(.06)	.55(.06)	1.41(.10)
At ease	.38(.06)	.51(.10)	.85(.06)	.57(.09)	1.25(.10)
Comfortable	.38(.05)	.33(.10)	.77(.06)	.40(.09)	.45(.15)
Happy	.55(.05)	.87(.08)	.96(.06)	.63(.06)	–
Glad	.51(.06)	.96(.09)	.84(.06)	.72(.06)	–
Pleased	.40(.06)	.96(.08)	.96(.06)	.64(.05)	–
Carefree	.35(.06)	–	.88(.07)	.45(.08)	1.38(.11)
Calm	.44(.05)	.38(.10)	.59(.10)	–	1.20(.10)
Friendly	–	–	–	–	1.15(.09)
Efficient	–	–	.75(.07)	–	–

Note. Only variables with a salient loading for at least one participant are shown. Values in parentheses are standard errors.

TABLE 3
Factor Loading Patterns for Five Participants on the Social Affection Factor

	Participant				
Item	*1*	*2*	*3*	*4*	*5*
Warmhearted	.41(.05)	1.03(.10)	.86(.06)	.57(.05)	1.26(.13)
Kindly	.48(.06)	.84(.08)	.79(.07)	.54(.06)	1.29(.11)
Affectionate	.47(.05)	.66(.13)	.51(.08)	.69(.07)	1.21(.12)
Friendly	.52(.05)	.89(.08)	.74(.06)	.68(.05)	–
Forgiving	.27(.07)	.68(.10)	.75(.07)	.29(.05)	–
Excited	–	.51(.13)	–	.45(.07)	–
Grouchy	–	–	–	– .63(.07)	– .47(.06)
Happy	–	–	–	–	1.44(.10)
Pleased	–	–	–	–	1.39(.11)
Glad	–	–	–	–	1.18(.12)
Relaxed	–	–	–	–	– .86(.19)
Enthusiastic	–	–	–	.43(.08)	–

Note. Only variables with a salient loading for at least one participant are shown. Values in parentheses are standard errors.

Model Fit

The model fit was quite acceptable ($\chi^2 = 4{,}529$; d$f = 2{,}812$; Akaike Information Criterion (AIC) = 27,005; RMSEA = .075). For reference purposes, a model

TABLE 4
Factor Loading Patterns for Five Participants on the Energy Factor

	Participant				
Item	*1*	*2*	*3*	*4*	*5*
Energetic	.59(.07)	1.18(.08)	.93(.07)	1.02(.07)	1.14(.10)
Enthusiastic	.48(.08)	.99(.11)	.74(.08)	.29(.07)	1.12(.11)
Active	.53(.05)	1.09(.09)	.92(.08)	.87(.06)	1.09(.10)
Peppy	.38(.05)	.88(.07)	.70(.07)	.98(.07)	1.12(.11)
Lively	.54(.06)	1.05(.07)	.89(.07)	.90(.06)	1.13(.10)
Excited	–	.59(.13)	.78(.09)	.34(.07)	1.56(.29)
Aroused	–	.51(.11)	.53(.07)	.58(.08)	–
Forceful	–	.83(.08)	–	.58(.08)	–
Vigorous	–	1.05(.08)	.92(.08)	.88(.07)	–
Affectionate	–	–	.30(.08)	–	–
Relaxed	– .21(.06)	–.68(.14)	–	–.74(.10)	–
Carefree	–	.53(.09)	–	–	–
Efficient	–	.89(.10)	–	–	–

Note. Only variables with a salient loading for at least one participant are shown. Values in parentheses are standard errors.

TABLE 5
Factor Loading Patterns for Five Participants on the Fatigue Factor

	Participant				
Item	*1*	*2*	*3*	*4*	*5*
Tired	.86(.07)	1.23(.09)	–	1.05(.08)	1.50(.15)
Weary	.61(.05)	1.23(.10)	–	.98(.08)	1.18(.15)
Drowsy	–	1.29(.10)	–	1.15(.08)	.87(.13)
Sluggish	.56(.06)	1.15(.11)	–	.86(.09)	–
Dull	–	–	–	.76(.08)	–
Calm	.16(.04)	–	–	–	–
Comfortable	–	– .39(.09)	–	–	–

Note. Only variables with a salient loading for at least one participant are shown. Values in parentheses are standard errors.

representing the traditional conception of factorial invariance with invariant factor loadings and freely estimated factor covariances showed apparently poorer model fit to our data ($\chi^2 = 6{,}530$; d$f = 2{,}852$; AIC = 28,925; RMSEA = .110). A chi-square difference test ($\Delta\chi^2$ / Δdf = 2,001 / 40) or differences in AIC between the models (ΔAIC / Δdf = 1,921 / 40) are not strictly proper because the two models are not nested. Moreover, the minor amount of “tinkering”

TABLE 6
Factor Loading Patterns for Five Participants on the Concentration Factor

	Participant				
Item	*1*	*2*	*3*	*4*	*5*
Concentrating	–	.76(.07)	–	.36(.06)	.51(.14)
Efficient	.60(.08)	–	–	.54(.08)	.84(.15)
Intent	–	.77(.09)	–	.28(.06)	–
Earnest	–	.83(.08)	–	.19(.06)	–
Serious	–	.36(.08)	–	–	–
Excited	–	–	–	–	– .95(.31)
At ease	–	–	–	–.60(.09)	–
Contented	–.26(.06)	–	–	–	–

Note. Only variables with a salient loading for at least one participant are shown.

we did in tailoring the loading patterns to individuals doubtless involves some capitalization on chance. We also ran a second comparison model freeing the salient factor loadings and the factor covariances across groups ($\chi^2 = 4{,}425$; $df = 2{,}779$; AIC = 26,967; RMSEA = .074), thereby allowing for a nested model comparison ($\Delta\chi^2$ / $\Delta df = 104$ / 33). Although this chi-square difference is statistically significant, it is small for the magnitude of these chi-square values, suggesting that the factorial invariance specification with fixed factor covariances does provide a more parsimonious description of the data while not markedly degrading model fit.

For additional support we averaged (using Fisher's z transformation) the factor intercorrelations found in the separate confirmatory p-technique factor analyses. The resulting values were very similar to those found in the confirmatory analysis solution (Table 1). A second-order, exploratory factor analysis was then performed on the matrix of primary factor intercorrelations, primarily for descriptive purposes. Two factors representing 78% of the total variance of the five primary factors were extracted and rotated to an oblique solution. The second-order factor loading pattern is shown in Table 7. The two factors, which correlate +.58, rather cleanly separate the primary factors into broad dimensions of affect and activity.

The second-order factor analysis reinforces an important point regarding the traditional view of factorial invariance. Because the one primary factor intercorrelation matrix (Table 1) describes all participants, the second-order factor solution also describes all participants. Thus, the second-order solution is invariant over participants and the traditional view of factorial invariance is upheld-but it applies at a level of abstraction one-step removed from conventional thinking.

TABLE 7
Second-Order Factor Loading Pattern

Factor	*Factor I (Affect)*	*Factor II (Activity)*
Well-Being	1.01	– .02
Social Affection	.88	.07
Energy	.10	.92
Fatigue	.05	– .71
Concentration	.08	.74

Note. Factor I and Factor II are correlated $r = +.58$.

DISCUSSION AND CONCLUSIONS

We have presented a rationale and method for dealing with idiosyncrasy while still focusing on generality in the quest to establish lawful behavioral relationships. The key role of invariant relationships in construct identification and measurement was recognized but an alternative conception of traditional factorial invariance was proposed that filters out idiosyncratic features before estimating relations among the constructs of interest. Our approach enables "tailoring" the manifest representation of underlying constructs to the individual by defining invariance at the level of construct interrelations rather than at the level of construct-observables relations.

Are there examples of such "tailoring" of concepts in the physical sciences? In addition to the geometric figures illustration involving area and volume given earlier, the same abstract equation—the damped linear oscillator (Boker & Nesselroade, 2002)—describes the behavior of a pendulum and an electronic tank circuit. The damping parameter represents mechanical friction in the former case and electrical resistance in the latter. The abstraction—the second-order differential equation representing both systems—is the same, but the respective systems are completely different. Thus, generality (the differential equation) is made specific (and empirically testable) by the electronic technician with his observables (meter readings) and by the clock-maker with different ones (pendulum behavior).

Good theories must be abstract to provide generality but they must also provide empirically testable deductions. Relationships among abstractions such as positive and negative affect are important to psychological theory and must allow the same kind of empirical tests of their implications. We are proposing to liberalize the empirical testing procedures to allow for idiosyncratic features of construct representation and measurement.

A central concern has to do with the abstract concept of invariance and how to give it useful operational meaning. Factorial invariance is often invoked in rather unusual circumstances as when one kind of behavior (e.g., self-report, autonomic nervous system activity) is used as a device to try to get at other kinds of behavior (feelings, beliefs, intentions). The hope of identifying a parsimonious set of behavior patterns underlies much of psychological research. But such patterns must have generality to be theoretically useful. Generalities come at the price of specificity so looking, as we have here, for invariant relationships a couple of steps removed from "the data" seems warranted. It is our belief, consistent with the empirical demonstration we have provided, that it is reasonable to seek such invariant relationships in these more difficult circumstances, but that one should seek them at an appropriate level of abstraction.

To the extent that invariance holds at the level of relationships among the factors e.g., homogeneous factor intercorrelation matrices, then a remarkable fact needs to be acknowledged. Factoring those intercorrelations (second-order factor analysis) will yield sets of factors that exhibit the more traditional factor loading pattern invariance across participants, thus restoring the concept of factorial invariance to its traditional role. It is in this sense that we use the term "filter" to describe the functioning of the p-technique factor analyses. They filter out the idiosyncrasies of the measurement approach at the first-order factor level, but restore the possibility of generality at the second-order level. Our philosophical view of the first-order factors seems to be compatible with the "realism" of constructs discussed by Borsboom et al. (2003) in that the same factors are "there" for different individuals even though the observable manifestations may differ somewhat for those individuals.

The alternative of studying fewer people in greater depth runs counter to the deeply engrained, large sample approaches characterizing graduate training and, therefore, general practice. But, we are not advocating putting more unaccounted for variability in the so-called error term. Rather, our proposal is designed to eliminate it from comparisons altogether.

Elsewhere (Nesselroade, 2006), a proposal related to the "p-technique factors as idiographic filters" idea was advanced that was intended to capitalize on the capacity of Dynamic Factor Analysis (DFA) techniques to represent "processes." One version of DFA (Browne & Zhang, 2005; Nesselroade, 2002) estimates the web of auto- and cross-regressions of latent variables as the basic features of a process, but allows the possibility that the observed "indicators" of those latent variables can be idiosyncratically "tuned" just as are the individual level factor loadings we have been discussing in this article. Thus, both approaches rely on individual factor loading patterns to play the role of "idiographic filters" but in this "process" case, the locus of invariant relationships is in the lag structure of the latent variables.[7]

[7]Michael Browne (2006) has challenged us to confront the question: "Do the factor correlations remain invariant at lag 0 only or at lag 0 and at lag 1 (or 2)? Equivalently are joint relationships between mood factors at the same time invariant or are changes in mood across time also invariant?" We plan to do so in a subsequent article.

We encourage a re-examination of factorial invariance from a technical point of view. Meredith (1993) and Millsap and Meredith (2006) have laid out the situation pretty definitively from that standpoint of the classical view of factorial invariance. For example, we focus on the factor correlation matrix whereas, traditionally, factor covariance matrices represent the preferred metric. But concerns such as idiosyncratic measurement units, for example, fall far outside the purview of the present discussion.

One of the often mentioned advantages of modeling with latent variables is the opportunity they provide to isolate measurement error from the estimated relationships among constructs (disattenuation). This is clearly a reliability issue. Our proposal to use first order latent variables (factors) to filter out particular kinds of variation from estimated relationships among constructs is just as clearly a validity issue or, one might add, a validity-enhancing issue.

Another obvious concern is falsifiability. Can one always achieve an acceptable model fit if the individual factor loading patterns are altered sufficiently? The answer may be "yes," if the patterns are idiosyncratically altered enough. We let "theory"—literature-based expectations regarding the adjective scales—guide the initial modeling efforts and departed only cautiously from those expectations. Clearly, safeguards must be enunciated and kept in place in such "tailoring" operations.

Finally, the individual level arguments we have made can be applied to group data. We believe that qualitative differences among individuals must not be aggregated as though they are mere quantitative differences (see e.g., Molenaar, 2004). The former should lead us to disaggregate, rather than aggregate, in informed ways. Then, perhaps, informed aggregation can take place at a different level of abstraction. Sub-populations may be as idiosyncratic as individuals thus calling for some "tailoring" of construct measurement at that level of analysis. The proposed alternative to classical factorial invariance invites much closer scrutiny because it may enable the "filtering out" of less consequential individual differences while enhancing important similarities—similarities that reflect nomothetic lawfulness.

AUTHOR NOTES

This work was supported by the Institute for Developmental and Health Research Methodology at the University of Virginia and was begun when all the authors were affiliated with that institution. Denis Gerstorf was supported by a Research Fellowship from the German Research Foundation (DFG). Nilam Ram also thanks the Max Planck Institute for Human Development for its generous support of his work on this article. We are grateful to Paul Baltes, Michael W. Browne, Jerry Clore, Jack McArdle, and Peter C. M. Molenaar for many helpful comments and suggestions.

REFERENCES

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt, Rinehart, and Winston.

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, *81*, 506–520.

Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measuring change*. Madison, WI: University of Wisconsin Press.

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intra-individual variability: Recovering the parameters of simulated oscillators in multiwave data. *Multivariate Behavioral Research*, *37*, 127–160.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–219.

Browne, M. W. (2006, January). Personal communication.

Browne, M. W., & Nesselroade, J. R. (2005). Representing psychological processes with dynamic-factor models: Some promising uses and extensions of ARMA time-series models. In A. Maydeu–Olivares & J. J. McArdle (Eds.), *Psychometrics: A festschrift to Roderick P. McDonald* (pp. 415–452). Mahwah, NJ: Erlbaum.

Browne, M. W., & Zhang, G. (2005). *User's guide: DyFA: Dynamic factor analysis of lagged correlation matrices*. Columbus, OH: Psychology Department, Ohio State University.

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

Cattell, R. B. (1963). The structuring of change by p-technique and incremental r-technique. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 167–198). Madison, WI: University of Wisconsin Press.

Cattell, R. B., & Scheier, I. H. (1961). *The meaning and measurement of neuroticism and anxiety*. New York: Ronald.

Corballis, M. C., & Traub, R. E. (1970). Longitudinal factor analysis. *Psychometrika*, *35*, 79–95.

Daly, D. L., Bath, K. E., & Nesselroade, J. R. (1974). On the confounding of inter- and intra-individual variability in examining change patterns. *Journal of Clinical Psychology*, *30*, 33–36.

Hamaker, E. L., Molenaar, P., & Nesselroade, J. R. (2007). The integrated trait state model. *Journal of Research in Personality*, *41*, 295–315.

Horn, J., McArdle, J. J., & Mason, R. (1983). When invariance is not invariant: A practical scientist's view of the ethereal concept of factorial invariance. *Southern Psychologist*, *1*, 179–188.

Kagan, J. (1980). Perspectives on continuity. In O. G. Brim & J. Kagan (Eds.), *Constancy and change in human development*. Cambridge, MA: Harvard University Press.

Keyser, C. J. (1956). The group concept. In J. R. Newman (Ed.), *The world of mathematics* (Volume *3*, pp. 1538–1557). New York: Simon and Schuster.

Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist*, *36*, 276–289.

Lamiell, J. T. (1988). *Once more into the breach: Why individual differences research cannot advance personality theory*. (Paper presented at the annual meeting of the American Psychological Association, August, Atlanta, GA)

Lebo, M. A., & Nesselroade, J. R. (1978). Intra-individual differences dimensions of mood change during pregnancy identified in five p-technique factor analyses. *Journal of Research in Personality*, *12*, 205–224.

Livson, N. (1973). Developmental dimensions of personality: A life-span formulation. In P. B. Baltes & K. W. Schaie (Eds.), *Life-span developmental psychology: Personality and socialization* (pp. 97–122). New York: Academic Press.

McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research*, *29*, 409–454.

McArdle, J. J. (2006, January). Personal communication.

Meredith, W. (1964a). Notes on factorial invariance. *Psychometrika*, *29*, 177–185.

Meredith, W. (1964b). Rotation to achieve factorial invariance. *Psychometrika*, *29*, 186–206.

Meredith, W. (1993). Measurement invariance, factor analysis and factor invariance. *Psychometrika*, *58*, 525–543.

Millsap, R. E., & Meredith, W. (2006). Factorial invariance: Historical perspectives and new problems. In R. Cudeck & R. C. MacCallum (Eds.), *100 years of factor analysis*. Mahwah, NJ: Erlbaum.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time-series. *Psychometrika*, *50*, 181–202.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology—this time forever. *Measurement: Interdisciplinary Research and Perspectives*, *2*, 201–218.

Molenaar, P. C. M., & Nesselroade, J. R. (under review). *The recoverability of p-technique factor analysis*.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus, the comprehensive modeling program for applied researchers user's guide*. Los Angeles: Author.

Nesselroade, J. R. (2002). Elaborating the different in differential psychology. *Multivariate Behavioral Research*, *37*, 543–561.

Nesselroade, J. R. (2006). Factoring at the individual level: Some matters for the second century of factor analysis. In R. Cudeck & R. MacCallum (Eds.), *100 years of factor analysis*. Mahwah, NJ: Erlbaum.

Nesselroade, J. R., & Ford, D. H. (1985). p-technique comes of age: Multivariate, replicated, single—subject designs for research on older adults. *Research on Aging*, *7*, 46–80.

Nesselroade, J. R., & Molenaar, P. C. M. (1999). Pooling lagged covariance structures based on short, multivariate time-series for dynamic factor analysis. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 224–250). Newbury Park, CA: Sage.

Nunnally, J. C. (1973). Research strategies and measurement methods for investigating human development. In J. R. Nesselroade & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological issues*. New York: Academic Press.

Proffitt, D. R. (2006). Distance perception. *Current Directions in Psychological Science*, *15*, 131–135.

Rosenzweig, S. (1958). The place of the individual and of idiodynamics in psychology: A dialogue. *Journal of Individual Psychology*, *14*, 3–20.

Rosenzweig, S. (1986). Idiodynamics vis-a-vis psychology. *American Psychologist*, *41*, 241–245.

Sameroff, A. J., Peck, S. C., & Eccles, J. L. (2004). Changing ecological determinants of conduct problems from early adolescence to early adulthood. *Development and Psychopathology*, *16*, 873–896.

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.

Tucker, L. R. (1966). Learning theory and multivariate experiment: Illustration by determination of generalized learning curves. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (pp. 476–501). Chicago: Rand McNally.

van Kampen, V. (2000). Idiographic complexity and the common personality dimensions of insensitivity, extraversion, neuroticism, and orderliness. *European Journal of Personality*, *14*, 217–243.

Zevon, M., & Tellegen, A. (1982). The structure of mood change: Idiographic/nomothetic analysis. *Journal of Personality and Social Psychology*, *43*, 111–122.