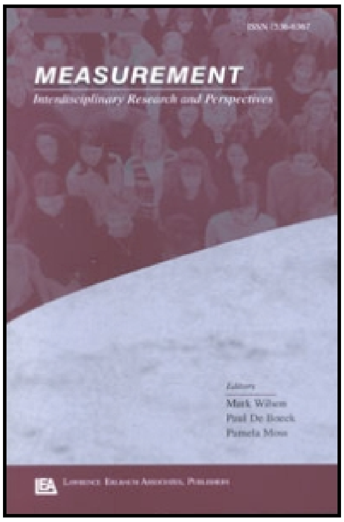

Measurement: Interdisciplinary Research & Perspective
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653679

Commentary: Theoretical Equivalence, Measurement Invariance, and the Idiographic Filter

Denny Borsboom [a]; Conor V. Dolan [a]
[a] University of Amsterdam.

MEASUREMENT, *5*(4), 236–263

COMMENTARIES

Theoretical Equivalence, Measurement Invariance, and the Idiographic Filter

Denny Borsboom and Conor V. Dolan
University of Amsterdam

Nesselroade, Gerstorf, Hardy, and Ram (this issue) propose to "filter out" idiosyncrasies of dynamic processes at the level of the individual through the application of dynamic factor analysis. The problem that they deal with is that individuals may differ in the items that are "salient" for a given construct, so that the same measurement model does not hold across subjects. The proposed method, however, allows for the same constructs to be measured by different items in different subjects, and therefore appears to bypass this problem. Nesselroade et al. accomplish this by first estimating a dynamic factor model for each individual, where the item-factor relations are free to vary across subjects, and then executing a higher order analysis in which the factor intercorrelations are required to be identical across subjects. They motivate this procedure partly by appealing to the literature on measurement invariance; the only difference, in their own words, is that they define "invariance at the level of construct interrelations rather than at the level of construct-observables relations."

We think that the attempt to accommodate interindividual differences in intraindividual dynamic structures is interesting and laudable. However, it is doubtful whether Nesselroade et al. really get what they want, that is, a procedure to bridge the gap between "nomothetic" and "idiographic" research. In particular, we see two problems with their approach. The first is conceptual, and has to do with the question of how we identify two variables (or factors) to be "the same," and what the role of measurement invariance could be in answering this question. Second, we argue that, although the proposed method may help to pool time series at a more general level than that of the individual, it does not, in its current form, integrate intraindividual processes and interindividual differences. We discuss these issues in turn.

Correspondence should be addressed to Denny Borsboom, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. E-mail: d.borsboom@uva.nl

WHEN ARE TWO FACTORS THEORETICALLY EQUIVALENT?

Nesselroade et al. intend to treat two factors as instantiations of the same theoretical attribute if the pattern of correlations (over time) that they display with other factors is the same. This is the crux of their approach; they do not want to rely on the assumption that the measurement model (i.e., the model that links theoretical constructs to observables) is invariant across subjects. Nevertheless, they do want to identify one subject's factor with that of another subject, even though these factors may be measured through different items. To achieve this, one does of course have to introduce *some* form of invariance *somewhere*. Nesselroade et al. choose to impose invariance over persons on the correlation matrix of these within-person factors. The question before us is thus: Is equality of latent correlations over individuals adequate to justify the inference that within-person factors are instantiations of the same attribute (i.e., are theoretically equivalent) even though these factors were measured through different items? To answer this question, we first consider the definition of measurement invariance (MI), and subsequently extend it to apply to the time series of different individuals.

The standard definition of MI is as follows (Mellenbergh, 1989). Item Y, an indicator of the latent variable X_b, is MI with respect to the variable S, if and only if $f(Y|X_b = x) = f(Y|X_b = x, S = s)$ for all (x, s), where the upper case letter denotes the variable, and the lower case denotes a values that the variable may assume (e.g., S = Sex and s = male or female). The function $f(Y|.)$ denotes the (conditional) density or distribution of Y. The subscript b serves to underline that in this definition the variable X is a source of variance between subjects, that is, an interindividual differences variable. In the context of the linear factor model, with S denoting groups, Meredith (1993), derived the constraints associated with strict factorial invariance, that is, measurement invariance as defined in the linear factor model for normally distributed indicators. These constraints are the equality (over groups) of factor loadings, intercepts, and residual variances.

Now, we suppose we can apply this same definition to the generalized P-technique scenario, where we administer Y repeatedly over time to a single individual. We may then consider measurement invariance over time to arise when $f(Y|X_w = x) = f(Y|X_w = x, T = t)$, where T denotes time. The subscript w serves to emphasize that the latent variable X is now a source of variance within the individual. Working on the conjecture that the implications for MI are the same as in the standard case (Meredith, 1993), this definition should imply a P-technique model with equal factor loadings, intercepts, and residual variances over time (homoskedasticity). Note that these requirements do not imply stationarity, as MI does not constrain the nature of time series of the variable X_w. For instance, in the univariate case, the time series of X_w could be a random walk, that is, a nonstationary process. We address some important consequences of this observation in the next section.

Assuming that MI with respect to time holds, we may further consider a definition of MI with respect to different individuals: $f(Y|X_w = x) = f(Y|X_w = x, I = i)$, where I is a nominal variable denoting individuals, and i is a given individual. Nesselroade et al. appear to accept that, in the context of the linear factor model, such a definition of measurement invariance would require the equality over individuals of factor loadings, intercepts, and residual variances (although in the text they focus mainly on the equality of factor loadings). However, they *discard* the requirement of equal factor loadings, because it is judged to be unrealistic given “possibly idiosyncratic features of behavior.” In their words, “[allowing] factor loadings to vary accommodate individuality, which might be due to uniqueness of expression of the construct, particularities of the measurement process, or both.”

However, relinquishing the conditions of MI comes at the high price of not really being sure that the latent variables of interest are comparable over individuals. Nesselroade et al. assume that there will be sufficient resemblance across subjects (i.e., in the observed configuration of factor loadings) that the “substantive labels given to the factors seem reasonable.” We are reticent to accept this because in our experience, there is huge variation in what scientists judge to be “reasonable.” Moreover, to rely on their judgments on this issue seems to reduce the issue of measurement invariance to one of face validity. This is a questionable move, in our view.

Perhaps more important, however, is that Nesselroade et al. give the impression that the requirement of measurement invariance could be *replaced* by the requirement of equal correlations between intraindividual factors. Thus, Nesselroade et al. appear to think that, in the absence of measurement invariance, such equal correlations can be interpreted as *evidence* that the factors in question are instantiations of the same attribute; hence that the researcher can *rely* on this evidence when judging whether two factors are “the same.” This, we think, is incorrect.

In this respect, it should be noted that the requirement of equivalent correlations with other factors is not necessary for two factors to be instantiations of the same attribute. It is, in fact, quite easy to identify situations in which the same attribute is differently related to other attributes in different persons. For instance, the intraindividual relation between bodily height and weight in adulthood can be expected to differ between people. Assuming that people shrink during adulthood, we can expect this correlation to be zero in one subpopulation (those whose weight randomly fluctuates around its mean), to be positive in a second subpopulation (those who systematically lose weight across adulthood), and to be negative in a third (those who systematically gain weight across adulthood). If equal correlations with other variables were necessary for theoretical equivalence of the measured variables, then we would be forced to conclude that, say, the tape measure used to measure length cannot measure the same attribute in each of these subpopulations, and the scales used to measure weight cannot do so either. After all, the correlation between these two attributes is not the same for members of the different subpopulations. This is clearly incorrect; hence, equal correlations with other variables cannot be a necessary condition for the theoretical equivalence of two factors as instantiations of the same attribute.

Moreover, it is easy to come up with examples where substantive considerations suggest that theoretical equivalence implies exactly the *opposite* of equivalent correlations; for instance, substantive theory may suggest that Hank's factor X_H should, in fact, correlate negatively with his Y_H, whereas Jane's factor X_J should correlate positively with her Y_J To give one example, it is well known in certain Amsterdam academic circles that the first author of this paper shows a distinct positive relation between alcohol consumption and sociability, whereas the opposite holds for the second author of this paper. Hence, if we were to apply a procedure along the lines proposed by Nesselroade et al. (let's say we have measured sociability using different items in the two authors of this paper), substantive considerations—that is, the factual knowledge of the disparate effect of alcohol—would require the relevant factors to show opposite, rather than equivalent, patterns of correlations. This does not mean that the idiographic filter of Nesselroade et al. is useless or wrong, but it does establish that the pattern of correlations implied by the hypothesis of theoretical equivalence may take literally *any* form depending on the substantive context. Clearly, then, such considerations do not fall within the scope of the theory of measurement invariance.

We conclude that the requirement of equal correlations between factors: (a) is not necessary for the theoretical equivalence of such factors, (b) may be undermined rather than supported by substantive considerations regarding the individuals in question and, (c) does not fall within the scope of the theory of measurement invariance. This requirement cannot therefore serve as an alternative to the requirement of measurement invariance.

INTEGRATING THE IDIOGRAPHIC AND THE NOMOTHETIC

Nesselroade et al. present their procedure partly as an attempt to bridge the gap that exists between the study of intra- and interindividual differences, or, in their terminology, between "idiographic" and "nomothetic" paradigms. With respect to this issue, Nesselroade et al. claim that their method "facilitates the aggregation of information across participants and enhances the detection of nomothetic relationships, while recognizing and isolating idiosyncrasy"; its objective is to establish "explicit invariant relationships at the individual level that will support a claim that the relationships hold equally well at the group level." It is questionable, however, whether the procedure suggested will actually carry us this far. To investigate this issue, we need to address the nature of the relation between pooled latent variable correlations, as obtained using the proposed idiographic filter, and the latent variable correlations obtained in standard interindividual differences research.

We imagine that the latent variables, denoted by X_b above, that is, the latent variables of R-technique factor analysis, fall within the nomothetic theoretical framework, and that the latent variables, denoted by X_w above, that is, the latent variables in P-technique factor analysis, fall within the idiographic framework. For conceptual clarity concerning what is meant by "the group level," it is important to note that the objective of the procedure proposed by Nesselroade et al. is to establish nomothetic relationships with respect to the intraindividual latent variables X_w, *not* the interindividual latent variables X_b. The question that Nesselroade et al.'s procedure addresses is this: Under which conditions may we take the observed processes to be indicative of the same underlying dynamic structure, even though measurement invariance over subjects is lacking? However, the question that appears to be central to the relation between interindividual differences and intraindividual processes is a different one, namely: When can differences in observed variables *within a single subject over time* be attributed to the same factors that cause differences in observed variables *between subjects at a given time point*?

In contrast to the question, addressed above, of whether within-subject factor correlations should be equal for two factors to be instantiations of the same theoretical attribute, this question *does* fall squarely within the scope of the theory of measurement invariance. Moreover, we submit that its answer lies in the *rigorous* application of MI to multi-subject generalized P-technique; that is, in an application that does *not* allow any variation in the measurement model that relates the observables to the latent structure within individuals. To see this, suppose we wish to interpret a within-Hank observed difference between two time points (say, Hank's observed score increased by 10 points in the time between the measurements) as arising due to changes in the *same* attribute that is responsible for the difference observed between Hank and Jane at a single time point (say, Hank scored 10 points higher than Jane). Under which conditions is such an inference justified?

This basically comes down to the question whether X_b and X_w, as discussed above, can be taken to refer to the same latent dimension. Now, assuming that the standard (multi-group R technique) constraints apply, the identity of X_b and X_w can be investigated in a multi-subject (generalized) P-technique factor analysis with equality constraints on factor loadings, intercepts, and residual variances.

However, this clearly requires us to model the trend in the time series, which may or may not be stationary; that is, we need to model the mean trends over time within subjects, as well as the mean differences between subjects at a given time point. If we do not do this, we cannot, as a matter of principle, test the hypothesis that the within-subject mean differences and the between-subject mean differences are differences on the same dimension. Now, if we cannot test this hypothesis, we cannot assume it is correct. And if we cannot assume that it is correct, we cannot interpret observed differences between subjects in terms of the same attributes that underlie observed differences within subjects. Hence, without testing this hypothesis, we cannot bridge the gap between interindividual and intraindividual variation, and, as a result, we cannot achieve the highly desirable integration of the "idiographic" and "nomothetic" research paradigms.

The filter proposed by Nesselroade et al., however, centers the time series of subjects. That is, it only applies to variation *around the mean* of any given subject, not to variation *of the means* of different subjects. Since the latter variation can be expected to make up part of the interindividual differences observed at any given time point (see also Hamaker, Nesselroade, & Molenaar, 2007), we may, under the assumptions of Nesselroade et al., still create any relation between interindividual differences that we please, regardless of the nature of the intraindividual processes. For instance, we may create a situation, in which the correlation between two observed variables is negative in every subject, while across subjects the correlation between these variables at any given time point is positive. This can be done by simply creating a process model that ensures that the covariation of X and Y *around* their respective means is positive *within* each of the subjects, while the covariation *of these means across subjects* is negative at any given time point (e.g., see Figure 2 in Hamaker et al. (2007), for a graphical illustration of this situation). So interpreted, however, the procedure of Nesselroade et al. cannot be taken to identify "explicit invariant relationships at the individual level that will support a claim that the relationships hold equally well at the group level" (unless, of course, Nesselroade et al. have in mind something entirely different from the kind of relations we discuss here).

We submit that, for within-subject differences on observables to be interpreted in the same theoretical terms as between-subject differences, one is forced to explicate how mean differences on the latent variables translate into mean differences on the observables, that is, one must include mean structures (trends) in the model. Furthermore, if one wants to attribute the within-Bill differences and within-Jane differences to the same source as the between-Bill-and-Jane differences, one need full invariance of the measurement model across subjects, for precisely the reasons that Meredith (1993) identified.

Alternatively, one may study the between-subject relations and the within-subject relations separately, as is done for instance in the model of Hamaker et al. (2007). But this does *not* require or guarantee that "explicit invariant relationships at the individual level [. . .] will support a claim that the relationships hold equally well at the group level." For in the model of Hamaker et al. (2007), the within-subject relations may be entirely different from the between-subject relations. In conclusion, the "integrating force" of the procedure of Nesselroade et al. would seem to be limited with respect to bringing nomothetic and idiographic research closer.

CONCLUSION

The procedure that Nesselroade et al. propose may be useful to establish that the dynamic structure of a process is the same across subjects, even though measurement invariance does not hold. However, it will work if and only if the requirement of equal correlations between factors is motivated on substantive grounds. The reason is that it is easy to come up with examples where the requirement of equal correlations is violated, even though the same factors are being measured, and even situations where one would expect the requirement to be violated if the same attribute were measured in different subjects. Therefore we cannot impose equal latent correlations and assume that satisfaction of this requirement indiscriminately supports the hypothesis that the factors in question are theoretically equivalent.

Moreover, to establish that differences between subjects arise from the same source as differences within subjects, the procedure proposed by Nesselroade et al. is incomplete, because it addresses neither mean differences between subjects, nor time-dependent differences of the mean within subjects. Therefore, for this purpose, one should seek to establish measurement invariance as classically defined, but now implemented at the level of the individual subjects. We conjecture that, if measurement invariance is not met, observed differences within subjects cannot be ascribed to the same sources as differences between subjects, for exactly the same reasons that one cannot, in the absence of measurement invariance, interpret observed mean differences between groups in terms of latent differences (Meredith, 1993). Hence, if measurement invariance is not met, the idiographic filter will not yield the possibility to interpret intraindividual differences in the same terms as interindividual ones.

ACKNOWLEDGMENTS

Borsboom's research was supported by NWO innovational research grant no. 451-03-068.

REFERENCES

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

Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, *13*, 127–143.

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

Nesselroade, J. R., Gerstorf, D., Hardy, S. A., & Ram, N. (2008) Idiographic filters for psychological constructs. *Measurement*, *5*, 217–235.

Where is the Individual? Comments on Nesselroade, Gerstorf, Hardy and Ram

K. Warner Schaie
Pennsylvania State University and the University of Washington

Nesselroade et al. have done a marvelous job in discussing the methodological issues for a meaningful revival of the idiographic versus nomothetic debate that has flared up periodically over the past seven decades (e.g., Allport, 1937; Bem & Allen, 1974; Rosenzweig, 1958; Molenaar, 2004; Zevon & Tellegen, 1982). Nesselroade et al. have previously attempted to resolve the paradox that all behavior occurs at the individual level but that scientific generalizations can only be made for populations of individuals by emphasizing the important role of P-technique factor analysis (e.g., Nesselroade, 2006; Nesselroade & Ford, 1985). In their present article Nesselroade et al. show more explicitly how concepts of intra- and interindividual variance can be utilized to treat the generalizability of sets of individual factor patterns at the second-order level.

A critical element of their proposal for filtering individual idiosyncratic content from psychological constructs in the study of individual differences is to shift invariance assumptions concentrating on the relationship between the observed and latent constructs to requiring invariance for the correlations among

Correspondence should be addressed to K. Warner Schaie, The Pennsylvania State University, 118 Henderson Bldg., University Park, PA 16802. E-mail: kws@psu.edu