Development and Psychopathology **22** (2010), 239–254
© Cambridge University Press, 2010
doi:10.1017/S0954579410000015

Special Section
Match Between Person-Oriented Methods and Theory

KEYNOTE ARTICLE

Matching method with theory in person-oriented developmental psychopathology research

SONYA K. STERBA AND DANIEL J. BAUER
The University of North Carolina at Chapel Hill

Abstract

The person-oriented approach seeks to match theories and methods that portray development as a holistic, highly interactional, and individualized process. Over the past decade, this approach has gained popularity in developmental psychopathology research, particularly as model-based varieties of person-oriented methods have emerged. Although these methods allow some principles of person-oriented theory to be tested, little attention has been paid to the fact that these methods cannot test other principles, and may actually be inconsistent with certain principles. Lacking clarification regarding which aspects of person-oriented theory are testable under which person-oriented methods, assumptions of the methods have sometimes been presented as testable hypotheses or interpreted as affirming the theory. This general blurring of the line between person-oriented theory and method has even led to the occasional perception that the method *is* the theory and vice versa. We review assumptions, strengths, and limitations of model-based person-oriented methods, clarifying which theoretical principles they can test and the compromises and trade-offs required to do so.

Several influential developmental theories of the last 25 years have emphasized the importance of studying individual adaption and maladaption from a holistic–interactionist perspective (e.g., Cairns, 1979; Cicchetti & Rogosch, 1996; Cicchetti & Schneider-Rosen, 1986; Kagan, 1994; Magnusson, 1985; Rutter, 1996; Sameroff, 1982; Sroufe & Rutter, 1984). This perspective asserts that an individual's prior behaviors, genetic makeup, and contextual risk or protective factors operate as an integrated whole; taken in isolation, they may lose their meaning and consequence for that individual's behavioral course. Yet the most commonly used longitudinal methods of the last quarter century for examining stability (e.g., panel and cross-lag models) and change (repeated measures analysis of variance, analysis of difference scores, analysis of residualized change) largely preclude this perspective. These highly restrictive *variable-oriented methods* (Block, 1971) assume that, given prior behavior scores, genetic markers, contextual risk and protective factors, and so forth, individuals are interchangeable units who, apart from random error, differ neither quantitatively nor qualitatively in behavioral course.

This disconcerting theory–method mismatch, together with the perception that the holistic–interactionistic perspective was "too loose and too general" (Magnusson & Torestad, 1993, p. 447), fueled an effort to distill this perspective into a concrete and unified set of core theoretical principles and to identify and explicate a set of concordant statistical methods (Bergman, 1998, p. 84; Bergman & Magnusson, 1991, p. 324, 1997, p. 293; von Eye & Bergman, 2003, p. 554; von Eye & Bogat, 2006, p. 392.) These theoretical principles define *person-oriented theory.* Their concordant methods have been labeled *person-oriented methods.* The combination of both defines the *person-oriented approach* to developmental psychopathology research. Key principles defining person-oriented theory are given in Table 1. In brief, these principles state that development is partly individual specific (*individual-specificity principle*), involves complex interactions, such as Person × Person × Context × Time interactions (i.e., Colic × Gender × Depressed Mother × Time; *complex-interactions principle*), involves interindividual differences in intraindividual change (*interindividual-differences/intraindividual-change principle*), and can be summarized by patterns of variables (*pattern-summary principle*) that are

This work was supported by an NIMH F31 MH080494 award (to S.K.S.) and an NSF SES-0716555 award (to D.J.B.). The authors thank Sy-Miin Chow for helpful comments on an earlier draft.

Address correspondence and reprint requests to: Sonya Sterba, L. L. Thurstone Psychometric Laboratory, Department of Psychology, The University of North Carolina at Chapel Hill, Campus Box 3270, Chapel Hill, NC 27599-3270; E-mail: ssterba@email.unc.edu.

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

Table 1. *Published person-oriented theoretical principles or tenets*

Principle	Description	Source
1. Individual specificity	"Functioning, process, and development of behavior are, at least in part, specific and unique to the individual."	von Eye & Bergman (2003), tenet 1, p. 554; see also Bergman & Magnusson (1991, 1997), tenet 1; von Eye & Bogat (2006), tenet 1; Bergman (1998), tenet 1
2. Complex interactions	"The process is complex and is conceptualized as involving many factors that interact at various levels which may be mutually related in a complicated manner."	Bergman & Magnusson (1997), tenet 2, p. 293; see also Bergman & Magnusson (1991), tenet 2; von Eye & Bergman (2003), tenet 2; von Eye & Bogat (2006), tenet 2; Bergman (1998), tenet 2
3. Interindividual differences in intraindividual change	"There is lawfulness and structure to intra-individual constancy and change and inter-individual differences in constancy and change."	von Eye & Bogat (2006), tenet 3, p. 392; von Eye & Bergman (2003), tenet 3; see also Bergman & Magnusson (1991, 1997), tenet 3; Bergman (1998), tenet 3
4. Pattern summary	"Processes develop in a lawful way that can be described as patterns of the involved factors."	von Eye & Bogat (2006), tenet 4, p. 392; von Eye & Bergman (2003), tenet 4; see also Bergman & Magnusson (1997), tenet 4
5. Holism	"The meaning of the involved factors is determined by the interactions among these factors."	von Eye & Bogat (2006), tenet 5, p. 392; von Eye & Bergman (2003), tenet 5; Bergman & Magnusson (1997), tenet 4
6. Pattern parsimony	"Although there is, theoretically, an infinite variety of differences with regard to process characteristics and observed states at a detailed level, at a more global level there will often be a small number of more frequently observed patterns."	Bergman & Magnusson (1997), tenet 5, p. 293; see also Bergman & Magnusson (1991), tenet 4; von Eye & Bergman (2003), tenet 6; von Eye & Bogat (2006), tenet 6; Bergman (1998), tenet 4

not meaningfully reducible to their component variables (*holism principle*) and are few in number (*pattern-parsimony principle*). Methods that have been identified as person oriented include less-restrictive variable-oriented methods (e.g., latent growth curve modeling), classification methods (e.g., longitudinal cluster analysis, latent class growth modeling, latent transition modeling, latent Markov modeling, longitudinally linked configural frequency analysis), hybrid classification methods (e.g., latent growth mixture modeling, mixed latent Markov modeling), and single-subject methods (e.g., *p*-technique factor analysis, dynamic factor analysis; as in Bergman & Trost, 2006; Curran & Willoughby, 2003; Curran & Wirth, 2004; Magnusson & Torestad, 1993; Molenaar, 2004, 2007; Muthén & Muthén, 2000; von Eye & Bogat, 2006).

The person-oriented approach has been gaining popularity in developmental psychopathological research (e.g., Cicchetti & Canon, 1999; Emde & Spicer, 2000; Gottlieb & Halpern, 2002; Sameroff & Mackenzie, 2003; Zahn-Waxler, Klimes-Dougan, & Slattery, 2000). This trend reflects the appeal of its underlying theory and methodology, and its potential for interrelating subdisciplines historically divided along variable-centric lines (e.g., neurological, social, emotional, and cognitive development). Nevertheless, the implementation of the person-oriented approach in developmental psychopathology research has been persistently hampered by two issues.

The first issue is an underappreciation of differences among person-oriented methods, which are of paramount importance when choosing an appropriate method. Methods labeled person-oriented differ with respect to which of the person-oriented theoretical principles they can *empirically test* and which they must *assume true*. In addition, no person-oriented method is compatible with all person-oriented principles. However, the similar labeling of many methods as "person-oriented" has been thought to obscure these important differences (e.g., Block, 2000), implicitly downplaying the informed compromises and trade-offs that must be made in selecting an appropriate method.

The second issue is a perceived blurring of the line between person-oriented theory and person-oriented methods. Certainly, the value of empirically assessing person-oriented principles, rather than taking them as de facto assumptions, has been recognized (Bergman & Trost, 2006, p. 604; Magnusson, 1998, pp. 51–52). However, person-oriented theorists have nearly exclusively promoted descriptive/heuristic person-oriented methods, which have limited ability for testing hypotheses about competing models (Bergman, Magnusson, & El Khouri, 2003, p. 45), and this has served to broadly downplay the statistical testing of these principles. Consequently, there is a lack of clarity in practice regarding which person-oriented theoretical principles can be investigated, via which methods. This lack of clarity has in turn led to instances where a priori assumptions of person-oriented methods are posited as empirical hypotheses to be tested. In addition, investigators sometimes appear to confuse person-oriented theory and person-oriented methods. Bergman and Trost (2006) even describe encountering the theory *defined* as the method: "one sometimes finds in the literature caricatures of definitions of the person-oriented approach such as 'it is

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

cluster analysis' or 'it is categorizing people.' Statements of this kind ignore the theoretical aspects of this approach" (p. 624).

Our aim is to clarify the trade-offs and compromises that must be made by the developmental psychopathology researcher who is deciding among methods labeled person-oriented and who is seeking to test person-oriented theories. These trade-offs include (a) which and how many person-oriented principles can be empirically tested by the method, (b) what areas of model robustness are sacrificed to render these person-oriented principles testable, and (c) how the method balances risk of *ecological fallacy* (drawing false conclusions about individual behavior from population behavior) against risk of *atomistic fallacy* (drawing false conclusions about aggregate behavior from individual behavior). To do so, we consider a set of person-oriented methods. We highlight their statistical assumptions, strengths, limitations, and areas of interpretational confusion. We then review which person-oriented principles they can and cannot be used to evaluate. We use published empirical examples from developmental psychopathological research to provide concrete grounding for each method.

The set of person-oriented methods we consider minimally overlaps with prior reviews of person-oriented theory and methods geared toward developmental psychopathologists (e.g., Bergman & Magnusson, 1997; von Eye & Bergman, 2003), as follows. First, prior theory/method reviews have almost exclusively covered one type of person-oriented method—*classification methods*—but we cast a wider net, including *hybrid classification methods* (newer and popular, but often misunderstood), *single-subject methods* (older and underused, but reemerging with more flexible software implementations), and *less-restrictive variable-oriented methods* (likely familiar, but not as a tool for person-oriented research). Second, prior theory/method reviews have almost exclusively considered *heuristic*, descriptive classification methods (e.g., K-means cluster analysis, hierarchical cluster analysis by Ward's method) rather than *model-based* alternatives (e.g., latent class and Markov models). Model-based methods have some unique and important strengths over descriptive/heuristic methods, such as the provision of standard errors for statistical hypothesis tests, the ability to contrast appropriateness of alternative models using overall model fit indices, and the allowance for uncertainty and measurement error that can otherwise bias statistical inference. For these reasons, model-based person-oriented methods are seeing increased interest and application in developmental psychopathology research (Bauer, 2007; Kaplan, 2008; Muthén, 2007; Nagin, 2005; Nesselroade, Gerstorf, Hardy, & Ram, 2007). Still, it remains difficult for person-oriented theorists to adjudicate among model-based person-oriented methods because these methods have typically been presented in isolation. Even versions of these methods suitable for capturing continuous versus discontinuous change have been discussed in separate literatures (e.g., Collins & Wugatler, 1992; Kaplan, 2008; Muthén & Muthén, 2000; Nagin, 2005). Hence, our review emphasizes interrelations among model-based person-oriented methods, where possible including descriptions of variants for capturing continuous and discontinuous (stage-sequential) change.

Our Keynote Article describes each method conceptually in the text with the aid of plots and diagrams. For the interested reader, accompanying model equations are presented in an online appendix (see http://www.unc.edu/~ssterba or contact the first author). Each method is identified as either being able to render a theoretical principle testable, conditionally testable, or untestable. When a principle is identified as testable or conditionally testable, the online appendix details one way in which such tests may be obtained.

Case I. Less-Restrictive Variable-Oriented Methods

Latent growth curve model (LGM)

The less-restrictive variable-oriented method considered here is the LGM (Bollen & Curran, 2006; Meredith & Tisak, 1990). It is identical to the hierarchical linear model growth curve (Willet & Sayer, 1994). The LGM is used to model quantitative individual differences in continuous developmental change for a population of persons. For instance, consider Gilliom and Shaw's (2004) linear LGM for externalizing symptoms; it posited straight line growth for all children, but quantitative differences in the levels and slopes of individuals' behavioral trajectories. Figure 1a provides a path diagram of Gilliom and Shaw's (2004) LGM of externalizing behavior at five time points (ages 2–6). Squares denote repeated measures of an observed variable (here, externalizing scale scores), which serve as indicators of one or more latent growth factors (denoted by circles). These latent growth factors represent aspects of change. The meaning of the growth factors is determined by the factor loadings (denoted by arrows connecting squares and circles), which often are not estimated, but rather fixed to specific values to determine the functional form, or "curve" of the latent trajectory. In Figure 1a, the loadings on the first latent growth factor (1, 1, 1, 1, 1) are shown to define it as the latent initial externalizing status (latent intercept), and the loadings on the second (0, 1, 2, 3, 4) are shown to define it as the latent rate of linear change in externalizing over time (latent slope).

One "nonrestrictive" feature of the LGM is the fact that it allows each person to have his or her *own* scores on the latent intercept and latent slope factors. These individual-specific scores are not estimated directly. Instead, the mean of each factor is estimated, along with the variance of the individual-specific scores about the mean (denoted by the curved double-headed arrow attached to each circle). This feature of the LGM allows persons to vary *quantitatively* in their latent initial status and latent rate of change. For example, Gilliom and Shaw (2004) found that children had an average of 1.7 externalizing symptoms at age 2 and a decreasing linear slope of −0.07. However, any individual's latent intercept score could quantitatively vary (higher/lower) around 1.7 and any individual's latent linear slope score could quantitatively vary (steeper/shallower) around −0.07. In contrast, a

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

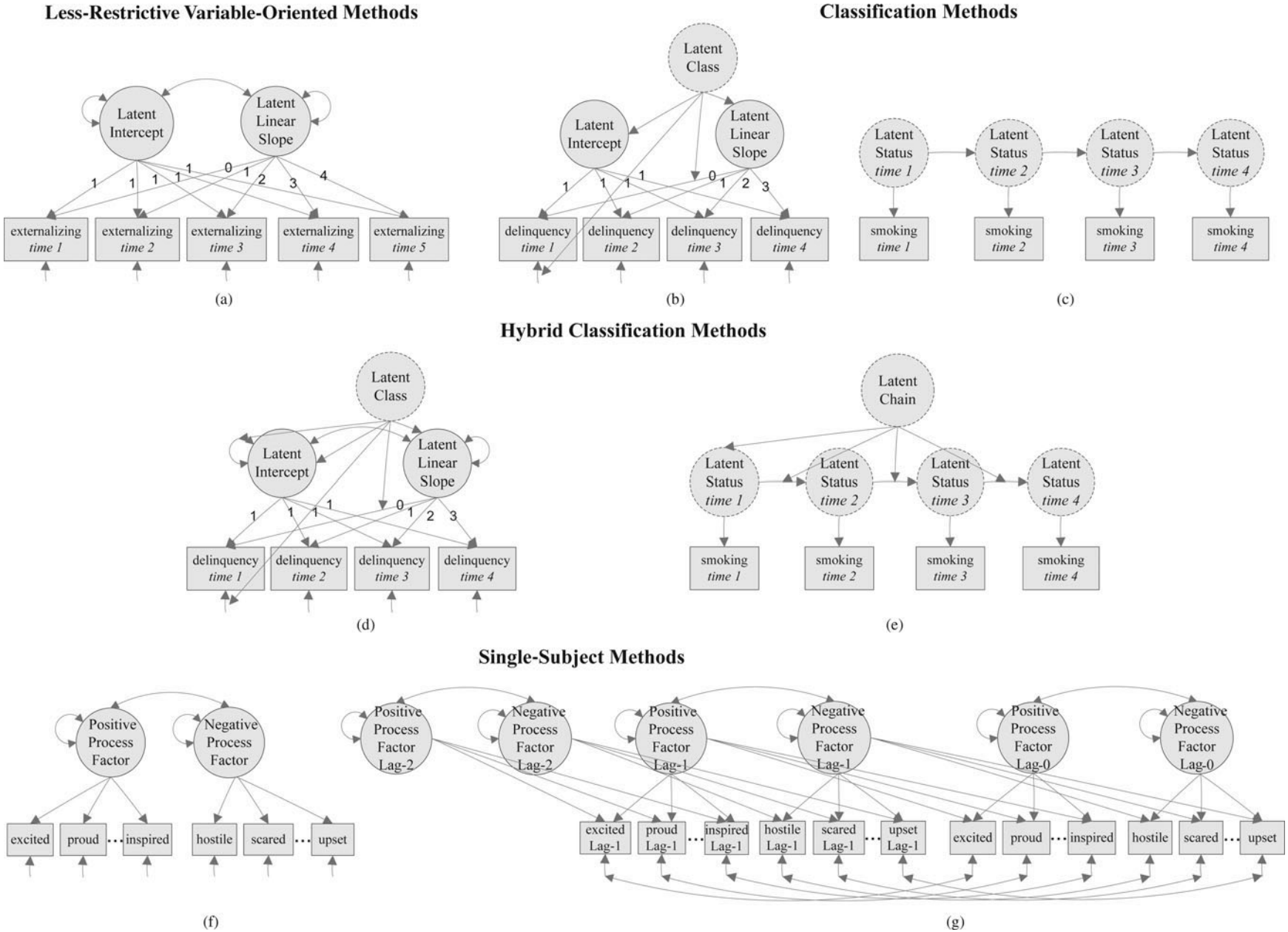


Figure 1. Path diagrams of alternative model-based person-oriented methods. (b, d, e) Latent classification or chain variables can influence other model parameters (thus allowing them to be class/chain varying); only a few examples are given here. (a) Latent growth model, (b) latent class growth model, (c) latent Markov model, (d) latent growth mixture model, (e) mixed latent Markov model, (f) *p*-technique factor model, and (g) dynamic factor model. [A color version of this figure can be viewed online at journals.cambridge.org/dpp]

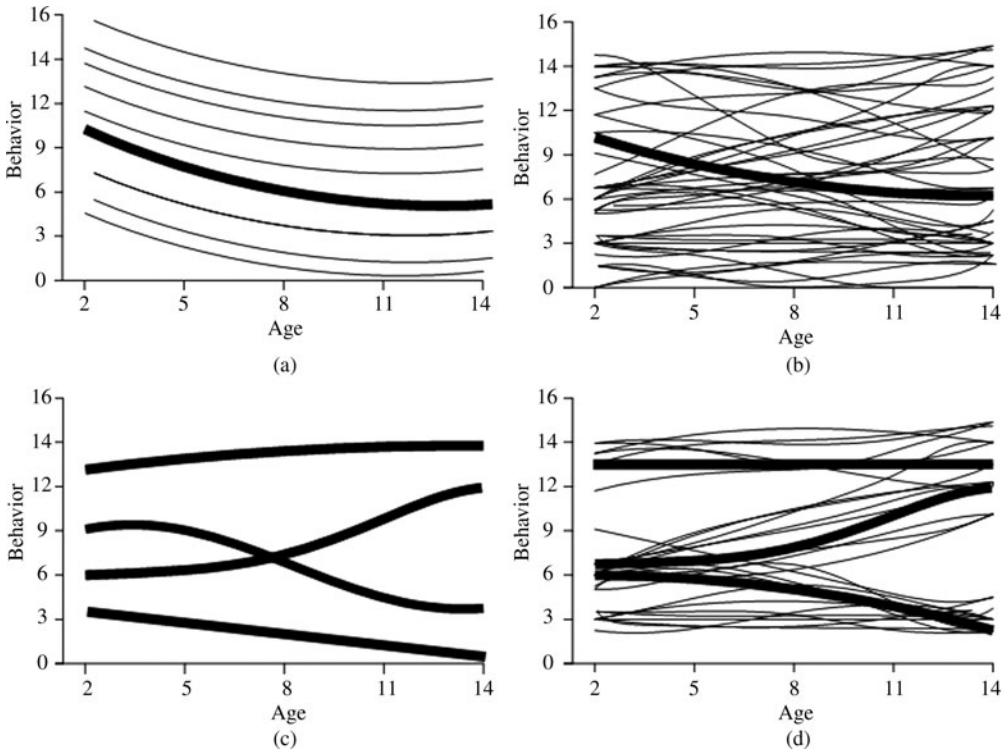


Figure 2. A schematic of model implied trajectories from alternative longitudinal models for continuous change. Thick black lines represent model implied mean trajectories, and thin black lines represent model implied individual trajectories. (a) A traditional variable-oriented model: a repeated measures analysis of variance (with fixed slopes and random intercept), (b) a latent growth curve model (with random intercept and random slopes), (c) a latent class growth model (with fixed intercept and fixed slope within class), and (d) a growth mixture model (with random intercept and random slope within class).

"restrictive" feature of the LGM is that it typically has factor loadings fixed to the *same* values for all persons (but see Mehta & West, 2000) and the *same* number of growth factors specified for all persons. This feature prevents different children from following different trajectory functions (e.g., linear vs. nonlinear).

Principles untestable with LGM

Like other variable-oriented methods, the LGM does not accord well with the pattern-summary principle and pattern-parsimony principle, as persons are allowed to vary only quantitatively, not qualitatively, in their latent initial status and latent rate of change, aside from groups defined by observed variables. Hence, these two principles are both untestable as well as inconsistent with LGM.

Principles conditionally testable using LGM

The LGM allows researchers to empirically assess the individual-specificity, complex interactions, and interindividual-differences/intraindividual-change principles to some extent. However, the interpretability of these results will be *conditional* on the legitimacy of disregarding the pattern-summary and pattern-parsimony principles. With respect to the individual-specificity and interindividual-differences/intraindividual-change principles, because each person has their own intercept and slope scores, the LGM can capture intraindividual change specific to each person, which is conditional on the constraint that all individuals follow the same general function of change (same factor loading pattern and same number of growth factors). In addition, because the LGM estimates the overall mean and variance of the individual-specific scores on the latent intercept and slope factors, LGM captures interindividual differences among the intraindividual trajectories. Finally, because the magnitude of these interindividual differences can be explicitly evaluated (e.g., by comparing a model that allows only interindividual differences in intercepts in Figure 2a with a model that also allows interindividual differences in slopes in Figure 2b), the LGM can be used to conditionally test some aspects of the interindividual-differences/intraindividual-change principle. Specifically, if the model in Figure 2b is preferable, this suggests that individuals change at differential rates over time, providing some support for the individual-specificity and interindividual-differences/intraindividual-change principles.

Regarding the complex-interactions principle, variable-oriented models are often described as "hav[ing] a limited capacity . . . for handling complex interactions" (Bergman & Magnusson, 1997, p. 298; see also Cairns, Bergman, & Kagan, 1998), leading some to advocate the use of classification methods to capture holistic patterns (e.g., Bergman, 2001).

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

On this point, however, Bauer and Shanahan (2007) noted that, to some extent, complex interactions can be recovered in variable-oriented analyses (or the less-restrictive LGMs) by including interaction terms in the fitted models. This strategy avoids the troublesome consequences of categorizing continuous predictors (MacCallum, Zhang, Preacher, & Rucker, 2002) and the possibility of creating groups that do not exist (Bauer & Curran, 2003), which have concerned methodologists implementing classification approaches (e.g., von Eye & Bergman, 2003). Furthermore, it is now quite straightforward to probe and plot higher order interactions of, say, person characteristics and/or contextual characteristics with time in a LGM model in order to conditionally evaluate the complex-interactions principle (Curran, Bauer, & Willoughby, 2004; Preacher, Curran, & Bauer, 2006). For example, Gilliom and Shaw (2004) consider four-way interactions of child negative emotionality, fearfulness, maternal negative control, and time. They found that when children had a combination of high negative emotionality, low fearfulness, and high negative maternal control, their latent externalizing linear slopes were stably elevated from age 2 to 6, rather than decreasing. Nevertheless, the interpretation of such interactions obviously becomes more difficult as the number of variables involved increases (e.g., four-, or five-way interactions).

Principles with limited testability using LGM

In practice, the LGM is used almost exclusively to study univariately defined trajectories (i.e., on a single outcome). According to the holism principle, the "externalizing" variable from Gilliom and Shaw (2004) is not meaningful once abstracted from the integrated constellation of variables that function together within the individual. The holism principle demands a multivariate approach. To close this gap somewhat, multiple univariate LGMs can be specified for different behaviors and interrelated in the same model (called parallel-process LGM; for extensions, see Bollen & Curran, 2004). The interdependency aspect of the holism principle (Magnusson & Torestad, 1993, p. 438) can be evaluated by testing whether intercepts and slopes of each process correlate. The reciprocity aspect of the holism principle (Magnusson & Torestad, 1993, p. 438) can be evaluated by testing whether instead intercepts of one process predict slopes of the other. For example, Gilliom and Shaw (2004) specified an LGM for externalizing and anothor univariate LGM for internalizing in the same model, and then tested whether the slope and intercept of one behavior predicted the slope and intercept of the other. For an example of three parallel processes, see Troop-Gordon and Ladd (2005). More than four or five parallel processes might often become difficult to estimate, however. Note that simply partialling out the variance associated with time-invariant or time-varying covariates from a univariate LGM does not test the holism principle because doing so would not capture its "moving bidirectional system" in which one process "serves as stimuli for the other but also changes as a result of the stimuli exchanges" (Magnusson & Torestad, 1993, p. 438).

LGM: Trade-offs/compromises

As summarized in Table 2, LGM is only able to test three person-oriented principles, and the interpretability of these tests is conditional on the legitimacy of a priori dismissing two other principles. However, LGMs are robust to a number of commonly violated distributional assumptions for the repeated measures and latent factors (see Bollen & Curran, 2006; Muthén & Asparouhov, 2006). Of the models considered in this article, LGM poses the least risk of atomistic fallacy (estimates of population parameters are based on data from all individuals) but most risk of ecological fallacy (some, but not all, aspects of change held equal across individuals). Promising new methods of detecting model fit or misfit for each individual (see Coffman & Millsap, 2006) and for detecting individual heterogeneity in factor loadings (Kelderman & Molenaar, 2007) can, however, help alleviate concerns about committing an ecological fallacy.

Case II. Classification Methods

We consider one classification method for continuous change and one for discontinuous change.

Latent class growth analysis (LCGA)

The first classification method considered here, LCGA (Muthén, 2001; Nagin, 1999, 2005; Nagin & Land, 1993), is used to model qualitative individual differences in continuous developmental change, for a population of persons. The conventional LGM accounts for associations among a repeatedly measured variable using a *single* population-level trajectory that varies *quantitatively* across persons, but LCGA accounts for the same associations among repeated measures using *multiple* class-specific trajectories that differ *qualitatively* between class but do *not* differ quantitatively within class, apart from random noise. This is illustrated in Figure 2c, where LCGA is shown to capture systematic associations among repeated measures using between-class mean differences in level and slope, but not systematic with-class variability in level and slope (Bauer & Curran, 2003, 2004). For example, Reinecke (2006) used a LCGA to evaluate trajectories of delinquency across four time points (7th–10th grade) and identified three trajectories: none (60%), low with slow linear increase (32%), and high with rapid linear increase (8%). Reinecke's (2006) LCGA is presented via path diagram in Figure 1b. The dotted circle denotes the latent classification variable that sorts children into class-specific delinquency trajectories. Arrows extending from the dotted circle point to what can vary over class to give classes qualitatively different trajectory shapes (e.g., factor loadings, means of latent intercept and slope factors, and within-occasion residual variances). Notably, the absence of the curved arrows that were in Figure 1a indicates the lack of systematic, quantitative variability among individuals' latent delinquency intercept and slope scores, within class.

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

Table 2. *Synopsis of match and mismatch between person-oriented theory and model-based person-oriented methods*

Person-Oriented Theoretical Principle	Person-Oriented Method			
	Less Restrictive Variable Oriented	Classification	Hybrid Classification	Single Subject
1. Individual specificity	Conditionally testable	Untestable[a]	Conditionally testable	Testable
2. Complex interactions	Conditionally testable	Conditionally testable	Conditionally testable	Untestable
3. Interindividual differences/ intraindividual change	Conditionally testable	Untestable[a]	Conditionally testable	Limited testability
4. Pattern summary	Untestable[b]	Untestable[a]	Conditionally testable[c]	Testable
5. Holism	Limited testability	Limited testability	Limited testability	Potentially testable
6. Pattern parsimony	Untestable[b]	Conditionally testable	Conditionally testable	Testable

[a]These principles are untestable and conditionally consistent or consistent with the method.
[b]These principles are untestable and inconsistent with the method.
[c]In the growth modeling but not Markov modeling framework.

Latent Markov model

The LCGA is used to model qualitative differences in *continuous* developmental change, but the second classification method considered here, the latent Markov model (Langeheine & van de Pol, 1990; Wiggins, 1973), is used to model qualitative individual differences in *discontinuous*, stage-sequential developmental change for a population of persons.[1] In the latent Markov model, an observed repeated measure (that is usually categorical, but see Schmittmann, Dolan, van der Maas, & Neale, 2005) serves as an indicator of a latent classification variable at each of several time points. For example, consider the path diagram of Mannan and Koval's (2003) latent Markov model for investigating alternative sequential patterns of adolescent smoking behavior; it is depicted in Figure 1c. The squares denote the repeated measure of self-reported smoking behavior, which was polytomous (four observed categories: never, irregular, quit, regular) and was administered at four time points (Grades 6, 8, 10, and 12). The dotted circles denote a latent classification variable at each time point; each of these latent classification variables has four *latent statuses* (never, irregular, quit, regular). Mannan and Koval (2003) found that, initially (i.e., at grade 6) 69% of students were in the never latent status, 24% were in the irregular latent status, 6% were in the quit latent status, and 1% were in the regular latent status. The arrows linking the squares and the dotted circles denote conditional response probabilities of self-reporting a particular smoking category given membership in a particular true/latent smoking status (important because adolescents appreciably underreport smoking; Mannan & Koval, 2003). These conditional response probabilities are held to be the same over time to identify the model. Regressing the latent classification variable at time $t - 1$ on the latent classification variable at time t allows persons to change/transition latent status membership over time. This forms a longitudinal *sequence* of latent statuses (e.g., *never* latent status at Grade 6, *irregular* latent status at Grade 8, *irregular* latent status at Grade 10, *regular* latent status at Grade 12). Mannan and Koval's (2003) four latent statuses at each of four time points yield 256 potential sequences (e.g., never–never–never–never or irregular–irregular–-regular–regular). They found that progressive sequences from irregular to regular and quit to regular were increasingly common through 10th grade, and that the longer students remained in the regular status, the less likely they were to transition out (also known as canalization of a developmental pathway). The fact that the latent classification variables are regressed on each other (instead of having the observed repeated measures regressed on each other) allows the transition probabilities between latent statuses to be measurement error free.

Principles untestable with classification methods

The LCGA and latent Markov model cannot test the pattern-summary principle. If there are associations among observed repeated measures, LCGA can *only* account for these associations via the estimation of multiple trajectory classes, even if in the real world, there is quantitative variation (growth factors) not qualitative variation (growth classes) in trajectories (Bauer, 2007; Bauer & Curran, 2004). In addition, if there are associations among observed repeated measures, the latent Markov model can *only* account for these associations via the estimation of multiple statuses/time point even if in the real world there is quantitative variation at each time point (latent continuum) not qualitative variation at each time point (latent discrete statuses). Hence, classes/statuses may or may not reflect qualitatively distinct subgroups. The interpretation

1. For simplicity, we discuss only the latent Markov model, and subsequently, the mixed latent Markov model, but points raised generalize to the latent transition model (Collins & Wugalter, 1992) and the mixed latent transition model (e.g., Nylund, 2007). The only difference between the latent Markov model and the latent transition model is that, in the former, the classification variable at each time point has one observed indicator and, in the latter, the latent classification variable at each time has multiple indicators.

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

of classes/statuses as real-world qualitative population subgroups is an additional *assumption* imposed by the analyst.[2] Confusion arises in practice because researchers often inadvertently imply that classification methods can affirm the pattern-summary principle. For example, Nurious and Macy (2008) initially state that "person-oriented research is based on the premise that meaningful subgroups exist" (p. 393), and then go on to review classification studies extracting $K = 3$ or 4 classes, claiming that "collectively, these results illuminate . . . the existence of distinct subgroups . . ." (p. 403). Ratelle, Guay, Vallerand, Larose, and Senecal (2007) state that their "first goal was to discover various motivational profiles that are naturally occurring among samples of college and high school students. The terminology *naturally occurring* is important here because it echoes a person-oriented approach in which the central goal is to identify the existing profiles by allowing them to emerge instead of forcing them through a priori categories (e.g., median split)" (pp. 734–735). To clarify, classification methods cannot be used to illuminate the existence, to discover, or to allow the emergence of qualitative population subgroups because classification methods extract classes/statuses irrespective of whether variation is truly qualitative or quantitative in the population.

Classification methods also cannot be used to directly test the individual-specificity and interindividual-differences/intraindividual-change principles. Classification methods include no systematic *individual*-specific components to development, only *class/status*-specific components, and permit no inter*individual* variability in intra*individual* change, only inter*class/status* variability in intra*class/status* change. In the LCGA this is because the model assumes that all persons within class homogeneously follow the same class trajectory and that within class persons' scores differ only due to random noise, not systematically. Similarly, in the latent Markov model transition probabilities from one latent status to another are allowed to vary *only* between statuses, over time, *not* within status, over persons (Langeheine, 1994). That is, all persons in the never–irregular–irregular–irregular–regular sequence are given the same probability of transitioning from an irregular to regular status from 10th to 12th grade, regardless of their current stressful life events or peer group pressures. Transition probabilities are only allowed to differ between, for example, the never–never–never–never smoking sequence and the never–irregular–irregular–regular smoking sequence. It is also equivalent to state that the latent Markov model assumes persons are interchangeable within status, and therefore within sequence. Furthermore, it is sometimes argued that, *if* the pattern-summary principle is entirely upheld/valid, then classification models are conditionally consistent with individual-specificity and interindividual differences/intraindividual change, even though no *truly* individual-specific or individually varying parameters are estimated (e.g., Cairns et al., 1998). That is, it is argued that classes/statuses are sufficient to summarize individual members (e.g., high entropy), while still doing justice to individual differences (as represented exclusively by between-class/status differences).

Principles conditionally testable using classification methods

Consequently, the pattern-parsimony principle can only be conditionally assessed, under the assumption that the pattern-summary principle is true. That is, in the LCGM, *if* the trajectory classes used as a statistical device to explain associations among repeated measures do represent qualitatively different population subgroups (i.e., if the pattern-summary principle is true), then we can meaningfully test whether there are few versus many of these real-world trajectory classes (pattern-parsimony principle). In the Markov framework, we typically would only similarly test the pattern-parsimony principle if we have multiple indicators of latent status per time point (not just one indicator, as in Mannan & Koval, 2003).[3]

In addition, the complex-interactions principle can only be conditionally assessed, under the assumption that the pattern-summary principle is true. Under this assumption, we can then meaningfully interpret interactions that are captured as changes in the prevalence of LCGM trajectory classes as a function of predictor values. We can also meaningfully interpret interactions that are captured as changes in latent Markov transition probabilities as a function of predictor values (see Nylund, 2007). Moreover, we can do so without having to correctly specify a particular functional form of the relationship between predictors and the repeated measure (as in the LGM). We recommend that when using classification methods researchers explicitly state that they are testing the pattern-parsimony and complex-interactions principles conditional on the assumption that pattern summary is true.[4]

Two important caveats are in order. First, although the pattern-parsimony principle is technically conditionally testable using classification methods, typically, a small number of latent classes/statuses (e.g., three to seven) is all that the LCGM or multiple-indicator latent Markov model can support (Bauer & Curran, 2004), eroding the falsifiability of this principle. Second, although one of the advantages of classification methods is the intuitive and accessible interpretation of potentially nonlinear, potentially complex interactions, in line with the complex-interactions principle, this advantage is unfortunately not capitalized on to a great extent in many

2. In some cases latent classes can be used simply to approximate a latent continuum without any assumption that the classes represent natural groups (see Nagin, 2005).

3. To aid interpretability, when we have just one indicator/time point in a latent Markov model, we typically just estimate same number of latent statuses/time point as there are observed response categories for that indicator (i.e., four in Mannan & Koval, 2003).

4. This recommendation would be useful for traditional variable-oriented analyses (e.g., longitudinal factor analysis) as well, where researchers often state that "k dimensions were identified in the population at each occasion," without adding "*conditional* on the assumption that dimensions rather than types are responsible for associations among the repeated measures."

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

applications. Often few (e.g., Reinecke, 2006) or no (e.g., Mannan & Koval, 2003) predictors and only lower order interactions are included. With only a few predictors, and only lower order interactions, an LGM might just as easily capture the effects as an LCGA, without having to categorize persons and risk creating artificial groupings. For example, Reinecke (2006) regressed gender and educational level on their LCGA latent classification variable and found that male adolescents with low education were more likely to be in the high delinquency trajectory and female adolescents were more likely to be in the nonoffending trajectory. These same conclusions have been obtained from a LGM (e.g., Windle, 2000).

Principles with limited testability using classification methods

The standard LCGA and standard latent Markov model do not allow investigation of the holism principle because they identify only univariately defined sequences or trajectories. However, parallel-process LCGAs or latent Markov models could be specified for multiple behaviors, and interrelated, to somewhat, but not fully, minimize this theory–method mismatch (see Flaherty, 2008; Nagin & Tremblay, 2001; and online appendix). If the conditional probability of membership in LCGA trajectory class *k* on one behavior given membership in trajectory class *j* on the other behavior was greater than chance, or if sequential transition probabilities for one behavior varied across current status membership on the other behavior, we would have evidence supporting the reciprocity aspect of the holism principle.

Classification methods: Trade-offs/compromises

Of all the methods discussed here, classification methods are most widely associated with the label person-oriented method (Bergman & Trost, 2006). Thus, it is notable that even these methods are *unable* to test some person-oriented principles. As recapped in Table 2, the classification methods reviewed here can conditionally test two principles (given that the pattern-summary principle is true) and can provide limited testing of one principle (holism) but cannot test the other three principles. Note also that the number and prevalence of extracted classes/statuses are not robust to a number of commonly encountered data conditions (see Eggleston, Laub, & Sampson, 2004; Jackson & Sher, 2008).

Conclusions from classification methods do not pose much risk of atomistic fallacy because classification methods are typically not used to make conclusions about the aggregate population. However, they do pose some risk of ecological fallacy when generalizing from the class level to the individual level (particularly given the aforementioned difficulties falsifying the pattern-parsimony principle). It can be tempting to downplay the risk of committing an ecological fallacy on substantive grounds with classes being a "good enough" depiction of development for a particular application area. Yet, this reasoning merits caution because it is seldom based on empirical evidence.

Case III. Hybrid Classification Methods

Thus far, we have seen that the conventional LGM assumes that all individuals have the same trajectory functional form (factor loadings) and the same number of growth factors. These assumptions constrict our ability to empirically assess the individual-specificity and interindividual-differences/intraindividual-change principles. We have also seen that the classification modeling assumption that individuals are interchangeable within class constricts our ability to empirically assess person-oriented principles individual specificity and interindividual differences/intraindividual change. The hybrid models considered here—growth mixture models (GMMs; for continuous change) and mixed latent Markov models (for discontinuous change)—relax these two assumptions.

GMMs

For continuous longitudinal change, a GMM (Muthén & Shedden, 1999; Verbeke & Lessafre, 1996) is obtained either by relaxing the LGM assumption that all variation in individual trajectories is quantitative, or by relaxing the LCGA assumption that all variation is qualitative. For instance, Reinecke (2006) converted the LCGA in Figure 1b into a GMM in Figure 1d by estimating *individual variability* for the latent delinquency intercept and latent delinquency slope factor *within-class*. This is denoted in Figure 1d by the addition of curved arrows to the latent intercept and latent slope factors. This difference between the LGM, LCGA, and GMM is also illustrated by comparing Figure 1a, 1b, and 1d. Figure 2d illustrates that the GMM allows associations among repeated measures to reflect a *qualitative* mixture of latent classes with different functional forms (e.g., an early-onset/decreasing class, an elevated-stable class, and an adolescent-limited class). GMM also allows each trajectory class to be heterogeneous, such that growth parameters (rate or magnitude of change) vary *quantitatively* across persons within each class.

Mixed latent Markov models

The mixed latent Markov model (e.g., Langeheine & van de Pol, 1990) relaxes the assumption of standard latent Markov models that individuals are interchangeable within a latent sequence. It does so by allowing the probability of transitioning from one latent status to another to vary, not only over time, but also over latent person-groups, called *chains* (pending identification restrictions). For instance, Mannan and Koval (2003) extended their latent Markov model into a mixed latent Markov model with two chains: a high-risk chain and a low-risk chain. The dotted circle in Figure 1e denotes the latent classification variable that sorts persons into latent chains, and the arrows extending from the dotted circle denote that the initial (Grade 6) latent status probabilities as well as the transition probabilities are allowed to vary across latent chain (pending identification restrictions). One can think of Mannan and Koval's (2003) mixed latent Markov model as

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

having a low-risk chain with 256 potential sequences and a high-risk chain with 256 potential sequences, or one can think of their mixed latent Markov model as having only 256 potential sequences, each with two chains. Under the latter conceptualization, there is *within-sequence heterogeneity* in initial status and in transition probabilities. That is, it is within-sequence, and there are interchain differences in intrachain change. Mannan and Koval (2003) also considered further restrictions on transition probabilities for the mixed latent Markov model (e.g., no backward transitions, nonzero transitions for one chain only), which are often needed, particularly when there are few measurement occasions (see Langeheine & van de Pol, 1990).

Principles conditionally testable using hybrid classification methods

Of the methods reviewed thus far, hybrid methods allow empirical assessment of the widest range of person-oriented principles, but some issues will be raised shortly about how successfully. The pattern-summary principle can be empirically assessed with GMM by comparing the fit of a GMM to a LGM, to discern whether classification of individuals is really necessary (e.g., Hirsh-Pasek & Burchinal, 2006; Sterba, Prinstein, & Cox, 2007). The pattern-parsimony principle can be empirically assessed with the GMM (and mixed latent Markov model, but typically this is only done if there are multiple indicators per time point). This assessment would involve determining whether the best-fitting number of trajectory classes (or statuses per chain) is equal to or less than a predefined "small" number.

Assuming that the pattern-summary principle is upheld, the individual-specificity principle is conditionally testable in a growth modeling framework by comparing the fit of models that require within-class homogeneity (LCGA) and models that allow systematic within-class heterogeneity across individuals (GMM; as done by Reinecke, 2006). In addition, under the same assumption, the individual-specificity principle is conditionally testable within a Markov modeling framework by comparing the fit of models that require within-class homogeneity (latent Markov) versus models that allow within-class heterogeneity across chains (mixed latent Markov; as done by Mannan & Koval, 2003). Moreover, compared to, say, the LGM, hybrid methods considerably relax the a priori assumptions that must be met to test the individual-specificity principle. To illustrate, in the GMM, tests of individual specificity are conditional on persons within class *k*, not all persons, having the same trajectory functional form and number of growth factors. Likewise, in the mixed latent Markov model, tests of individual specificity are conditional on the assumption that persons within a particular *chain*, not all persons, have the same initial status and transition probabilities. Again, assuming that the pattern-summary principle is upheld, the interindividual-differences/intraindividual-change principle can be empirically assessed, within the Markov framework, by testing whether transition probabilities vary across chain or, within a growth modeling framework, by testing whether slope growth factors have significant variances within class. Finally, the complex-interactions principle would be conditionally testable in the hybrid classification models for the same reasons as discussed under classification models.

The results of such hypothesis testing regarding the interindividual-differences/intraindividual-change, individual-specificity, pattern-summary, pattern-parsimony, and complex-interactions principles may, however, be misleading in hybrid models because these models are particularly vulnerable to specification errors in ways that impinge on the ability to test these hypotheses. One difficulty is that hybrid classification models allow both within-class/chain variation *and* between-class/chain variation to reproduce the associations among the repeated measures—*sometimes statistically equivalently* (Bartholomew & Knott, 1999; Bauer & Curran, 2004; Horn, 2000; Molenaar & von Eye, 1994). In the Markov framework, this means that it is difficult to distinguish, on purely statistical grounds, whether (a) transition probabilities vary across time but not across chains, (b) transition probabilities vary across chains, but not across time, or (c) transition probabilities differ across both time and chains, (providing there are enough time points and/or indicators to even allow all of these sources of variability; Collins, Hyatt, & Graham, 2000). Similarly, in the latent growth framework, this means that it is difficult to distinguish, on purely statistical grounds, whether a model with (a) more growth factors within class but fewer classes or (b) fewer growth factors within class but more classes is most realistic (Bauer & Curran, 2003). Moreover, the GMM may fit better than the LCM for reasons having nothing to do with the underlying latent structure: that is, simply because classes better account for skew/kurtosis of repeated measures than do factors (Bauer & Curran, 2004; Lubke & Neale, 2006). In sum, whereas the LGM decomposes the associations among repeated measures to highlight relations among variables (and assumes *no* classes), the LCGA decomposes the same associations to highlight the relations among persons (and assumes *no* factors). The GMM attempts to do both simultaneously, which renders it nonrobust to a number of assumption violations. For example, in the GMM, if a researcher misspecifies the growth model within class *k*, the model may, unbeknownst to the researcher, compensate for this by overextracting *spurious* additional latent classes (Bauer, 2007; Bauer & Curran, 2003, 2004; Lubke & Neale, 2006). This can lead to recovery of *spurious* Person × Context × Time interactions (Bauer & Curran, 2003), limiting empirical evaluation of the complex-interactions principle.

For these reasons it is undesirable for authors to explicitly state (e.g., Keller, Spieker, & Gilchrist, 2005; Schaeffer, Petras, Ialongo, Poduska, & Kellam, 2003, or otherwise imply, e.g., Mun, Windle, & Schainker, 2008; Sterba et al., 2007) that finding more than one GMM class constitutes direct evidence of unobserved population subgroups. Finding more than one GMM class only constitutes statistical evidence that a mixture distribution for the repeated measures is preferable to a (usually) normal distribution. The researcher then *chooses* whether or not to overlay the additional *assumption*

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

that the classes represent unobserved population subgroups (Bauer, 2007). To clear up this confusion, we recommend that researchers using hybrid methods explicitly state that the empirical assessment of the aforementioned principles is conditional on this additional assumption.

Principles with limited testability using hybrid classification methods

Standard hybrid methods also cannot fully test the holism principle for the same reasons as were discussed in the classification methods section. However, the same modeling extensions that were described in the classification methods section, can better approximate it.

Compromises and trade-offs: Hybrid classification methods

As summarized in Table 2, the hybrid methods reviewed here allow conditional testing of five principles—if *additional* assumptions are invoked, including that extracted classes truly represent population subgroups—and allow limited testing of the sixth principle. As discussed above, however, these hybrid models are not robust to specification errors, and so these tests require cautious interpretation, sensitivity analyses (to examine the effects of minor model modifications on extracted classes), and external validation (Bauer & Curran, 2004).

Of the models considered thus far, hybrid methods pose least risk of ecological fallacy. For example, GMM holds fewer parameters equal across all persons than the LGM and holds fewer parameters equal across class members than the LCGA. Hybrid methods also pose little risk of atomistic fallacy, as they are not typically used to draw population-level conclusions.

Case IV. Single-Subject Methods

Methods considered thus far have all assumed that some aspects of a longitudinal behavioral course, such as the number of latent growth factors or the probability of transitioning from a particular latent status to another, are descriptive of *all* individuals (for LGM and latent Markov models) or *some* individuals (for LCGA, GMM, and mixed latent Markov models). Moreover, standard versions of methods considered thus far (latent growth curve, LCGA, GMM, latent Markov, and mixed latent Markov) have only been able to define longitudinal stability and change univariately (i.e., on a single, behavioral outcome such as externalizing, delinquency, or smoking). Single-subject methods such as dynamic factor analysis (Molenaar, 1985) or *p*-technique factor analysis (Cattell, Cattell, & Rhymer, 1947), relax these assumptions. They allow the description of a *continuous*, longitudinal etiologic process to be completely tailored to a particular individual. They *exclusively* capture intraindividual etiologic processes, and they can define these processes *multivariately*, not simply univariately.

However, achieving these objectives requires a fundamentally different kind of data set than was employed in methods discussed thus far. Data sets used for the methods discussed thus far (growth modeling or Markov modeling) included several occasions, many individuals, and one dependent variable. For these methods, the data matrix is composed of individuals (rows) $\times$ occasions (columns), for a single dependent variable. Data sets used for dynamic factor analysis and *p*-technique, in contrast, include many occasions, many dependent variables, and *one* individual (i.e., a multivariate time series). Consider, for example, Shifren, Hooker, Wood, and Nesselroade's (1997) data set, which included 71 occasions of measurement and 20 dependent variables from the Positive and Negative Affect Schedule (PANAS) measured on one individual (i.e., upset, afraid, nervous, hostile, ashamed, distress, irritable, scared, guilty, jittery, interest, inspired, proud, alert, active, determined, attentive, enthusiastic, strong, and excited). To implement dynamic factor analysis and *p*-technique models, they arranged their data matrix with 71 occasions as rows $\times$ 20 variables as columns for that single individual. Whereas cross-sectional factor analyses of the PANAS scale (e.g., Watson, Clark, & Tellegen, 1988) have previously found that two factors well explain within-time across-persons associations in PANAS symptoms, Shifren et al. wanted instead to find out what factor structure explains across-time within-person associations in PANAS symptoms.

Next, we discuss how single-subject methods capture longitudinal behavioral processes *within individual* and how they define such processes *multivariately*. Single-subject methods summarize the regularity of an individual's longitudinal series of observed repeated measures using what are termed latent *process factors.* These process factors explain the across-occasion associations among observed repeated measures, within a single person. Specifying one process factor to account for the single person's PANAS repeated measures effectively posits a unidimensional emotion liability, and specifying more than one factor effectively posits a multidimensional emotion liability (e.g., a positive process and a negative process as shown in Figure 1f). By comparing alternative models with less or more process factors, we can test *how many* underlying dimensions are needed to explain interoccasion symptom associations for that particular person. Because the process factors in Figure 1f are each defined by 10, not 1, measured variables, the person's longitudinal behavioral process is defined *multivariately*, not univariately. Comparatively, note that whereas conventional factor models for data on many persons at one time point assume that the same number of factors holds across individuals (i.e., interindividually) within time, these single-subject methods (dynamic factor, *p*-technique) assume that the same number of process factors holds across time (i.e., interoccasionally) within person.

Dynamic factor analysis

The *p*-technique factor model in Figure 1f is critically limited in that it does not account for the fact that the single person's

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

time series of internalizing symptoms is time ordered, and thus may have serial dependencies. For example, if a person's positive emotion process factor had a high score at time $t-1$ because he/she was strongly feeling inspired, proud, and excited because of getting accepted into graduate school, he/she may also tend to have a high positive emotion process factor score at time t because some of those feelings are likely to linger into the next day. Dynamic factor models expand upon the *p*-technique factor model in Figure 1f to account for such serial dependencies. There are several ways that dynamic factor models can account for serial dependencies (Nesselroade, McArdle, Aggen, & Meyers, 2002); we describe and illustrate one in Figure 1g. First, lagged versions of the dependent variables are created—here lagged symptoms, for example, lag-1 irritable, lag-1 enthusiastic (see Wood & Brown, 1994). Then, to account for serial dependencies, lag-0 measured variables load on lag-0 and lag-1-positive and -negative emotion process factors, and lag-1 measured variables load on lag-1 and lag-2 process factors. Error terms associated with corresponding lag-0 and lag-1 measured variables are allowed to correlate, as shown in Figure 1g (Hershberger, 1998). By comparing alternative models with and without lagged factors/symptoms, we can tell whether there is any meaningful serial dependency, that is, whether the process factors are *dynamic*. For didactic implementations, see Hershberger (1998), Nesselroade et al. (2002), Wood and Brown (1994), and Zhang (2006).

Principles testable using single-subject methods

Single-subject methods such as dynamic factor analysis and *p*-technique factor analysis would seem to be the quintessential person-oriented methods. An advantage of single-subject methods for person-oriented research is that they do *not* assume the individual-specificity principle, pattern-summary principle, pattern-parsimony principle, or holism principle to be true or false a priori, but instead allow them to be empirically evaluated. Evaluation of the holism principle is possible if a researcher has access to multivariate time series data for one person. However, procedures for doing so have not yet been formalized, perhaps because of the underutilization of single-subject methods for person-oriented research. Evaluation of the individual-specificity, pattern-summary, and pattern-parsimony principles, described below, requires access to multivariate time series data for more than one person, such as for 12 persons, in Shifren et al. (1997).

To test the pattern-summary and pattern-parsimony principles, first, a separate dynamic factor model would be fit to each of several persons' time series data. Shifren et al. (1997) fit this model to 5 of the 12 persons' data. (Shifren et al. found that the other 7 persons had essentially no variability in their scores on the PANAS symptoms, thus preventing estimation of process factors to explain interoccasion variability.) Second, conventional tests for *measurement invariance* would be conducted to compare these persons' models, for instance, to determine if the number and nature of process factors is the same for Person 1 as for Persons 2, 3, and so forth. More specifically, the same series of measurement invariance tests that are typically used to compare conventional factor models across levels of a variable (e.g., male vs. female) for many persons can be used to compare dynamic factor models across individuals (Person 1 vs. Person 2) for many time points (see Borsboom & Dolan, 2007; and online appendix). In doing so, Shifren et al. (1997) found that a two-factor lag-1 model (as in Figure 2g) fit best for two persons, a one-factor lag-1 model (not shown in Figure 2) fit best for two persons, and a two-factor lag-0 model (as in Figure 2f) fit best for the last person. Within the pairs of people having same best-fitting model model, a similar pattern of significant factor loadings was found; however, the magnitude of the loadings differed somewhat. If Shifren et al. (1997) had found full measurement invariance across these five persons, this would have suggested that the within-individual longitudinal affect processes are similar across individuals (called *poolability*). See Nesselroade and Molenaar (1999) and Chow, Nesselroade, Shifren, and McArdle (2004) for details concerning poolability and Browne and Zhang (2005) for implementation. The pattern-summary principle would be supported to the extent that *groups* of persons share the same best-fitting dynamic factor model (i.e., the same number of process factors and similar loadings of symptoms on process factors). This principle was minimally upheld in Shifren et al. (1997), because groups of two persons shared the same best-fitting model. The pattern-parsimony principle would be supported to the extent that the number of groups yielding fundamentally different best-fitting dynamic-factor models is small. This principle was not upheld in Shifren et al. (1997), because three different varieties of best-fitting models were found among five persons. The individual-specificity principle would be supported to the extent that some parameters (e.g., the within-lag covariance between positive and negative process factors) still vary within subgroups sharing same best-fitting model and/or some persons have unique best-fitting models. This principle was upheld in Shifren et al. (1997) because one person had a unique best-fitting model.

Principles untestable with single-subject methods

Drawbacks of these single-subject methods for person-oriented developmental research are that empirical evaluation of the complex-interactions principle and interindividual-differences/intraindividual-change principle are still difficult. Regarding the complex-interactions principle, investigation of Person × Context × Time interactions is probably most easily accomplished once time series data is pooled across multiple persons. In a dynamic factor analysis or *p*-technique analysis for a single person, incorporating time-invariant person, or context, or Person × Context predictors of process factors would not be helpful because they would just be constant values for that one person, and incorporating time-variant person, or context, or Person × Context predictors of process factors would only be useful for explaining interoccasion variability.

In addition, the interindividual-differences/intraindividual-change principle is difficult to assess using the standard dynamic factor analysis model (or *p*-technique model) be-

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

cause these models assume *no systematic intraindividual change*, growth, or development in means, variances, or covariances of measured variables (called *stationarity*; Jones, 1991; Nesselroade & Molenaar, 2003), and any such trends are typically removed prior to model fitting. However, mean trends can now be incorporated into the dynamic factor analysis model following Molenaar, de Gooijer, and Schmitz, (1992), for example, using the implementation in Hershberger (1998; also see online appendix). Still, nonstationarity in variances and covariances of measured variables is not easily relaxed (Molenaar, 1994). Past implementations of dynamic factor analysis used mainly *short-term* intensive longitudinal designs (often hourly or daily rather than weekly, monthly, or yearly), which decreased the likelihood of finding marked developmental trends in mean behavior across occasions. Hopefully, the burgeoning interest in all sorts of intensive longitudinal designs and daily diary studies among developmentalists (e.g., Walls & Schafer, 2006) will fuel further advances in dynamic factor analysis theory and software to relax some of these remaining strict assumptions.

Single-subject methods: Trade-offs/compromises

As summarized in Table 2, dynamic factor analysis tests four person-oriented principles but has difficulty handling the other two. Adopting single-subject methods and, for example, finding support for the holism or individual-specificity principles, can serve as a potent wake-up call regarding the dangers of using methods that define an etiologic process univariately or methods that automatically assume between-individual differences and within-individual differences have interchangeable meaning. However, it can be difficult to know what next steps to take when little aggregation is warranted according to strict measurement invariance testing criteria (e.g., Curran & Wirth, 2004; Nesselroade et al., 2007; and Commentaries). Researchers are currently conducting robustness studies to understand the consequences of pooling dynamic factor models that are only partially invariant across persons (Nesselroade et al., 2007). Of the models considered in this article, the single-subject methods, together with tests for poolability when time series for multiple persons are available, pose no risk of ecological fallacy and also little risk of atomistic fallacy.

Conclusions

The person-oriented approach is becoming increasingly influential in developmental psychopathological research (e.g., Cicchetti & Canon, 1999; Emde & Spicer, 2000; Gottlieb & Halpern, 2002; Sameroff & Mackenzie, 2003). Researchers have been advised that "it is often more natural to employ person-oriented methods if a person-oriented perspective is believed to be valid" (von Eye & Bergman, 2003, p. 578). Accordingly, person-oriented methods, particularly model-based varieties, are seeing increasing use in developmental psychopathology research (e.g., Bauer, 2007; Kaplan, 2008). However, there is a considerable gap between the breadth of person-oriented methods available, and those actually used in empirical applications. For example, a brief review of articles that reported using person-oriented methods and were published in *Development and Psychopathology* over the past 10 years indicated that classification-type methods were used in the vast majority of cases. Specifically, 12% of these articles used less restrictive variable-oriented methods, 85% used model-based or heuristic classification or hybrid classification methods, and 3% used single-subject methods. We submit that the overwhelming dominance of classification-type methods has more to do with their frequent linkage to person-oriented theory in prior reviews and less to do with their carefully weighed suitability in each application. Typically, applications did not provide a rationale for (a) why the selected person-oriented method was chosen over other person-oriented methods, (b) whether or how the selected person-oriented method enabled testing relevant aspects of person-oriented theory, or (c) what the assumptions or statistical vulnerabilities of the selected person-oriented method were.

In response, our review showed that methods labeled person-oriented differ widely in the person-oriented theoretical principles that they can and cannot empirically test, must assume true, or are incompatible with. However, this in itself is not a problem. No statistical models, person oriented or otherwise, are perfect representations of the underlying substantive metatheory, allow testing of all relevant hypotheses, require no restrictive assumptions, and are robust to misspecification and misinterpretation. However, the problem lies in that our review also found that the trade-offs/compromises distinguishing particular person-oriented methods are underappreciated, and that the distinctions between person-oriented methods and theory are also underappreciated or even blurred.

We feel that for the person-oriented approach to continue to fruitfully affect developmental psychopathology research, the dialogue about theory–method match needs to become more concrete and specific. We need to move beyond simply making casual allusions to the person-oriented approach when categorizing persons in some manner, and we need to improve our awareness of what hypotheses new model-based person-oriented methods can and cannot test. First, we recommend that researchers explicitly state the methodological assumptions of person-oriented models used and carefully posit hypotheses that are testable using these models. Second, we call attention (echoing Molenaar, 2004, 2007) to the underutilization of single-subject person-oriented methods despite the upsurge in intensive longitudinal designs, and even though they can empirically test aspects of person-oriented theory that the often-used classification methods cannot. Third, in this review we hoped to dispel the myth that categorization/classification is required to implement the person-oriented approach. We urge researchers to explore the limits of higher order interaction effects that can be recovered using less-restrictive variable-oriented methods before turning to classification methods (which may create groups that do not exist) to aid interpretation of these effects. Fourth, we find it unfortunate that standard versions of popular model-based person-oriented methods (except for single-subject methods) provide only univariately defined trajectories/

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

sequences. We urge researchers to incorporate parallel-process modeling extensions, along with class-specific predictors and outcomes, to try to approximate a more holistically defined longitudinal behavioral course. With the continued refinement of person-oriented theory and with the continued development of more flexible and realistic companion methods, we expect the compromises and trade-offs that need be made in pursuit of a person-oriented theory–method match to shrink in kind.

Future Directions

In this Keynote Article, we evaluated the testability of a widely cited—but certainly not utterly comprehensive—set of person-oriented theoretical principles using popular model-based person-oriented methods. One direction for future work is to articulate which *non*model-based (i.e., heuristic, descriptive) methods can test which principles. Model-based person-oriented methods have frequently been dismissed on grounds of being complex, complicated, and often requiring strong assumptions (e.g., Bergman et al., 2003, p. 45); for these reasons, Bergman et al. (2003) state "our bias here is descriptive" (p. 46). However, the trade-off of model-based methods' typically greater ability to specify and test precise hypotheses has not been made explicit, in the context of testing person-oriented principles.

Another direction for future work is to articulate which model-based methods allow testing of other concepts from holistic–interactionistic theory not highlighted here. For example, prominent concepts such as the dynamic, self-organizing, nature of individual development and its potential dependence on turning-point events were not covered here. *Dynamics* of individual development can refer to a variety of time-dependent processes, such as a self-regulatory mechanism that oscillates between limits, converges on an attractor state, or progresses or regresses across stages, some of which become more or less likely over time. Some methods discussed here (Markov models) do allow time-dependent shifting (as described for Mannan & Koval's, 2003, smoking behavior example). Others (growth models) are less dynamic in that persons follow one growth curve for the duration of a study and no self-regulation is assumed to occur (Boker & Graham, 1998). However, methods such as differential equations models, although not traditionally labeled as person-oriented models, are specifically geared toward studying self-regulating or organizing systems (see Boker & Graham, 1998). Turning-point events have been accommodated in a variety of models, such as classification methods (Nagin, Pagani, Tremblay, & Vitaro, 2003) and less-restrictive variable-oriented methods (e.g., Cudeck & Klebe, 2002), but discussion of these topics was outside the scope of the present review.

References

Bartholomew, D. J., & Knott, M. (1999). *Latent variable models and factor analysis* (2nd ed.). London: Arnold.

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research*, *24*, 757–786.

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*, *8*, 338–363.

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*, *9*, 3–29.

Bauer, D. J., & Shanahan, M. J. (2007). Modeling complex interactions: Person-centered and variable-centered approaches. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 255–284). Mahwah, NJ: Erlbaum.

Bergman, L. R. (1998). A person-oriented approach to studying individual development: Snapshots and processes. In R. B. Cairns, L. R. Bergman, & J. Kagan (Eds.), *Methods and models for studying the individual* (pp. 83–121). London: Sage.

Bergman, L. R. (2001). A person approach in research on adolescence: Some methodological challenges. *Journal of Research on Adolescence*, *16*, 28–53.

Bergman, L. R., & Magnusson, D. (1991). Stability and change in patterns of extrinsic adjustment problems. In D. Magnusson, L. R. Bergman, & B. Torestad (Eds.), *Problems and methods in longitudinal research: Stability and change* (pp. 323–346). New York: Cambridge University Press.

Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology*, *9*, 291–319.

Bergman, L. R., & Magnusson, D., & El-Khouri, B. (2003). *Studying individual development in an interindividual context: A person-oriented approach*. Mahwah, NJ: Erlbaum.

Bergman, L. R., & Trost, K. (2006). The person-oriented versus variable-oriented approach: Are they complementary, opposites, or exploring different worlds? *Merrill–Palmer Quarterly*, *52*, 601–632.

Boker, S. M., & Graham, J. (1998). Dynamical systems analysis of adolescent substance abuse. *Multivariate Behavioral Research*, *33*, 479–507.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models: A synthesis of two traditions. *Sociological Methods and Research*, *32*, 336.

Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft.

Block, J. (2000). Three tasks for personality psychology. In L. R. Bergman, R. B. Cairns, L. G. Nilsson, & L. Nystedt (Eds.), *Developmental science and the holistic approach* (pp. 155–164). Mahwah, NJ: Erlbaum.

Borsboom, D., & Dolan, C. V. (2007). Theoretical equivalence, measurement invariance, and the idiographic filter. *Measurement*, *5*, 236–242.

Browne, M. W., & Zhang, G. (2005). DyFA: Dynamic factor analysis of lagged correlation matrices: Version 2.03 [Computer software and manual]. Retrieved from http://faculty.psy.ohio-state.edu/browne/

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach*. Hoboken, NJ: Wiley.

Cairns, R. B. (1979). *Social development: The origins and plasticity of interchanges*. San Francisco, CA: Freeman.

Cairns, R. B., Bergman, L. R., & Kagan, J. (1998). *Methods and models for studying the individual*. London: Sage.

Cattell, R. B., & Cattell, A. K., & Rhymer, R. M. (1947). *P*-technique demonstrated in determining psychophysical source traits in a normal individual. *Psychometrika*, *12*, 267–288.

Chow, S.-M., Nesselroade, J. R., Shifren, K., & McArdle, J. J. (2004). Dynamic structure of emotions among individuals with Parkinson's disease. *Structural Equation Modeling*, *11*, 560–582.

Cicchetti, D., & Cannon, T. D. (1999). Neurodevelopmental processes in the ontogenesis and epigenesist of psychopathology. *Development and Psychopathology*, *11*, 375–393.

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, *8*, 597–600.

Cicchetti, D., & Schneider-Rosen, K. (1986). An organization approach to childhood depression. In M. Rutter, C. Izard, & P. Read (Eds.), *Depression in young people, clinical and developmental perspectives* (pp. 71–134). New York: Guilford Press.

Coffman, D., & Millsap, R. (2006). Evaluating latent growth curve models using individual fit statistics. *Structural Equation Modeling*, *13*, 1–27.

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

Collins, L. M., Hyatt, S. L., & Graham, J. W. (2000). Latent transition analysis as a way of testing models of stage sequential change in longitudinal data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds), *Modeling longitudinal and multilevel data: Practical issues, applied approaches and specific examples* (pp. 147–162). Mahwah, NJ: Erlbaum.

Collins, L. M. & Wugatler, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research*, *27*, 131–157.

Cudeck, R., & Klebe, K. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods*, *7*, 41–63.

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2004). Testing and probing main effects and interactions in latent curve analysis. *Psychological Methods*, *9*, 220–237.

Curran, P. J., & Willoughby, M. T. (2003). Implications of latent trajectory models for the study of developmental psychopathology. *Development and Psychopathology*, *15*, 581–612.

Curran, P. J., & Wirth, R. J. (2004). Interindividual differences in intraindividual variation: Balancing internal and external validity. *Measurement*, *2*, 219–247.

Eggleston, E. P., Laub, J. H., & Sampson, R. J. (2004). Methodological sensitivities to latent class analysis of long-term criminal trajectories. *Journal of Quantitative Criminology*, *20*, 1–26.

Emde, R. N., & Spicer, P. (2000). Experience in the midst of variation: New horizons for development and psychopathology. *Development and Psychopathology*, *12*, 313–331.

Flaherty, B. P. (2008). Testing the degree of cross-sectional and longitudinal dependence between two discrete dynamic processes. *Developmental Psychology*, *44*, 468–480.

Gilliom, M., & Shaw, D. S. (2004). Codevelopment of externalizing and internalizing problems in early childhood. *Development and Psychopathology*, *16*, 313–333.

Gottleib, G., & Halpern, C. T. (2002). A relational view of causality in normal and abnormal development. *Development and Psychopathology*, *14*, 421–435.

Hershberger, S. L. (1998). Dynamic factor analysis. In G. A. Marcoulides (Ed.), *Modern methods for business research*. Mahwah, NJ: Erlbaum.

Hirsh-Pasek, K., & Burchinal, M. (2006). Mother and caregiver sensitivity over time: Predicting language and academic outcomes with variable- and person-centered approaches. *Merrill–Palmer Quarterly*, *52*, 449–485.

Horn, J. L. (2000). Comments on integrating person-centered and variable-centered research on problems associated with the use of alcohol. *Alcoholism: Clinical and Experimental Research*, *24*, 924–930.

Jackson, K. M., & Sher, K. J. (in press). Comparison of longitudinal phenotypes based on alternate heavy drinking cut scores: A systematic comparison of trajectory approaches III. *Psychology of Addictive Behaviors*.

Jones, K. (1991). The application of time series methods to moderate span longitudinal data. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change* (pp. 75–87). Washington, DC: American Psychological Association.

Kagan, J. (1994). *Galen's prophecy: Temperament in human nature*. New York: Basic Books.

Kaplan, D. (2008). An overview of Markov chain methods for the study of stage-sequential developmental processes. *Developmental Psychology*, *44*, 457–467.

Kelderman, H., & Molanaar, P. C. M. (2007). The effect of individual differences in factor loadings on the standard factor model. *Multivariate Behavioral Research*, *42*, 435–456.

Keller, T. E., Spieker, S. J., & Gilchrist, L. (2005). Patterns of risk and trajectories of preschool problem behaviors: A person-oriented analysis of attachment in context. *Development and Psychopathology*, *17*, 349–384.

Langeheine, R. (1994). Latent variable Markov models. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research* (pp. 373–395). Beverly Hills, CA: Sage.

Langeheine, R., & van de Pol, F. (1990) A unifying framework for Markov modeling in discrete space and discrete time. *Sociological Methods and Research*, *18*, 416–441.

Lubke, G., & Neale, M. (2006). Distinguishing between latent classes and continuous factors: Resolution by maximum likelihood? *Multivariate Behavioral Research*, *41*, 499–532.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*, 19–40.

Magnusson, D. (1985). Implications of an interactional paradigm for research on human development. *International Journal of Behavioral Development*, *8*, 115–137.

Magnusson, D. (1998). The logic and implications of a person-oriented approach. In R. B. Cairns, L. R. Bergman, & J. Kagan (Eds.), *Methods and models for studying the individual* (pp. 33–62). London: Sage.

Magnusson, D., & Torestad, B. (1993). A holistic view of personality: A model revisited. *Annual Review of Psychology*, *44*, 427–452.

Mannan, H. R., & Koval, J. J. (2003). Latent mixed Markov modelling of smoking transitions using Monte Carlo bootstrapping. *Statistical Methods in Medical Research*, *12*, 125–146.

Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, *5*, 23–43.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107–122.

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

Molenaar, P. C. M. (1994). Dynamic latent variable models in developmental psychology. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research*. London: Sage.

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

Molenaar, P. C. M. (2007). Psychological methodology will change profoundly due to the necessity to focus on intra-individual variation. *Integrative Psychological & Behavioral Science*, *41*, 35–40.

Molenaar, P. C. M., de Gooijer, J. G., & Schmitz, B. (1992). Dynamic factor analysis of nonstationary multivariate time series. *Psychometrika*, *57*, 333–349.

Molenaar, P. C. M., & von Eye, A. (1994). On the arbitrary nature of latent variables. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research* (pp. 226–242). Thousand Oaks, CA: Sage.

Mun, E. Y., Windle, M., & Schainker, L. M. (2008). A model-based cluster analysis approach to adolescent problem behaviors and young adult outcomes. *Development and Psychopathology*, *20*, 291–318.

Muthén, B. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1–33). Mahwah, NJ: Erlbaum.

Muthén, B., & Asparouhov, T. (2006). Growth mixture modeling: Analysis with non-Gaussian random effects. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Advances in longitudinal data analysis*. Boca Raton, FL: Chapman & Hall/CRC Press.

Muthén, B. O. (2007). Latent variable hybrids: Overview of new and old models. In G. R. Handcock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 1–24.) Charlotte, NC: Information Age Publishing.

Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*, 882–891.

Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*, 463–469.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric group-based approach. *Psychological Methods*, *4*, 139–157.

Nagin, D. S. (2005). *Group-based modeling of development* (pp. 1–78). Cambridge, MA: Harvard University Press.

Nagin, D. S., & Land, K. C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology*, *31*, 327–362.

Nagin, D. S., Pagani, L., Tremblay, R. E., & Vitaro, F. (2003). Life course turning points: The effect of grade retention on physical aggression. *Development and Psychopathology*, *15*, 343–361.

Nagin, D. S., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods 6*, 18–34.

Nesselroade, J. R., Gerstorf, D., Hardy, S. A., & Ram, N. (2007). Idiographic filters for psychological constructs. *Measurement*, *5*, 217–235.

Nesselroade, J. R., McArdle, J. J., Aggen, S. H., & Meyers, J. M. (2002). Dynamic factor analysis models for representing process in multivariate time-series. In D. S. Moskowitz & S. L. Hershberger (Eds.), *Modeling intraindividual variability with repeated measures data: Methods and applications* (pp. 235–265). Mahwah, NJ: Erlbaum.

Nesselroade, J. R., & Molenaar, P. C. M. (1999). Pooling lagged covariance structures based on short, multivariate time series for dynamic factor

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press

analysis. In R. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 223–250). London: Sage.
Nesselroade, J. R., & Molenaar, P. C. M. (2003). Quantitative models for developmental processes. In J. Valsiner & K. J. Connolly (Eds.), *Handbook of developmental psychology* (pp. 622–639). London: Sage.
Nurious, P. S., & Macy, R. J. (2008). Heterogeneity among violence-exposed women: Applying person-oriented research methods. *Journal of Interpersonal Violence*, *23*, 389–415.
Nylund, K. (2007). *Latent transition analysis: Modeling extensions and an application to peer victimization.* Doctoral dissertation, University of California, Los Angeles.
Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*, 437–448.
Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., & Senecal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of Educational Psychology*, *99*, 734–746.
Reinecke, J. (2006). Longitudinal analysis of adolescents' deviant and delinquent behavior. *Measurement*, *2*, 100–112.
Rutter, M. (1996). Developmental psychopathology as an organizing research construct. In D. Magnusson (Ed.), *The lifespan development of individuals: Behavioral, neurobiological, and psychosocial perspectives* (pp. 394–413). New York: Cambridge University Press.
Sameroff, A. J. (1982). Development and the dialectic: The need for a systems approach. In W. A. Collins (Ed.), *The concept of development* (pp. 83–103). Mahwah, NJ: Erlbaum.
Sameroff, A. J., & Mackenzie, M. J. (2003). Research strategies for capturing transactional models of development: The limits of the possible. *Development and Psychopathology*, *15*, 613–640.
Schaeffer, C. M., Petras, H, Ialongo, N., Poduska, J., & Kellam, S. (2003). Modeling growth in boys' aggressive behavior across elementary school: Links to later criminal involvement, conduct disorder, and antisocial personality disorder. *Developmental Psychology*, *39*, 1020–1035.
Schmittmann, V. D., Dolan, C. V., van der Maas, H. L. J., & Neale, M. (2005). Discrete latent Markov models for normally distributed response data. *Multivariate Behavioral Research*, *40*, 461–488.
Shifren, K., Hooker, K., Wood, P., & Nesselroade, J. R. (1997). Structure and variation of mood in individuals with Parkinson's disease: A dynamic factor analysis. *Psychology and Aging*, *12*, 328–339.
Sroufe, A., & Rutter, M. (1984). The domain of developmental psychopathology. *Child Development*, *55*, 17–29.
Sterba, S. K., Prinstein, M. J., & Cox, M. J. (2007). Trajectories of internalizing problems across childhood: Heterogeneity, external validity, and gender differences. *Development and Psychopathology*, *19*, 345–366.
Troop-Gordon, W., & Ladd, G. W. (2005). Trajectories of peer victimization and perceptions of the self and schoolmates: Precursors to internalizing and externalizing problems. *Child Development*, *76*, 1072–1091.
Verbeke, G., & Lesaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association*, *91*, 217–221.
von Eye, A., & Bergman, L. R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology*, *15*, 553–580.
von Eye, A., & Bogat, K. (2006). Person-oriented and variable-oriented research: Concepts, results, and development. *Merrill–Palmer Quarterly*, *52*, 390–420.
Walls, T. A., & Schafer, J. L. (2006). *Models for intensive longitudinal data*. New York: Oxford University Press.
Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scale. *Journal of Personality and Social Psychology*, *54*, 1063–1070.
Wiggins, L. M. (1973). *Panel analysis*. Amsterdam: Elsevier.
Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, *116*, 363–381.
Windle, M. (2000). A latent growth curve model of delinquent activity among adolescents. *Applied Developmental Science*, *4*, 193–207.
Wood, P., & Brown, D. (1994). The study of intraindividual differences by means of dynamic factor models: Rationale, implementation, and interpretation. *Psychological Bulletin*, *116*, 166–186.
Zahn-Waxler, C., Klimes-Dougan, B., & Slattery, M. J. (2000). Internalizing problems of childhood and adolescence: Prospects, pitfalls, and progress in understanding the development of anxiety and depression. *Development and Psychopathology*, *12*, 443–466.
Zhang, Z. (2006). *Codes for Mplus using MLE based on block-Toeplitz to estimate dynamic factor models*. Retrieved July 14, 2008, from http://www.psychstat.org/us/article.php/71.htm

https://doi.org/10.1017/S0954579410000015 Published online by Cambridge University Press