Measuring Change:  
What Individual Growth  
Modeling Buys You

John B. Willett  
Harvard University

The measurement of change is critically important in psychology and the social sciences because when people acquire new skills, when they learn something new, when they grow intellectually and physically, when their attitudes and interests develop, they change in fundamental and interesting ways. It is only by being able to measure change over time accurately that we can map out critical time-dependent phenomena at the core of human development. Unfortunately, despite its obvious importance, there has been much controversy over the years about the measurement of change. In the past, influential methodologists have examined the technical properties of change measurement and convinced themselves, and everyone else, that it is impossible to measure change well. But recent developments demonstrate that not only can change be measured, but, with a little foresight and planning, it can be measured extremely well.

Those widely publicized and false conclusions about the many difficulties of measuring change are rooted in a simple misconception—that individual change should be viewed only as an increment, as the difference between “before” and “after.” Early methodologists viewed each person who was changing as having acquired a “quantum” of achievement, or attitude, or value, or whatever, between the “pre” and the “post” measurement. They argued that investigators should be concerned only with the size of the acquired chunk. This perception is mistaken. Individual change takes place continuously over time and must not be viewed as a “before”
and “after” phenomenon. In fact, it is our failure to perceive of change as a continuous process of individual development over time that has hamstrung the creation of decent statistical methods for its measurement.

Recently, however, a clearer day has dawned. Methodologists have finally recognized that individual change can be measured very well indeed, providing that measurement is moved beyond the limitations of the “before and after,” or two wave, research design. The continuous process of human development can be easily and efficiently documented simply by measuring each person repetitively over extended periods of time. In this chapter, I introduce this multiwave perspective and I argue that, with a little planning, our questions about change can be answered easily. Because there has been so much controversy, I begin by commenting briefly on traditional two-wave methods for the measurement of change.

TRADITIONAL METHODS OF MEASURING CHANGE: TWO-WAVE DATA AND THE DIFFERENCE SCORE

Traditionally, in research on change, investigators collected only two waves of data on each person in the sample. That is, they observed each person’s “status”—their achievement, attitude, or any other attribute of empirical interest—at the beginning and the end of the investigation period and they used these “pre” and “post” data to construct a two-wave within-person summary of individual change. Then, to figure out whether change was systematically related to other variables that described the individual’s background, training or treatment, the within-person change summaries were simply regressed on, or correlated with, variables selected to describe the background, training, and treatment. These latter between-person analyses were usually referred to generically as “the detection of correlates and predictors of change.”

This approach seems sensible and straightforward, but its simplicity masks hidden pitfalls. For one thing, for the strategy to be effective, the researcher must be able to construct a decent measure of individual change for everyone in the sample from the two waves of obtained data. Methodologists differ as to how this should be done, and psychometric history is littered with strategies proposed for the purpose. Some of these strategies are better than others, but all are inferior to the multiwave approach introduced later in this chapter.

Measures of Individual Change

Possibly the simplest two-wave measure of change is the observed difference score, obtained by subtracting the initial measurement from the final, for each person. Originally the difference score was highly favored but then
fell into disrepute, being much maligned through the 1960s and 1970s. This led to the birth of a coterie of alternative two-wave measures of change, including the residual change score. Although I do not advocate the use of two-wave change-measurement in practice, I comment briefly here on these measures in order to clarify the controversy that continues to swirl.

Before we can engage in a productive discussion of strategies for the measurement of change, however, a critical distinction must be made—the classical distinction among observed, true, and error scores. Because of the uncompromising vicissitudes of nature, when a test or rating instrument is administered to an individual, the observed measurement combines a measure of the person’s true capability with whatever random error happens to accompany the measurement. Of course, our research interest is usually focused squarely on true status—a commodity that, if measurement error is large, may differ considerably from observed status. Similarly, when a group of people are changing over time on some important attribute, it is not the fallible observed changes that are of critical interest but the underlying true changes. Measures of observed change simply provide a fallible lens through which we hope the hidden nature of true change may be discerned.

From this perspective, the observed difference score is a reasonable commodity. It is intuitively appealing, easy to compute, and is an unbiased estimator of the underlying true change. Despite these clear and unequivocal advantages, however, it has been resoundingly criticized for its purported unreliability and its correlation with initial status—criticisms that, it is now clear, were completely unwarranted.

Some have said that the difference score is always unreliable; others argued that it cannot be both reliable and valid simultaneously (Bereiter, 1963; Linn & Slinde, 1977). In general, neither of these claims is correct. They are misconceptions arising from misinterpretation of the concepts of reliability and validity in the context of change over time (Rogosa, Brandt, & Zimowski, 1982; Rogosa & Willett, 1983, 1985; Willett, 1988). Principally, by misinterpreting the observed pretest–posttest correlation as an index of construct validity in expressions for the reliability of the difference score, authors have misleadingly focused on hypothetical situ-
ations in which there is little variation in true change across people. By choosing to examine situations in which the interperson variability in true change is close to zero, they have ensured that the numerator of the expression for the reliability of the difference score (which is defined as the ratio of the variances of true and observed change) is very small and so the reliability is almost zero. In reality, in perfectly ordinary situations, the reliability of the difference score can be quite respectable (Rogosa & Willett, 1983). And when interperson variability in true change is large, the reliability of the difference score can even be greater than the reliabilities of the constituent pretest and posttest scores (Willett, 1988).

Anyway, even if the difference score were always unreliable, this would not necessarily be a problem for the measurement of within-person change. Low difference score reliability does not imply unilaterally that within-person change has been measured imprecisely. Low reliability often occurs in practice because most of the people in the sample are changing at about the same rate (especially over the short-term). So even though the 20 points (say) that everyone has changed can be measured very precisely, the changes of different people cannot be distinguished from one another and so the difference score appears unreliable. This problem of interpretation does not call the difference score itself into question—we know quite precisely that everyone has changed by 20 points—but it does undermine the credibility of reliability as a worthwhile indicator of measurement quality.

The difference score has also been falsely condemned for at least three other reasons, all of which originate in critics’ misunderstanding of the association between it and pretest status (Linn & Slinde, 1977; Plewis, 1985). If there is a positive correlation between change and pretest status, for instance, then people with high initial status will tend to have high gains subsequently. If the correlation is negative, then people with low initial scores will tend to change more rapidly than those with high initial scores. This has led some critics to claim that any measure of change that is not independent of initial status must be unfair because it gives “an advantage to persons with certain values of the pretest scores” (Linn & Slinde, 1977, p. 125). But such prejudice is not reasonable—why should change and status be unrelated? The intimate connection between change and status is a consequence of growth history. Current status is a product of prior change; current change determines future status. A correlation between change and status is an almost inevitable fact of life.

Second, investigators have estimated the correlation between change and initial status empirically and have worried because their findings often disagree with each other. Even when investigating the same popu-
lation in the same setting with the same measures, some find the correlation to be positive, some zero, and some strongly negative. But why should they expect a single value for this correlation? When different people are changing in different ways, individual trajectories are likely to crisscross with time and the correlation between change and initial status can fluctuate markedly as different times are selected as the occasion for "initial" measurement (see Rogosa & Willett, 1985). Unless some important occasion can be agreed upon substantively to be the initial time, researchers should expect to disagree when they ask what the correlation between change and initial status is, because the answer, of course, is that it depends on the time that you define as being initial.

Third, the difference score has been inappropriately criticized because some people have convinced themselves that its correlation with pretest score is always negative. In reality, this claim is false (for an example of a positive correlation between the difference score and change, see Thorndike, 1966). Those that condemn the difference score in this way have usually committed one of two mistakes. Either they have spuriously created the negative correlation themselves by standardizing pretest and posttest scores to the same standard deviation before computing the difference (an ill-advised process that should be avoided because it destroys legitimate growth information), or they are being confused by the vagueness of statistical estimation. In this latter case, they have typically used the sample correlation of observed initial status and the difference score as an estimate of the population correlation of true initial status and true change (the real parameter of interest). Unfortunately, because pretest measurement error appears with a negative sign in the difference score, this estimator is negatively biased. It is often negative even when the underlying true correlation is positive. However, the bias is easily corrected (see Willett, 1988) and anyway, as Rogosa et al. (1982) note, an unbiased estimate of within-person change—the difference score—should not be rejected because a poorly-conceived estimator of the association between true change and true initial status is biased. It is the latter that needs fixing, not the former.

The creation of the residual-change score was motivated by an unnecessary desire to create measures of change that were uncorrelated with pretest score. Residual-change scores are obtained by estimating the residuals that would be obtained in the population regression of true final status on true initial status. They are intended to describe the true change that each person would have experienced, if everyone had "started out equal." Various methods have been proposed for their computation (see Rogosa et al., 1982). Much energy has been dissipated in the psychometric literature detailing the properties of the many estimators of residual
change and considerable argument has been aroused. When discussing residual-change scores, methodologists disagree as to exactly what is being estimated, how well it is being estimated and how it can be interpreted. In addition to the many technical and practical problems that arise in the empirical application of residual change scores, there also remain unresolved issues of logic and substance that are too extensive to detail here (but, see Rogosa et al., 1982; Rogosa & Willett, 1985; Willett, 1988). I strongly advise the researcher to avoid residual change scores as measures of within-person change.

Detecting Correlates and Predictors of Change

Once a two-wave measure of change has been computed for each person, the relationship between change and other "background" variables is often investigated by the common-or-garden methods of correlation and regression analysis. For instance, to find out if changes in reading achievement are related to gender, the data analyst might simply correlate pre/post differences in test score with a dummy variable representing gender. Unfortunately, although straightforward, this rudimentary strategy is flawed. As we have seen, the difference score is a fallible measure of change that contains both the true change and measurement error. This latter random noise attenuates the between-person findings, leading the obtained sample correlations to underestimate the true relationship between change and the covariates. This problem can be avoided by correcting between-person analyses for the fallibility of the difference score (for methods of correction, see Willett, 1988).

Ironically, of course, this process of disattenuation requires access to information in addition to the two available waves of measured data, usually in the form of an external estimate of the reliability of the pretest and posttest scores. However, if the additional information is available then it becomes possible to detect relationships between true change and the correlates of change. The lesson here is not that the disattenuation is possible but that acceptable between-person analyses of two-wave data can only be conducted if the researcher possesses other information in addition to the pair of observed pretest and posttest scores. This emphasizes the real and fundamental weakness of the two-wave design—the collection of pretest and posttest data provides insufficient information to answer interesting questions about individual change over time. To do a decent job, additional data are always required and the best way to collect those data is not through the estimation of reliability at pretest and posttest but through the collection of more than two waves of data and the use of individual growth modeling.
MODERN METHODS OF MEASURING CHANGE: MULTIWAVE DATA AND INDIVIDUAL GROWTH MODELING

So, taking a "snapshot" of a person's observed status "before" and "after" is not the best way to reveal the intricacies of his or her progress. Changes may be occurring smoothly over time with some complex and substantively interesting trajectory, and the crude measurement of pretest and posttest scores just cannot reveal the interesting features of that trajectory. To do a good job of mapping out individual temporal trajectories, a truly longitudinal perspective must be adopted—a sample of people must be followed carefully over time, with multiple "waves" of data being collected on their status at sensibly spaced intervals. If the attribute of interest is changing steadily and smoothly over a long period of time, perhaps three or four widely spaced measurements on each person may be sufficient to capture the shape and direction of the change. But, if the trajectory of individual change is complex, then many more closely spaced measurements may be required (see Willett, 1989).

A word of caution, however: Like many other common forms of statistical analysis, the individual growth modeling approach is only applicable if it truly makes intuitive sense to measure change in the outcome variable. At the very least, this means that the outcome variable must have three properties. First, it must be a continuous variable, at either the interval or ratio level. Second, it must be equatable from occasion to occasion—that is, each scale point on the measure must retain an identical meaning as time passes. Third, it must remain construct valid for the entire period of observation. If any of these conditions are violated, then the method of individual growth modeling is being applied inappropriately.

Two Types of Question That Can Be Asked About Change

As an example of the type of longitudinal data that are required to measure change, Table 11.1 contains 5 waves of longitudinal data on the continuous dependent variable, tolerance of deviant behavior, for a small subsample of 16 youths drawn randomly from a larger sample of 168 adolescents in the National Youth Survey. Every year of the study—at ages 11, 12, 13, 14, and 15—each participant completed a 9-item instrument that asked whether it was wrong for someone their age to do one of the following: cheat on tests, purposely destroy property of others, use marijuana, steal something worth less than $5, hit or threaten someone

---

2I thank Stephen W. Raudenbush for providing these data.
### TABLE 11.1
Tolerance of Deviant Behavior as a Function of Age

<table>
<thead>
<tr>
<th>Subject ID Number</th>
<th>Age at Reported Tolerance of Deviant Behavior</th>
<th>Predictors of Change</th>
<th>Exposure to Deviant Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>0009</td>
<td>2.23</td>
<td>1.79</td>
<td>1.90</td>
</tr>
<tr>
<td>0045</td>
<td>1.12</td>
<td>1.45</td>
<td>1.45</td>
</tr>
<tr>
<td>0268</td>
<td>1.45</td>
<td>1.34</td>
<td>1.99</td>
</tr>
<tr>
<td>0314</td>
<td>1.22</td>
<td>1.22</td>
<td>1.55</td>
</tr>
<tr>
<td>0442</td>
<td>1.45</td>
<td>1.99</td>
<td>1.45</td>
</tr>
<tr>
<td>0514</td>
<td>1.34</td>
<td>1.67</td>
<td>2.23</td>
</tr>
<tr>
<td>0669</td>
<td>1.79</td>
<td>1.90</td>
<td>1.90</td>
</tr>
<tr>
<td>0624</td>
<td>1.12</td>
<td>1.12</td>
<td>1.22</td>
</tr>
<tr>
<td>0723</td>
<td>1.22</td>
<td>1.34</td>
<td>1.12</td>
</tr>
<tr>
<td>0918</td>
<td>1.00</td>
<td>1.00</td>
<td>1.22</td>
</tr>
<tr>
<td>0949</td>
<td>1.99</td>
<td>1.55</td>
<td>1.12</td>
</tr>
<tr>
<td>0978</td>
<td>1.22</td>
<td>1.34</td>
<td>2.12</td>
</tr>
<tr>
<td>1105</td>
<td>1.34</td>
<td>1.90</td>
<td>1.99</td>
</tr>
<tr>
<td>1542</td>
<td>1.22</td>
<td>1.22</td>
<td>1.99</td>
</tr>
<tr>
<td>1552</td>
<td>1.00</td>
<td>1.12</td>
<td>2.23</td>
</tr>
<tr>
<td>1653</td>
<td>1.11</td>
<td>1.11</td>
<td>1.34</td>
</tr>
</tbody>
</table>

*Note.* Longitudinal data on a random subsample of 16 youths drawn from the National Youth Survey, with: (a) 5 annual measurements of perceived tolerance of deviant behavior, (b) scores on two potential predictors of change: gender (0 = male, 1 = female), and perceived exposure to deviant behavior during adolescence (assumed time-invariant).

without reason, use alcohol, break into a building or vehicle to steal, sell hard drugs, or steal something worth more than $50 (Raudenbush & Chan, 1992). Responses to each item were registered on a 4-point scale with higher scores indicating greater tolerance for deviant behavior (1 = very wrong, 2 = wrong, 3 = a little bit wrong, 4 = not wrong at all) and were averaged across items to provide a scale score. The longitudinal data in Table 11.1—and in the full data set—allow us to answer “within-child” research questions about tolerance of deviant behavior. For instance, inspection of the “Age 11” column in the table indicates that children seem to enter the study differing in levels of tolerance; at age 11, for instance, Child #1552 is not very tolerant of deviant behavior at all, whereas Child #0009 is quite tolerant. Then, inspection of scores across columns 2 through 6 suggests some children, such as Child #0514, become more tolerant of deviant behavior as they age, whereas other children, such as Child #0624, seem to stay at about the same level of tolerance. Overall, although there is considerable scatter in the observed scores over time and across children, the table seems to indicate that adolescents become gradually more tolerant of deviant behavior as they age.
11. MEASURING CHANGE

The table also contains information on two potential predictors of change: (a) $G_p$, the gender of the $p$th adolescent ($0 = \text{male}$, $1 = \text{female}$), and (b) $E_p$, individual $p$'s reported exposure to deviant behavior in the first year of data collection, at age 11.\textsuperscript{3} Values of these predictors distinguish children from each other and allow us to ask between-child questions about how children's change in tolerance differs from child to child. For instance, if we were to conduct an analysis of data from the entire sample of 168 youths, we could answer questions like: Is change in the tolerance of deviant behavior related to gender? Do children who are more heavily exposed to deviant behavior at age 11 tend to become tolerant of deviant behavior more rapidly than children who were less heavily exposed?

Notice that, in inspecting the data, I am distinguishing questions at two linked levels—the within-child level (which, in keeping with standard terminology, we refer to as level-1) and the between-child level (which we refer to as level-2). At level-1, we focus on the individual and try to understand and summarize the changes that are occurring to each person under study. These are questions about intraindividual change: How rapidly is each child changing? What is the shape of each child's trajectory of change? At level-2, we ask questions about how the individual changes differ from person to person. These are questions about interindividual differences in change: Are interindividual differences in change of tolerance toward deviance during adolescence related to respondent gender and initial exposure to deviant behavior? These two levels—the within-person and the between-person—constitute a natural hierarchy and, as we see, they provide an important framework for the measurement of change.

Simple Exploratory Methods for Answering Questions About Change

One obvious exploratory way of answering level-1 within-person questions is to examine each child's change over time separately by plotting an empirical growth-trajectory for each child—that is, for each child, create a graph of observed status displayed against time. I illustrate this in Fig. 11.1 for youth #0514 of Table 11.1. In the figure, observed tolerance of deviant behavior is plotted on the vertical axis and the child's age (in years) is plotted on the horizontal axis.

\textsuperscript{3}Adolescents' exposure to deviance was also self-reported using a 9-item instrument in each of the 5 years of the study. Participants were asked how many of their peers engaged in the same nine activities identified in the Tolerance of Deviant Behavior instrument. For each item, ratings were obtained on a 5-point scale which ranged from 0 (none of them) to 4 (all of them). Scale scores were computed by averaging across items. In these analyses, to simplify the presentation, I have used only the participant's initial scale score at age 11 as a predictor of change.
Notice that there is a trend evident in the data—the observed tolerance of deviant behavior of Child #0514 seems to be generally increasing with age. I have summarized this trend by superimposing a within-child trend line on the plot, obtained by simply regressing tolerance of deviant behavior on age, for this child alone. In displaying the summary trend line, I have assumed that change in tolerance of deviant behavior is linear with age. I have made this assumption because inspection of the data suggested that it was a reasonable decision. Consequently, the steeper the slope of the summary trend line, the faster Child #0514’s tolerance of deviant behavior is increasing with age—in other words, when growth is linear, the slope of the child’s summary trend line acts as a measure of individual change.

Of course, as we will see later, we are not limited to the selection of a linear individual growth summary. Many other possible mathematical
functions are also available—both those that depend linearly on time, and those that do not. Choice of an appropriately shaped trajectory to represent true individual change is an important first step in any analysis of change. Ideally, theory will guide the rational choice of trajectory so that subsequent analyses have meaningful substantive interpretations. Often, however, the mechanisms governing the change process are poorly understood and so a low-order polynomial—a linear or a quadratic curve, for instance—is used to approximate the trajectory. Also, in much social and psychological research, only a restricted portion of the individual's lifespan is observed and few waves of data are collected, and so the selected trajectory must be mathematically simple if analyses are to be conducted successfully. Thus, the most popular trend used to summarize individual change over time is often a linear function of time, as here.

The empirical growth trajectories of everyone under investigation can be collected together informatively in a single picture. This provides a simple and straightforward way of exploring level-2 questions about between-person differences in change because eyeball comparisons of empirical trajectories across people can help detect systematics in the way that the individual growth trajectories differ from person to person. Figure 11.2 displays a collection of empirical growth trajectories for the 16 youths in Table 11.1. In this collection of empirical growth trajectories, the observed data themselves have been omitted in order to avoid clutter. Notice that there is evidence of heterogeneity in observed change across children—in some, tolerance of deviant behavior increases with age, in some tolerance remains fairly stable over adolescence, and some even become less tolerant of deviant behavior as they age. In Fig. 11.2, I have also coded the empirical growth trajectories to represent a child's exposure to deviant behavior at age 11 (dashed line = exposure below the subsample median, solid line = exposure above the subsample median). Notice that heavily-exposed children tend to become tolerant of deviant behavior more rapidly over time. Similar plots could be created to display the effects of gender, and of other interesting predictors.

Choosing a Level-1 Statistical Model to Represent Intra-Individual Change

These simple exploratory ideas can be formalized to provide a rigorous statistical basis for the measurement of change. To measure change with multiwave data, a pair of linked statistical models must be specified, one for each level in the hierarchy (Bryk & Raudenbush, 1987; Rogosa & Willett, 1985).

At level-1, individual change over time is represented by specifying an individual growth model that describes the temporal dependence of
FIG. 11.2. A collection of OLS-fitted growth trajectories in tolerance of deviant behavior between ages 11 and 15 for the subsample of 16 randomly-selected adolescents whose empirical growth records are displayed in Table 11.1.

Individual status on time. In drawing the empirical growth trajectories in the example in Figs. 11.1 and 11.2, for instance, a linear function of time was chosen as a valid and reasonable representation of individual change in tolerance of deviant behavior as a function of age. In this case, the observed tolerance of deviant behavior $Y_{ip}$ of youth $p$ on the $i$th occasion of measurement (that is, at age $t_i$) is represented by:

$$Y_{ip} = [\pi_{tp} + \pi_{tp} (t_i - 11)] + \varepsilon_{ip},$$  \hspace{1cm} (1)

where the structural component of the model, representing the dependence of true status on time, has been bracketed to separate it from $\varepsilon_{ip}$, the random error that accrues on each measurement.
The structural part of an individual growth model contains unknown constants referred to as the individual growth parameters, whose values determine the trajectory of true individual change over time. In equation 1, for instance, there are two such parameters: \( \pi_{0p} \) and \( \pi_{1p} \). The second parameter \( \pi_{1p} \) is the slope of the straight-line growth model and represents the rate at which person \( p \)'s true tolerance of deviant behavior is growing over time (that is, it is the increase in true tolerance per unit increase in age—see Fig. 11.3). If \( \pi_{1p} \) is positive, then person \( p \)'s true status is increasing with time, if it is negative then person \( p \)'s true status is decreasing with time. The other parameter \( \pi_{0p} \) is the intercept of the straight-line growth model, representing person \( p \)'s true tolerance of deviant behavior at age 11 (that is the true initial tolerance of person \( p \) on entry into the study). In Fig. 11.3, I present a plot of the true linear trajectory corre-

![Graph showing true tolerance over age with \( \pi_{0p} \) and \( \pi_{1p} \) marked.]
sponding to equation 1 with graphical interpretations of the two individual growth parameters included.

Notice that, in this model, the investigator can easily control the substantive interpretation of the intercept parameter, \( \pi_{0p} \). By subtracting 11 years from the child’s age before multiplying the individual slope parameter in the model (see equation 1), the origin of the time axis has been reset (or recentered) to an age of 11 years. This provides the individual intercept parameter with an interpretation that is interesting substantively in the context of this study—it represents true tolerance on entry into the study at age 11 (had the time axis not been recentered, and its origin left at age zero, then the intercept parameter would have represented true tolerance of deviant behavior at birth). In other investigations, other temporal recenterings or reparameterizations may prove more substantively reasonable.

Choosing a Level-2 Statistical Model to Represent Interindividual Differences in Change

Under the individual growth modeling strategy, once a level-1 model has been selected to represent change in a particular domain, everyone in the sample is assumed to have the same generic functional form for their growth but different people can have different values of the individual growth parameters. For instance, as in the case of Fig. 11.2, when within-person change is linear, individuals may differ in both their intercepts and slopes. More interestingly, the individual growth parameters may also differ from person to person in a systematic way—that is, in ways related to variables that describe critical attributes of the people involved (the correlates and predictors of change). Under the straight-line growth model for tolerance that we have adopted, for instance, the investigator can ask: Is there a systematic relationship between true initial tolerance and the predictors gender and exposure? Between true rate of change in tolerance and these same predictors?

Hypothesized relationships between individual growth parameters and any predictors of change can be formalized in level-2 statistical models for between-person differences in change (Bryk & Raudenbush, 1987; Rogosa & Willett, 1985). Equation 2, for instance, presents level-2 models for the true intercept and true slope of the straight-line tolerance growth model in equation 1 as a function of the predictors of change listed in Table 11.1:

\[
\begin{align*}
\pi_{0p} &= \gamma_{00} + \gamma_{01}\text{GENDER}_p + \gamma_{02}\text{EXPOSURE}_p + u_{0p} \\
\pi_{1p} &= \gamma_{10} + \gamma_{11}\text{GENDER}_p + \gamma_{12}\text{EXPOSURE}_p + u_{1p}
\end{align*}
\] (2)
where \( u_{ip} \) and \( u_{ip} \) represent the level-2 residuals (those parts of \( \pi_{ip} \) and \( \pi_{ip} \) that are "unexplained" by the two selected predictors of change).

In the level-2 models, the \( \gamma \) coefficients summarize the population relationship between the individual growth parameters and the selected predictors of change. They can be interpreted in the same way as regular regression coefficients—so, a non-zero value of \( \gamma_{\theta_1}, \gamma_{\theta_2}, \gamma_{\theta_3}, \) or \( \gamma_{\theta_4} \) indicates that the corresponding covariate is a predictor of true initial tolerance or true rate of change in tolerance, respectively. For instance, if the initial (age 11) tolerance of girls tends to be higher than that of boys (i.e., if they have larger values of \( \pi_{ip} \)) then \( \gamma_{\theta_1} \) will be positive (since GENDER = 1 for females). If youths who had greater exposure to deviant behavior at age 11 have higher rates of change in tolerance then \( \gamma_{\theta_2} \) will be positive. And so on.

**Methods of Model Estimation**

In an investigation of change, the within-person and between-person statistical models—like those in equations 1 and 2—are fitted to longitudinal multiwave data and their parameters estimated and interpreted. A variety of methods are available for carrying out this model-fitting and parameter estimation. Some of these methods are very straightforward and can easily be implemented on popular commercially-available statistical computer packages; others are more sophisticated and require dedicated computer software. Here, in order of increasing complexity, I briefly review a taxonomy of these data-analytic strategies.

*A Simple Two-Step Exploratory Strategy—Estimating the Individual Growth Parameters in Separate Level-1 Regression Analyses.* Perhaps the simplest analytic approach is to fit the level-1 individual growth model to the observed growth data by ordinary least-squares (OLS) regression analysis separately for each individual in the data set (in the same way that fitted trend-lines were created earlier for Figs. 11.1 and 11.2). This person-by-person growth modeling provides estimates of the individual growth parameters for each person which can then be collected together to become the dependent variables in subsequent level-2 between-person data-analyses. For instance, in the case of the straight-line growth model of equation 1, we can first obtain individual intercept and slope estimates to represent the initial tolerance and rate of change in tolerance by regressing observed tolerance on age (minus 11 years; see equation 1) for each of the youths in the sample. These estimates can then be collected together and related directly to GENDER and EXPOSURE, or other predictors, in follow-up level-2 correlation or regression analyses. This strategy is straightforward and easy to implement, and the OLS-estimated
intercept and growth rates provide more precise measurement of individual change than is possible with the difference score (see below).

**Improving the Two-Step Exploratory Strategy—Accounting for Between-Person Heterogeneity in the Precision of the Individual Growth Parameter Estimates.** Due to the idiosyncrasies of measurement, some people may have empirical growth records whose entries are smoothly ordered and for whom the growth data fall very close to the underlying true growth trajectory. Other less fortunate people have more erratic growth records and their data points are scattered, deviating widely from the underlying growth trajectories. These differences in scatter affect the precision (the standard errors) with which the level-1 individual growth parameters can be estimated. In the case of our example, youths with “smooth and systematic” growth records will have highly precise estimates of initial tolerance and rate of change in tolerance (that is, their parameter estimates will have smaller standard errors); youths with “erratic and scattered” observed growth records will have less precise intercept and slope estimates (that is, larger standard errors). The level-2 analyses of the relationships between the estimated individual growth parameters and the predictors of change will be improved (made asymptotically efficient) if this between-person variation in the precision of the first-round growth parameter estimates is taken into account (Hanushek, 1974). To achieve this, we use weighted least-squares (WLS) regression analysis to fit the corresponding level-2 models using weights that are inversely related to the standard errors of the first-round parameter estimates (examples of the application of this strategy and an expression for appropriate weights can be found in Willett & Ayoub, 1991, and Willett, 1988, respectively). This approach is an improvement on the unweighted analysis because it ensures that individuals whose growth parameters are determined most precisely play a more important role in determining the outcomes of the level-2 analyses.4

**Using Dedicated Computer Software to Fit the Within-Person and Between-Person Statistical Models Simultaneously.** In this chapter, I have argued that high quality measurement of change is possible via the collection of longitudinal data and the fitting of within-person and

---

4 Rogosa and his colleagues (Rogosa et al., 1982; Rogosa & Willett, 1985; Willett, 1994; Williamson, 1986; Williamson, Appelbaum, & Ephanich, 1991) describe an alternative method of improving level-2 analyses associated with the exploratory approach. Based on the marginal maximum-likelihood methods of Blomqvist (1977), they suggest a reliability-based disattenuation of level-2 correlational analyses using a reliability estimate obtained by treating the individual level-1 residual variances as estimates of a common measurement error variance.
between-person models like those in equations 1 and 2. Such models constitute an algebraic hierarchy describing the statistical structure of growth data to be analyzed. Methodology for fitting such hierarchical models has advanced rapidly. Currently, there are several dedicated computer programs available for simultaneously estimating all of the parameters of such models, and for providing associated standard errors and goodness-of-fit statistics. Kreft, de Leeuw, and Kim (1990) provided a comprehensive description of four of them and compare their functioning. They concluded that, in general, “all four programs tend to converge to the same solution” (p. 100). One of them—called HLM—has been widely used in the empirical literature and is well supported (Bryk & Raudenbush, 1987, 1992).

Using Covariance Structure Analysis to Estimate the Parameters of the Level-2 Models. Recently, pioneering authors have demonstrated how the analysis of change can be conducted by the methods of covariance structure analysis using computer software like LISREL (Jöreskog & Sörbom, 1989). Meredith and Tisak (1984, 1990; see also Tisak & Meredith, 1990) provided a technical framework for representing level-2 differences in individual development and gave examples of how model parameters can be estimated. Their work extends earlier research on longitudinal factor analysis (Tucker, 1958; Rao, 1958) and subsumes more traditional approaches to the analysis of panel data, such as repeated-measures analysis of variance. McArdle and his colleagues demonstrated the application of the covariance structure approach to a variety of developmental problems in psychology. For instance, they showed how these methods can be used to estimate average growth curves and to detect inter-individual differences in change, in a single domain and simultaneously in many domains (McArdle, 1986a, 1986b, 1989, 1991; McArdle & Epstein, 1987). They showed how average growth curves can be compared across groups (McArdle, 1989; McArdle & Epstein, 1987). They described how covariance structure methods can be used in convergence analysis—in which segments of average growth curves, estimated in overlapping cohorts, can be linked into a single continuous trajectory (McArdle, Anderson, & Aber, 1987; McArdle & Hamagami, 1991). Finally, Muthén and his colleagues described the technical basis for, and provided data-analytic examples of, the modeling of multilevel data using covariance structure methods (Muthén, 1989; Muthén & Satorra, 1989). Of particular interest are two papers (Muthén, 1991, 1992) in which interindividual variation in the individual growth parameters is systematically related to selected time-invariant predictors of change. Willett and Sayer (1994, 1996) reviewed in detail how the methods of individual growth modeling map onto the covariance structure approach and offer documented examples of their use in the detection of
predictors of change and in the investigation of relationships among simultaneous changes in several domains of measurement.

TWO CRITICAL ADVANTAGES OF INDIVIDUAL GROWTH MODELING

There are two critical advantages that the collection of multiwave data and the application of individual growth modeling hold for the measurement of change. The first of these—which we call the technical advantage—is concerned with the reliability with which change can be measured, and the second—which we call the substantive advantage—is concerned with the complexity and richness of the research questions that can be addressed under the new methodology. In the following two subsections, I comment on each of these advantages.

A Technical Advantage—Improved Reliability For the Measurement of Change

Analyses of change are automatically improved if the level-1 estimation of the individual growth parameters is made more precise. In practice, improvements in the precision of individual growth parameter estimation are easily achieved simply by collecting additional waves of data on each person in the sample. In the case of straight-line growth, for instance, the standard errors of fitted growth-rates decrease dramatically as extra waves of data are added to the data collection. It is for this reason that multiwave longitudinal research designs provide superior for the measurement of change, because then researchers can control the quality of their findings by simply adding waves of data collection to the research design until some desired level of precision is reached (see Willett, 1989).

When the precision of the individual growth parameter estimates is improved, the reliability with which change can be measured usually increases. In Fig. 11.4, I illustrate this link by displaying the estimated reliability of rate of change in tolerance of deviant behavior as a function of the number of waves of data for the sample of 168 adolescents introduced earlier. In the figure, I display the reliabilities that would have been obtained if both additional, and fewer, waves of data had been collected, all else remaining equal. Under the original design—that is, with 5 yearly waves of data collected—the estimated reliability of change measurement was moderate, about .6. But, all else remaining constant, if only 2 waves of data had been collected then the estimated reliability of change would have been close to zero, about .1. On the other hand, if investigators had collected 3 additional waves of data, for a total of 8 waves, then the reliability of the change measurement would have risen
Estimated Reliability of Rate of Change

FIG. 11.4. Estimated reliability of the OLS-estimate of rate of change in tolerance of deviant behavior as a function of number of waves of data collected, for a sample of 168 adolescents (assuming linear individual growth, homoscedastic measurement error and equally spaced waves of data collection).

to approximately .8. In general, the monotonic nature of the relationship between the reliability of change and the number of waves of data collected suggests that it is always of benefit to add waves of data collection to a research design, but that the biggest impact on reliability occurs for designs that originally contained fewer waves.

A Substantive Advantage—Richer Formulations for Representing Individual Change

Although it is advantageous—from a technical perspective—to use research design to control the reliability with which change can be measured, the real beauty of the individual growth modeling approach is that
it thrusts the investigator into a universe of richer and more interesting research questions. This transition is a direct consequence of the flexibility that the investigator can exercise in picking the functional form of the mathematical model used to represent individual growth at level-1. Not only does the investigator pick which individual growth model is used, but he or she also chooses the particular parameterization of that model that seems most appropriate. I comment briefly on model selection and reparameterization here, beginning with the latter.

The same individual growth model can be parameterized in many different ways, with each reparameterization permitting a different set of research questions to be addressed. Even the simple straight-line growth model in equation 1 can be reparameterized so that different questions can be addressed. One interesting reparameterization abandons the straightforward notion of including individual intercept and slope parameters in favor of a pair of parameters that simultaneously represent a person's initial and final tolerance on entry into, and exit from, the study, respectively. Under this latter reparameterization, the straight-line model in equation 1 becomes:

$$\gamma_i = \pi_{0i} \left( \frac{15 - t_i}{4} \right) + \pi_{1i} \left( \frac{t_i - 11}{4} \right) + \varepsilon_i \tag{3}$$

At first glance, the reparameterized model looks unwieldy and strange. Notice that it contains no classical "intercept" term and that, counter-intuitively, age has been transformed to create a pair of new level-1 predictors, $(15 - t_i)/4$ and $(t_i - 11)/4$, that are included in the model simultaneously. However, equation 3 remains a straight-line growth model because tolerance is still hypothesized to be linear with age. The two individual growth parameters that grace the model are still labeled $\pi_{0i}$ and $\pi_{1i}$ as before, but now their interpretation differs somewhat. In particular, because of the way that age has been transformed and included in the model, the two individual growth parameters now operate as a pair of simultaneous "intercepts" with $\pi_{0i}$ representing true tolerance at age 11 (i.e., initial tolerance or tolerance on entry into the study, as before) and $\pi_{1i}$ representing true tolerance at age 15 (i.e., final tolerance or tolerance on exit from the study).\(^5\) This reparameterization is presented in Fig. 11.5—notice that the individual growth trajectory plotted in the figure is identical to that already displayed in Fig. 11.3, but that the individual growth parameters now display their new interpretation.

\(^5\)These interpretations can be checked by the substitution of age 11, or age 15, for $t_i$ in equation 3.
FIG. 11.5. The hypothetical trajectory of true linear growth corresponding to Equation [3], showing the function of the individual growth parameters: (a) $\pi_{1p}$ representing true initial status at age 11, and (b) $\pi_{1p}$ representing true final status at age 15.

By reparameterizing the linear growth model in this way, a different set of research questions can be addressed, in which the investigator makes hypotheses simultaneously about entry- and exit-level tolerance. When longitudinal data are available on each child, this simultaneous investigation of initial and final true tolerance is superior to the more-obvious piecemeal approach that addresses the same questions by the separate analyses of observed tolerance at age 11 and at age 15, for two reasons. First, under the individual growth modeling approach, analyses of initial and final tolerance are conducted simultaneously, with a considerable saving of time and effort. Second, the individual growth modeling analysis make use of all the longitudinal data that is available on each child—this means that data collected at the noninteresting ages of 12, 13, and 14 are being used in addressing questions about tolerance at
ages 11 and 15, along with the data that were obtained at age 11 and age 15. This ability to use each person's entire longitudinal data stream in an investigation of tolerance at any given age, or simultaneously at any pair of ages, leads to a considerable increase in power for the statistical tests that are ultimately performed.

Another interesting extension of the linear growth model is the creation of growth functions that are discontinuous or "piecewise" with time. For instance, suppose we hypothesize the presence of a sudden discontinuity in individual growth in tolerance of deviant behavior at age 13, say. Before and after age 13, let us assume that individual growth in tolerance proceeds smoothly and linearly, perhaps with different intercepts and slopes, but that there is a sudden discontinuous "jump" in the growth trajectory at age 13. We can easily create an individual growth model that represents this type of behavior, as follows:

$$Y_p = \pi_{0p} + \pi_{1p}(t - 11) + \pi_{3p}D_t + \pi_{3p}a_t + \epsilon_p$$  \hspace{1cm} (4)

In this model, time has three representations. First, in the second term on the right-hand side of the equation, $t$, represents the usual main effect of linear age, which, as before, has been recentered at age 11. This ensures, as before, that the first individual growth parameter, $\pi_{0p}$, represents true tolerance at age 11 on entry into the study and that the second individual growth parameter, $\pi_{1p}$, represents the rate of change in true tolerance prior to the age 13 discontinuity. Second, $D_t$ is a time-varying dichotomous predictor that has been coded so that it takes on value 0 at all ages up to, but not including, age 13 and value 1 at age 13 and beyond. This ensures that the third individual growth parameter in the model, $\pi_{3p}$, represents any sudden vertical elevation, or "jig," in the tolerance trajectory at age 13. The third representation of time, the level 1 predictor $a_t$, represents the individual's age after the age-13 discontinuity—it is coded so that it takes on value zero at all ages up to, and including, age 13 but is incremented as the years pass subsequent to age 13 (i.e., $a_t$ equals 1 at age 14, 2 at age 15, and so on). This forces the fourth parameter in the individual growth model, $\pi_{3p}$, to represent the change in linear slope across the age 13 discontinuity. All four of the new individual growth parameters in equation 4 are labeled on the true tolerance trajectory displayed in Fig. 11.6.

Although more complex than the simple linear growth models introduced earlier, equation 4 is very useful when individual change is thought to take place in fits and starts. This type of development may occur when a treatment or intervention has been administered to children at a particular age, or when some crisis or loss has occurred at some specific age. The formulation is also appropriate for representing the discontinuous
type of individual change that could occur at specific transition points in a child's academic, social or moral development. Finally, the discontinuous individual growth model in equation 6 can itself be generalized and reparameterized in several ways. First, by recentering age at different years, we can change the interpretations associated with the growth parameters in the model. Second, by combining a version of equation 3 with equation 4, we can create a discontinuous individual growth model that has been reparameterized in terms of true status at ages 11, 13, and 15, say. Third, provided that it is known for each child, the age at discontinuity need not be the same for all children. Fourth, the growth
model can be formulated to contain multiple discontinuities in an ordered sequence.

However, at this juncture, perhaps a word of caution is needed. As one might expect, under the principle that "you don't get something for nothing," the adoption of increasingly complex mathematical models to represent individual change is not without cost to research design. As parameters are added to the level-1 model for individual growth, additional waves of data must be collected so that the model can ultimately be fit. At a bare minimum, the investigator must plan on collecting at least one more wave of data than there are individual growth parameters in the level-1 model. Thus, the simple linear growth models in equations 1 or 3 require at least 3 waves of data, the more complex discontinuous linear growth model in equation 4, which contains four individual growth parameters, requires at least 5 waves of data. And, of course, these are the minimum data requirements—if greater precision is required, or the investigator wishes to test alternative hypotheses about the shape of individual change over time, then further waves of data must be collected.

Not only can the linear growth model be reparameterized and extended to address sets of alternative research questions but, under the individual growth modeling approach, we are free to adopt substantively interesting nonlinear models to represent curvilinear change if that is appropriate. This is useful because, in practice, individual change trajectories are often complex and curvilinear, particularly over the long term. For instance, if we believe that change is a quadratic function of age then a second-order polynomial can be used to represent individual growth. In this case, the observed status of the pth person at age \( t \) is represented by:

\[
Y_{tp} = \pi_{0p} + \pi_{1p}(t_i - 11) + \pi_{2p}(t_i - 11)^2 + \epsilon_{tp}
\]

(5)

It is the presence of the additional quadratic parameter, \( \pi_{2p} \), in equation 4 that permits the individual growth trajectory to be curved. As the magnitude of \( \pi_{2p} \) increases, the curvature of the individual growth trajectory becomes more extreme and, as is illustrated in Fig. 11.7, when \( \pi_{2p} \) is negative the hypothesized individual growth trajectory is concave to the time axis, and when it is positive the trajectory is convex to the time axis.\(^6\)

\(^6\)In a quadratic model, the slope and curvature parameters together provide another interesting interpretation—they specify the time at which the growth curve peaks or troughs. This can be seen by differentiating equation 5 and setting its first derivative to zero, in which case the time at which the quadratic growth curve "flips over" is:

\[
11 - \frac{\pi_{1p}}{2\pi_{2p}}
\]
The other two parameters in the model have interpretations similar, but not identical, to those specified earlier for the companion linear model. Because age has been recentered on the first occasion of measurement (when each adolescent was 11 years old), \( \pi_{0i} \) continues to represent the true tolerance of deviance at age 11 for the \( p \)th youth. However, because of the presence of the curvature parameter in the growth model, the slope parameter, \( \pi_{1pi} \) now represents the \( p \)th youth’s instantaneous rate of true change in tolerance of deviance at age 11 (that is, it is the slope of the tangent to the true growth curve at age 11). Adopting the quadratic model to represent change in tolerance of deviant behavior through middle adolescence permits the investigator to address level-2 research questions about
all three of the individual growth parameters—their initial tolerance, their initial rate of change in tolerance, and their curvature. In particular, by selection of a quadratic individual growth model, the empirical researcher can investigate predictors of interindividual variation in curvature: Is growth in the tolerance of deviant behavior more curved for boys than for girls? Does the curvature depend upon their early exposure to deviance?

As with the straight-line model, the quadratic individual growth model in equation 5 can be reparameterized in a variety of interesting ways. For instance, as in the reparameterization that links equation 1 and equation 3 for the linear growth model, the quadratic model can be modified to permit the simultaneous investigation of individual true status at ages 11 and 15, as follows:

$$Y_{\text{op}} = \pi_0 + \pi_1 \left( \frac{15 - t_1}{4} \right) + \pi_2 \left( \frac{t_1 - 11}{4} \right) + \pi_3 \left( \frac{15 - t_1}{4} \right) \left( \frac{t_1 - 11}{4} \right) + \epsilon_i$$  \hspace{1cm} (6)

or at any other pair of ages in which the investigator is interested. And, similar to the earlier reformulation of the straight-line model to permit growth that proceeds in fits and starts, we can easily provide for discontinuities in quadratic individual growth, as follows:

$$Y_{\text{op}} = \pi_0 + \pi_1 t_1 + \pi_2 (t_1 - 11) + \pi_3 (t_1 - 11)^2 + \pi_4 D_i + \pi_5 a_i + \pi_6 a_i^2 + \epsilon_i$$ \hspace{1cm} (7)

where, under the previous definitions of the level-1 predictors $D_i$ and $a_i$, there is not only a vertical discontinuity in true status as age 13 (of magnitude $\pi_6$), but also discontinuities in linear slope ($\pi_4$) and in curvature ($\pi_5$). The possibilities are limited only by the investigator’s ingenuity and the availability of longitudinal data sufficient to fit the selected model.

In general, the individual growth modeling approach can accommodate any level-1 model that is linear in the individual growth parameters. For instance, the quadratic individual growth model in equation 5 is linear in the individual growth parameters despite the fact that it represents curvilinear change—a growth model is defined as being linear in the parameters if the true status of individual $p$ on occasion 1 is a weighted linear composite of the individual growth parameters, with numerical weights that are constant or that depend upon the known times of measurement. Keats (1983) defined such models as having the property of dynamic consistency.\textsuperscript{7} Many common growth functions are dynamically

\textsuperscript{7}For such models, the “curve of the averages” (obtained by taking the population of true growth curves and plotting the average of the true scores at each value of time) is identical to the “average of the curves” (obtained by averaging the individual growth parameters over the population and plotting a curve with parameters equal to these averages).
consistent, including the quadratic model cited above and all other polynomial models, regardless of their order. Other potentially important individual growth models such as the logistic model (which provides an important theoretical representation of human development from the perspective of some psychological theories—see Fischer & Pipp, 1984) is not linear in the individual growth parameters in its usual formulation. However, recent work by van Geert (1991) suggests that, when the logistic model is formulated as a difference equation (see van Geert, 1991, equation 17a), it becomes linear in its parameters and is therefore amenable to manipulation in the ways that I have described, under the individual growth modeling approach. Finally, a further exciting implication of van Geert's extensive work on models for cognitive development is that other more complex, and even chaotic, nonlinear models for change may also be subsumable within the individual growth modeling framework.

CONCLUDING DISCUSSION

In this chapter, I have briefly reviewed both traditional and modern methods for the measurement and analysis of individual change over time. I have argued that well-known traditional methods are limited by their reliance on the two-wave or pre-post research design and that modern methods have resolved these difficulties by the collection and analysis of multiwave data. In particular, when extended longitudinal data are available on each person in the sample, the investigator can take advantage of the power and flexibility of individual growth modeling. Under this latter approach, one mathematical model (the level-1 model) is specified to describe the true trajectory of individual development and another model (the level-2 model) is specified to represent relationships between the shapes of the individual trajectories and selected predictors of change. A taxonomy of statistical methods, at different levels of analytic complexity, are available for fitting the level-1 and level-2 models to data and for testing hypotheses about their parameters.

The individual growth modeling approach to the measurement of change offers many critical advantages to the thoughtful empirical researcher, including the following:

- The individual growth modeling approach can accommodate any number of waves of longitudinal data. A key message of this chapter has been the notion that the classical pre-post research design is limited for the measurement of change. Two waves of data are simply not sufficient for a thoughtful and precise investigation of change over time. For a better job, multiple waves of longitudinal data are required on each
person in the sample—in our example, adolescents’ tolerance of deviant behavior was measured annually for each of 5 years during adolescence. Not only can the method of individual growth modeling incorporate these multiple waves of data into the statistical analyses but their inclusion leads naturally to higher precision for the estimation of the individual growth trajectory and to greater reliability for the measurement of change. It also permits the testing of richer and more sophisticated hypotheses about the nature of change over time.

- **Occasions of measurement need not be equally spaced.** In our example, measures of adolescents’ tolerance of deviant behavior were each separated by one year throughout the entire period of data collection. However, equal spacing of the occasions of measurement is not a requirement of the method. Data may be collected at irregular intervals either for convenience (at the beginning and end of each of several school years, perhaps) or because the investigator wishes to estimate certain features of the trajectory more precisely (by clustering data-collection points more closely at times of greater research interest, say).

- **Different individuals can have different data-collection schedules.** In our example, tolerance of deviant behavior was measured for all adolescents on the same set of 5 occasions. This is not a requirement of the method, however. In fact, every individual in the data set may have their own unique data-collection schedule—in terms of both the number of waves of data collected and in the spacing of those waves. One implication of this flexibility is that the usual problems of missing data are less important in the analysis of change than in traditional repeated-measures analysis of variance, say. Of course, even though the method can easily incorporate individuals with fewer than a full complement of longitudinal data, it pays to remain cautious in the face of missing data because data points may not be missing at random and those individuals with fewer data points do contribute less information to the analyses.

- **Individual change can be represented by straight-line, curvilinear or discontinuous trajectories.** In this chapter, for convenience, I have used a straight-line growth model to introduce the method of individual growth modeling, but I have demonstrated ultimately how the method can incorporate nonlinear and discontinuous growth models when they are appropriate. The ability of the method to describe individual development by complex curvilinear and discontinuous mathematical functions opens up a new world of richer research hypotheses and permits the investigator to address more complex and interesting questions about the nature of individual change.
Multiple predictors of change can be included at level-2. The analysis of systematic inter-individual differences in change need not be limited to the detection of the effect of a single predictor of change. Within normal constraints imposed by the requirements of statistical power and the tenets of common sense, multiple predictors of change can be included in all level-2 models. Predictors can either represent the main effects of important correlates of change or, by suitable preprocessing of the data set to create cross-products among interesting combinations of predictors, statistical interactions among potential correlates can also be investigated.

The methods of individual growth modeling have liberated empirical researchers who are interested in the investigation of human development. Now, they are limited only by their ability to collect sufficient waves of longitudinal data and to reformulate their research questions as sensible individual growth models. After many years of dissension and distrust in the measurement of change, the analytic method itself is no longer perceived as the problem, as Lee J. Cronbach believed when he asked, 

"How we can measure change, and should we?" (Cronbach & Furby, 1970).

REFERENCES


11. MEASURING CHANGE


