CHAPTER 10

CONSIDERING ALTERNATIVE METRICS OF TIME

Does Anybody Really Know What "Time" Is?

Lesa Hoffman
University of Nebraska-Lincoln

Longitudinal studies (i.e., in which each person is observed at multiple occasions) are a cornerstone of research in psychology and human development and have become increasingly common across fields, such as education and business. Although many developmental questions have initially been addressed using cross-sectional studies, such between-person comparisons of people of different ages at a single point in time are often subject to well-known biases, including cohort effects, self-selection effects, mortality effects, and other problems (for more extended discussion, see Baltes, Cornelius, & Nesselroade, 1979; Baltes & Nesselroade, 1979; Hofer & Sliwinski, 2006; Schaie, 1965, 2008). Longitudinal studies can offer significant advantages over cross-sectional studies, in that not only can they provide cross-sectional, between-person information about interindividual

Advances in Longitudinal Methods in the Social and Behavioral Sciences, pages 255–287

variation (i.e., when the longitudinal study begins as a cross-sectional study of persons at different ages), but because they also provide within-person information about intraindividual change or variation over time (and between-person differences in those within-person changes).

Extensive methodological work has focused on the development of statistical models of change, such as for describing and predicting how scholastic achievement of children grows over time, how job performance of employees changes over time, how marital satisfaction waxes or wanes over time, how physical and cognitive function in older adults declines over time, and so forth. Indeed, much of this book focuses on the development or refinement of longitudinal models to be able to ask and answer questions of increasing complexity. Yet in order to make informed use of such exciting advances for longitudinal models, we must make an important assumption—we presuppose to know exactly what "time" is. That is, what is thought to be the fundamental causal process by which one should index change? Such deliberation on the possible metrics for indexing time (that in turn can reflect different processes) becomes important whenever persons differ at the onset of a study in the time metric of interest (e.g., persons of different ages at baseline). To illustrate, let us consider in more detail three of the examples already given: modeling growth in children' scholastic achievement, increases in employee job performance, and changes in marital satisfaction over time.

Considering Time

First, with respect to growth in children's scholastic achievement over time, we might reasonably assume that learning proceeds as a function of grade in school, given that each child is observed in multiple grades (i.e., there is within-person variation in grade as "time"), even if the children begin the study in different grades (i.e., there is also between-person variation in grade as "time"). Although children in the same grade may still differ in age, to the extent that scholastic achievement is a consequence of instruction (and not biology), then grade in school is likely to be more relevant for indexing learning over time than chronological age. Any effect of age differences between children in the same grade could still be accounted for by including the age at which they entered school as a person-specific covariate. Thus, in this context, what "time" should be seems relatively straightforward (even if more than one option is possible).

Second, in considering employee joh performance over time, we might assume that performance improves as a function of years of experience on the job, which we can monitor through repeated performance observations (i.e., "time" represents within-person variation in work experience

at their organization). But complications arise because employees often begin their position with different employment histories (i.e., "time" includes between-person variation in work experience also). For instance, consider the process of university promotion and tenure. Whether or not tenure requirements have been met is often evaluated as a function of years of experience (i.e., most candidates will apply for tenure 5–7 years into a tenure-track position), but at what point does the "work experience" that is relevant for promotion and tenure really begin? Does the relevant experience begin only at the entry point into that tenure-track position, or does it begin earlier, at the point of receipt of the doctoral degree, or npon completion of postdoctoral training or internship? What about persons who enter a tenure-track position after completing several years in a similar position at another university? If multiple candidates with different amounts of previous work experience (e.g., only graduate work, postgraduate work, or a previous tenure-track position) apply for tenure at the same time, the expectations for their accomplisbments at that point in time may heavily depend on how their relevant "time on the job" is conceptualized and measured. Thus, in this instance, what "time" should be is debatable, but with significant repercussions resulting from each possible alternative.

In other contexts, the best choice for "time" may be even less obvious. For instance, in studying changes in marital satisfaction over time, a logical first choice for time might be "time in marriage" (i.e., marital satisfaction may wax or wane the longer one is married, or due to within-person variation in "time in marriage"), even if couples differ in how long they've been married when they enter the study (i.e., there is also between-person variation in "time in marriage"). However, there is likely to be considerable heterogeneity in how quickly different couples may decide to marry. If one believes that relationship satisfaction progresses as a function of the length of the *overall* relationship (rather than just the length of the marriage), then "time in *relationship*" may more accurately index observed changes in relationship satisfaction than would "time in *marriage*." But some couples may meet and begin dating immediately, whereas others may be friends initially and then later decide to begin dating. In that case, "time in relationship" would need to be further distinguished as "time in *any* relationship" versus "time in a *romantic* relationship." Furthermore, some couples may have more volatile relationships, such that they may break up, but then later decide to get back together (perhaps doing so multiple times). Should any time in between their relationship epochs "count" within whichever time metric is used to index change in relationship satisfaction over time?

In addition, alternative theoretical viewpoints of relationship dynamics may require very different metrics of time for indexing change in relationship satisfaction. What if one believes that relationship satisfaction changes due to changing responsibilities of the spouses or partners (e.g., the transi-

tion to parenthood)? If so, how long a couple has been together (however defined) would be less relevant than how long they have been parents together (e.g., the age of their first child). Further still, couples may decide to end their unions at different points in their relationship, and so perhaps declines in satisfaction could be described more parsimoniously by tracking change as a function of a meaningful event, such as beginning counseling, separation, or divorce. In that case, time would be measured backwards in order to describe change over time as a function of the *dissolution* of the relationship instead. That is, couples would be aligned with respect to how soon they will be apart, rather than how long they've been together. Finally, because the couples that are still together after a long relationship are not a random subsample of all couples who had begun a relationship (because they have chosen to stay together), inferences about changes in relationship satisfaction need to be viewed as conditional on this self-selection process. Thus, what "time" should be in this context is anything but clear-cut!

The Focus of This Chapter

The point of the three preceding examples was to illustrate how the decision to index change over time should reflect the theoretical process thought to be responsible for any observed change, and thus how several alternative metrics of time (corresponding to different theoretical orientations) may be useful for tracking change in a given outcome as a result. The purpose of the present work is to thoroughly explore these issues surrounding the choices we make for the specification of time within longitudinal models. The title of this chapter is based on a song recorded by the band Chicago in 1969, in which the lyrics query: "Does anybody really know what time it is? . . . Does anybody really care (about time)?" I believe that we should indeed care about time (at least when conducting longitudinal analysis), and so the goal of this chapter is to present how ofteu unrecognized assumptions about the treatment of time can have important consequences for subsequent model interpretation. I use a working example examining change in cognitive functioning in older adults as a function of three alternative metrics of time (time since birth, tiine until death, and time since dementia diagnosis) to address two general issues: (1) what "time" sbould be and (2) how "time" should be modeled.

What Should Time Be?

It is important to note that the question of what "time" should be is not relevant within persons, in which all metrics for indexiug time are indistinguishable. For instance, in the previous relatiouship example, as each year passes relative to a person's status at the beginning of the study, he or she has been married one year longer, in the relationship one year longer, in the romantic relationship one year longer, has been a parent one year longer, and may be one year closer to separation or divorce. Within persons, these alternative metrics of time cannot be distinguished—within persons, time is just time. Between persons, however, all time metrics are not equivalent—people may begin the study at different points in "time" (e.g., they may have been married longer, have been a parent longer, or may begin the study closer to separation or divorce). As a result, both the amount of interindividual variation in "time" and its relationship with a given outcome are likely to differ depending on what "time" is. Thus, the first question is, given the presence of both between-person and within-person variation in time, from what point should we start counting—how should time be aligned between persons?

How Should Time Be Modeled?

Second, how should our model of change account for the potentially different effects of between- and within-person differences in "time"? Different model specifications (in addition to different time metrics) may result in different conclusions for the description and prediction of change over time. Such considerations have been described more generally for multilevel models as they relate to distinguisbing individual effects from contextual effects (e.g., Raudenbush & Bryk, 2002; Snijders & Bosker, 1999), or distinguishing time-specific effects from individual effects of time-varying covariates (e.g., Hedeker & Gibbons, 2006; Hoffman & Stawski, 2009), but relatively little attention has been paid to this issue in the context of indexing time in longitudinal studies, in which the same concerns are also relevant. Thus, the second question is, within a chosen time metric, how should the model be specified to best account for *all* sources of variation in time?

EXAMPLE DATA

Sample

These questions surroundiug alternative metrics and models of time are addressed using data from the Octogenarian Twin Study of Aging (as described in Johansson et al., 2004), in which observations were collected longitudinally from same-sex twin pairs. One twin from each pair was randomly selected for use in the current analyses, which included 173 persons (65% women) who were sampled on np to five occasions over an 8-year period (i.e., every 2 years). Other relevant characteristics of the analysis sample are summarized below.

Outcomes

Two cognition outcomes were examined. The Mini-Mental Status Exam (MMSE; Folstein, Folstein, & McHugh, 1975) is a general test of orientation and memory that is often used to identify persons suspected of having dementia. The questions are relatively easy and thus most participants without cognitive impairment score at ceiling. The second outcome was a more sensitive measure of memory, the Memory-in-Reality Object Recall Test (Johansson, 1988), in which participants were asked to place real-life objects in a three-dimensional model of an apartment and were later given a free-recall test for those objects. For ease of interpretation, both outcomes were *T*-scored to the same scale ($M = 50$, $SD = 10$). Additional information about the original OCTO-Twin sample and these measures of cognition can be found in Johansson et al. (1999, 2004).

Decomposing Variance

Before building longitudinal models, it is useful to decompose the outcome variability into *between-person* variability in the mean level of the outcome over time (i.e., cross-sectional variability representing interindividual differences) and *within-person* variability around a person's mean outcome over time (i.e., longitudinal variability representing intraindividual change and fluctuation). These sources of variation can be quantified by estimating what is called an "empty" longitudinal (multilevel) model, as shown in equation 10.1:

$$\begin{aligned} &\text{Level 1: } y_{ti} = \beta_{0i} + e_{ti} \\ &\text{Level 2: } \beta_{0i} = \gamma_{00} + U_{0i} \end{aligned} \tag{10.1}$$

in which y_{ti} is the outcome at time t for individual i. The level-1 model constructs an outcome at each occasion as a function of an individual intercept, represented by the placeholder β_{0i}, and a time- and individual-specific residual e_{ti}. The level-2 model then describes how each person's intercept is constructed: Here, as a function of the fixed intercept γ_{00}, which is the grand mean of the outcome over time, and the random intercept U_{0i}, which is the deviation from the grand mean of individual i's mean over time. Thus, U_{0i} represents between-person (BP) variation in the person-means (estimated as $\tau^2_{U_0}$) and e_{ti} represents within-person (WP) variation around those person-means (estimated as σ^2_e). We can then form a ratio of these two variance components in order to calculate an intraclass correlation (ICC), as shown in equation 10.2:

$$\text{ICC} = \frac{\text{var}(U_{0i})}{\text{var}(U_{0i}) + \text{var}(e_{ti})} = \frac{\tau^2_{U_0}}{\tau^2_{U_0} + \sigma^2_e} = \frac{\text{BP variation}}{(\text{BP} + \text{WP}) \text{ variation}} \tag{10.2}$$

in which the ICC is the proportion of variance that is between persons (or equivalently, the average correlation across occasions assuming compound symmetry). In this example, the ICCs for the MMSE and Object Recall outcomes were .50 and .42, respectively, indicating that approximately half of their variance was in mean level over time (between-person, cross-sectional variance).

Accelerated Longitudinal Designs

The example data were collected in an *accelerated longitudinal design* (i.e., overlapping cohort design, cohort sequential design; Bell, 1953; McArdle & Bell, 2000). Accelerated designs are useful for studying human development in a shorter time frame than that in which the development actually occurs. That is, accelerated designs can be useful when one wishes to do a longer-term longitudinal study, but just doesn't have the time. An example of an accelerated design is shown in Figure 10.1, in which the top panel depicts the sampling of different age cohorts (every 5 years from age 50 to 85), each of which is sampled for 10 years. The aim of such an accelerated design would be to capture a general age curve by overlapping the observed age cohorts, as shown in the bottom panel. If the overlapping age cohorts converge onto the same age trajectory, one can then model the developmental trajectory over a larger span of time (i.e., 45 years total in Figure 10.1) than would be directly possible using only longitudinal information (i.e., only 10 years within any person in Figure 10.1).

Choosing amongst Alternative Metrics of Time

The participants analyzed in the current example were selected because they each had known dates of birth, known dates of death, and estimated ages of onset of dementia (including Alzheimer's disease, vascular dementia, or dementia with a mixed or unknown etiology; type of dementia was not distinguished for the purposes of this example). Thus, there are at least three alternative metrics of time that could be used to index change (and thus with which to align between-person variability in "time") in these longitudinal data. Let us consider each in turn.

Age as Time

First, we could index change as a function of age, or *time since birth*, given that participants ranged from age 79 to 100 years of age ($M = 84$ years,

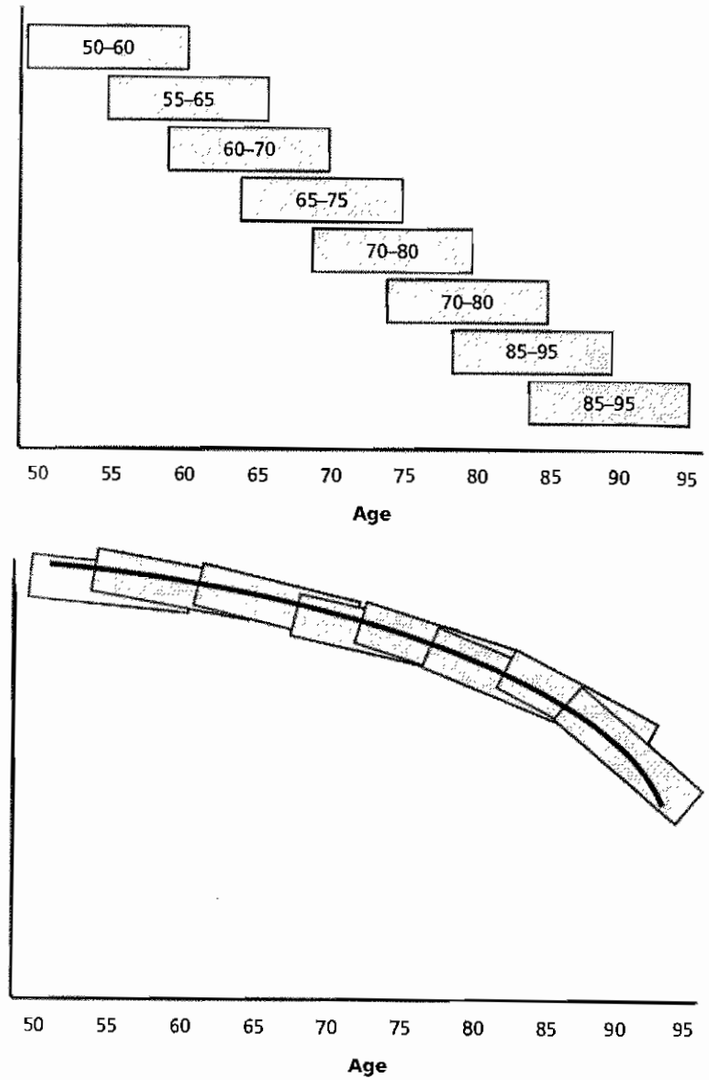

Figure 10.1 Example of an age-accelerated longitudinal design.

$SD = 3$ years) at the study beginning (baseline). This has been the most common approach by far within the cognitive aging literature. In constructing a model where age is time, individual differences would be organized around the mean outcome at a particular age (e.g., the mean of 84 years), and change would be specified as a function of the distance from that age. The use of age as time implies a theoretical model whereby cognition declines as a function of *time since birth*, such that aligning individuals based on their current age should be informative for describing individual differences in both the level and change in cognition over time (in which "time" would be age here).

Although age is measured longitudinally, if persons differ in age at baseline (i.e., they range from 79 to 100 years here), then age is also measured cross-sectionally. Therefore, it is useful to index the relative amount of cross-sectional versus longitudinal information available in the age predictor variable. We can do this by estimating an ICC for age (using equation 10.2) as derived from specifying time-varying age as an outcome in an empty model (using equation 10.1). In the example data, the ICC for age was .47, indicating that 47% of the variance in age was cross-sectional (i.e., due to initial age differences), whereas 53% was longitudinal (i.e., due to observed age changes during the study). Thus, although our theoretical model is based on the idea of "aging," only about half of the variance in age will be directly informative about within-person age changes. The rest of the age variance will be informative about preexisting differences between persons of different ages instead.

Death as Time

An alternative approach that has become increasingly popular within cognitive aging is to index change as a function of years to death instead of age, or *time to death*, given that participants ranged from −16 to 0 years to death ($M = -6$ years, $SD = 4$ years) at baseline. In constructing a model in which death is time, individual differences would be organized around the mean outcome at a particular distance from death (e.g., the mean of −6 years), and change would be specified as a function of the distance from that point. The use of death as time implies a theoretical model whereby cognition declines as a function of *impending death* (i.e., terminal decline), such that aligning individuals based on their current distance from death should be informative for describing individual differences in level and change in cognition over time, rather than based on their current age. In other words, it matters how many years one has left, not how many years one has already had. The ICC for time to death was .24, indicating that 24% of the information in death as time is between persons who enter the study at different durations to death (i.e., cross-sectional variance), whereas 76% is within-persons (i.e., longitudinal variance) as they grow closer to the end of their lives during the study.

Dementia as Time

A third option for "time" is based on a common important event: Given that everyone in the sample has been or will be diagnosed with dementia, we could also index change using proximity to the dementia diagnosis, or *time to dementia.* Participants ranged from −12 to 18 years from diagnosis ($M = 0$ years, $SD = 5$ years) at baseline, indicating that some participants entered the study already having been diagnosed with dementia, whereas others received a diagnosis at some point during or after the study. The use of dementia as time implies a theoretical model whereby cognition declines as a function of *dementia disease progression,* such that aligning individuals based on their current time with the disease (without regard to age or years to death) should be informative for describing individual differences in level and change in cognition over time. The ICC for time to dementia was .71, indicating that 71% of the information in dementia as time is between persons who enter the study with different amounts of disease progression (i.e., cross-sectional variance), whereas 29% is within persons relative to their own progression observed during the study (i.e., longitudinal variance). One significant limitation in using dementia as time, however, is that age of dementia onset can ouly be estimated, in contrast to observable events that have defined dates, such as birth or death. Thus, the variable for time to dementia diagnosis will contain measurement error, whereas the variables for time since birth or time to death (that are known rather than estimated) should uot.

Each metric of time (time since birth, time to death, and time to dementia) forms an accelerated design iu its own right, in that persons differ at baseline in each measure, and they also differ in how much of their information is actually cross-sectional (24–71%). Furthermore, the baseline values for these time dimensions are surprisingly uncorrelated. Age is only correlated with time to death at $r = .23$ and with time to dementia at $r = .17$, although time to death aud time to dementia are correlated more highly at $r = .52$ (given that dementia can be a cause of death). Thus, these alternative metrics of time will align different persons in very different ways.

Time as Time

Finally, a less obvious choice for organizing individual differences is simply "time" itself, or *time in study,* which ranges from 0 to 8 years with an ICC of exactly 0, indicating that time in study represents solely longitudinal information. The use of time in study as time makes no theoretical statement whatsoever—individuals are simply organized around their baseline level of performance and their change from baseline. Thus, when used by itself, time in study ignores individual differences in time since birth, time to death, and time to dementia. As we will see later, however, such an uninformative metric for time may actually be useful in empirically distinguishing among those distinct temporal processes.

Visualizing Alternative Metrics of Time

A useful descriptive exercise in evaluating alternative metrics of time is to construct plots of individual trajectories with an overlaid trajectory of the model-estimated means at each measurement occasion within that time metric. Figure 10.2 shows four such plots for the MMSE outcome: time since birth (top left), time to death (top right), time to dementia (bottom left), and time in study (bottom right). Because MMSE is also used to assess the presence of dementia, the time to dementia plot is somewhat circular, but it nevertheless illustrates an idealized scenario in which the mean trajectory (as shown by the heavy black line) is a good descriptor of the patterns shown in the individual trajectories. That is, we can informally judge the appropriateness of a given metric of time hy the similarity of the mean and individual trajectories—time to dementia appears to be a useful way to describe change in MMSE. An example of a poor match between the mean and individual trajectories can be seen for time since birth, which shows a much shallower rate of decline across age on average than what is shown by any individual. The same is true to a lesser extent for time to death. In contrast, the mean slope across time in study seems to match the individual trajectories fairly well, although there is noticeably greater heterogeneity in mean level relative to that shown in the accelerated time metrics (age, death, or dementia).

Figure 10.3 shows the same types of plots for the Object Recall memory outcome. Here the "best" metric of time is not nearly as evident, although the same general patterns appear: Time to dementia arguably seems to organize the individual trajectories around the mean trajectory mostly closely, followed by time to death and time in study, followed by time since birth.

Modeling Alternative Metrics of Time

Age as Time

Those mean and individual trajectories can then be modeled via fixed and random effects, in which fixed effects represent sample average effects and random effects represent deviations from each of those sample average effects for a given individual. Given that age is a commonly used metric of time, we can begin by constructing a model for change including fixed and random effects of age. Furthermore, we can approximate the apparent

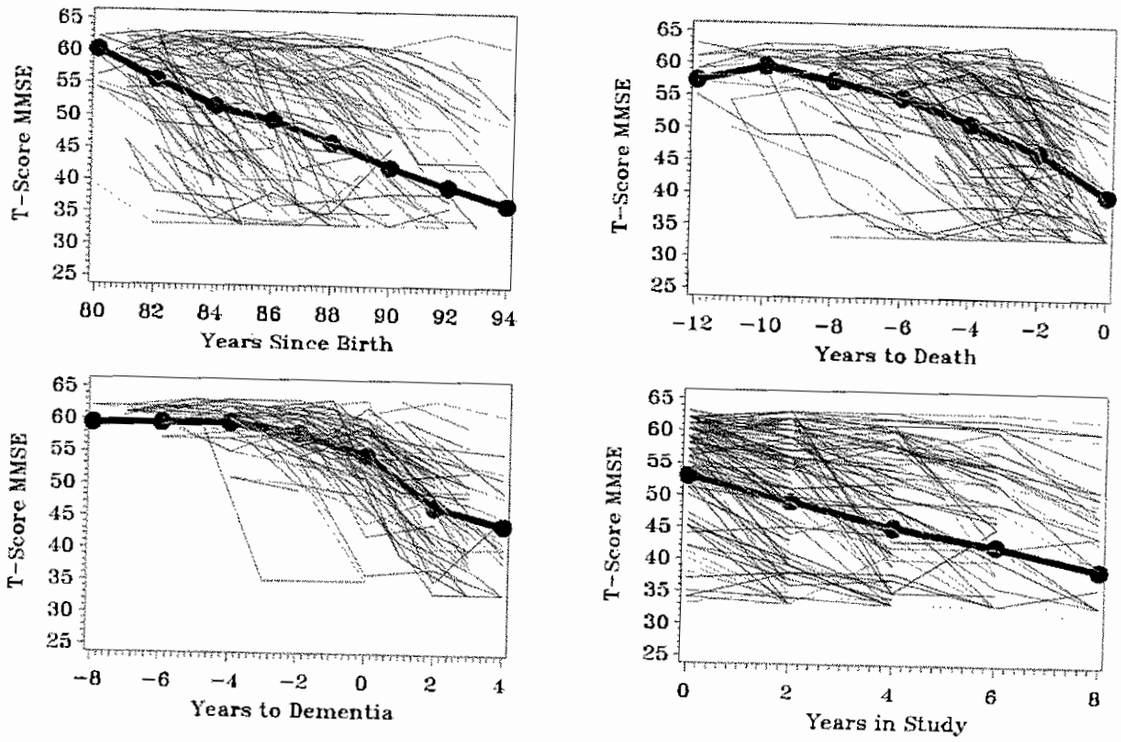

Figure 10.2 Individual (thin lines) and mean (thick line) trajectories for Mini-Mental Status Exam (MMSE) across years since birth (age; top left), years to death (top right), years to dementia diagnosis (bottom left), and years in study (bottom right).

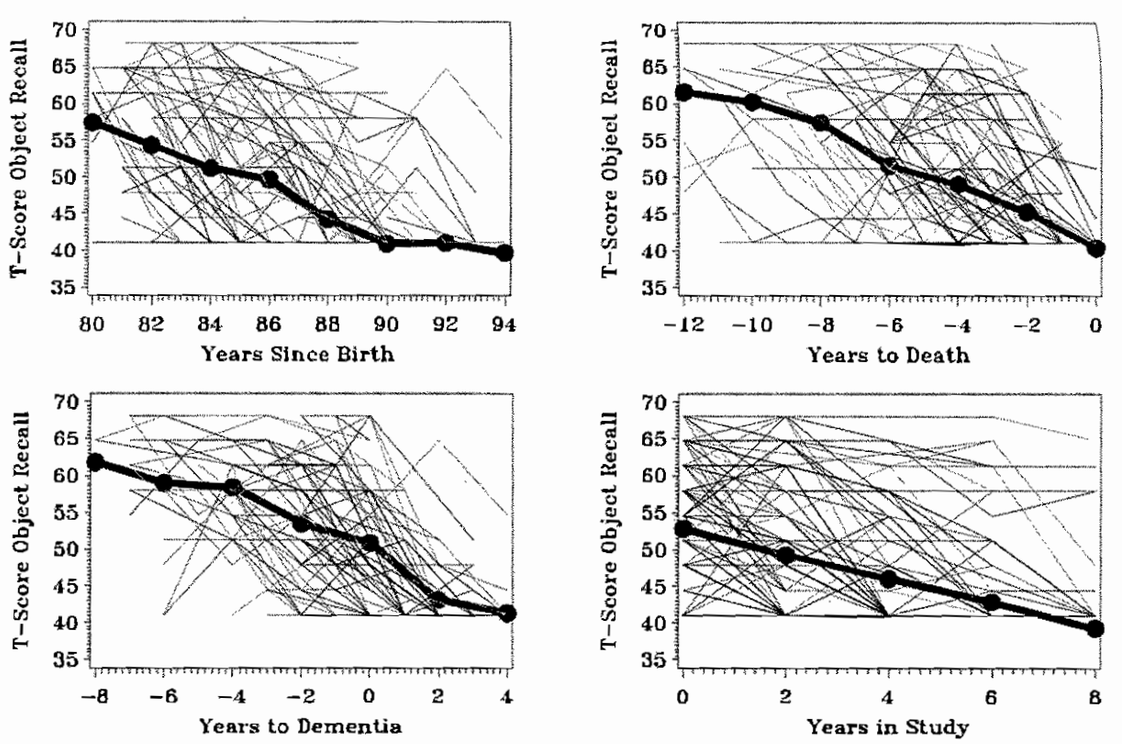

Figure 10.3 Individual (thin lines) and mean (thick line) trajectories for Object Recall memory across years since birth (age; top left), years to death (top right), years to dementia diagnosis (bottom left), and years in study (bottom right).

nonlinearity in the pattern of change across age with a quadratic effect of age, as shown in equation 10.3:

$$\begin{aligned}
\text{Level 1:}\quad & y_{ti} = \beta_{0i} + \beta_{1i}(\text{Age}_{ti} - 84) + \beta_{2i}(\text{Age}_{ti} - 84)^2 + e_{ti} \qquad (10.3)\\
\text{Level 2:}\quad & \beta_{0i} = \gamma_{00} + U_{0i}\\
& \beta_{1i} = \gamma_{10} + U_{1i}\\
& \beta_{2i} = \gamma_{20} + U_{2i}
\end{aligned}$$

in which y_{ti} is again the outcome at time t for individual i. The level-1 model constructs an outcome at each occasion using an individual intercept, linear age slope, and quadratic age slope, as represented by the placeholders β_{0i}, β_{1i}, and β_{2i}, as well as a time- and individual-specific residual e_{ti}. The subtraction of 84 from age moves the reference point for the model from that of a newborn (if using original age) to an 84-year-old (if using age centered at 84 instead). The level-2 model then constructs each person's growth terms: as a function of the fixed (average) effect for the sample (γ_{00}, γ_{10}, and γ_{20}) and the individual random effect (U_{0i}, U_{1i}, and U_{2i}). Each term varies across persons in this example model (but in practice the necessity of each fixed or random effect should be tested). The fixed intercept (γ_{00}) is the expected outcome at age 84, and thus the intercept variance ($\tau^2_{U_0}$) describes individual differences in the outcome level at age 84. Given the quadratic age slope, the fixed linear age slope (γ_{10}) is the instantaneous linear rate of change per year of age as evaluated at age 84, and thus the linear age slope variance ($\tau^2_{U_1}$) describes individual differences in the linear age slopes also as evaluated at age 84. The fixed quadratic age slope (γ_{20}) is half the rate of acceleration or deceleration; twice the quadratic slope is how the linear age slope changes per year of age. The quadratic age slope variance ($\tau^2_{U_2}$) describes individual differences in the quadratic age slopes (which are constant across age). The U_{0i}, U_{1i}, and U_{2i} level-2 random effects are assumed to have a multivariate normal distribution across persons, whereas the e_{ti} level-1 residuals are assumed to have constant variance across persons and occasions, with no covariance across occasions within persons or between persons.

Although not obvious, the age as time model in equation 10.3 makes an important assumption of *convergence*, that is, it assumes that persons of differing initial ages all converge onto the same trajectory (i.e., as shown in Figure 10.1). More conceptually, this assumption of age convergence means that the only reason that younger and older people differ is their age, or that the cross-sectional (BP) effects of age are equivalent to the longitudinal (WP) effects of age (i.e., an assumption of ergodicity). More succinctly, convergence means that it only matters *what* age you are; it does not matter *when* you were that age. Convergence is not likely to hold to the extent that the initial age range is large or to the extent that cohort effects, selection effects, or mortality effects are likely to be present. In these example data, 47% of the variation in age is cross-sectional, and thus it is an open question whether the cross-sectional and longitudinal effects of age are equivalent (i.e., whether the effects of age show convergence).

As discussed by Sliwinski, Hoffman, and Hofer (2010), age convergence can be tested empirically by using a variant of the grand-mean-centering approach that is used to decompose effects across levels in other multilevel contexts (i.e., distinguishing individual from group effects; distinguishing time-specific from individual effects), as shown in equation 10.4:

$$\begin{aligned}
\text{Level 1:}\quad & y_{ti} = \beta_{0i} + \beta_{1i}(\text{Age}_{ti} - 84) + \beta_{2i}(\text{Age}_{ti} - 84)^2 + e_{ti} \qquad (10.4)\\
\text{Level 2:}\quad & \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{AgeT1}_i - 84) + U_{0i}\\
& \beta_{1i} = \gamma_{10} + \gamma_{11}(\text{AgeT1}_i - 84) + U_{1i}\\
& \beta_{2i} = \gamma_{20} + \gamma_{21}(\text{AgeT1}_i - 84) + U_{2i}
\end{aligned}$$

in which the age at baseline (AgeT1_i) has been added as a time-invariant predictor to each level-2 equation. Although the decomposition of effects across levels typically requires computation of the mean at the higher level (i.e., mean age across time rather than age at baseline), in this case the mean age is likely to be biased by missing data—persons who dropped out of the study earlier would have a lower mean age, eveu if they were born in the same year. Thus, age at baseline is used as a more direct representation of age cohort (although birth year should be used in studies spanning multiple years at the first occasion). If significant effects of age cohort are found on the intercept (γ_{01}) or the linear or quadratic age slopes (γ_{11} or γ_{21}), this implies *age nonconvergence*, or that age cohort has an incremental effect (i.e., a contextual effect), even after controlling for current age. In other words, it would matter *when* you were age 84. Estimated parameters from the age as time models with age cohort are given in the second columns of Tables 10.1 and 10.2 for MMSE and Object Recall, respectively, and the corresponding model-predicted trajectories are depicted in the top panels of Figure 10.4. The quadratic age slope variance was nonsignificant and was thus not retained.

The parameters from the age as time model for MMSE can be understood as follows. Because of their interactions with age cohort, the fixed intercept, linear age slope, and quadratic age slope are interpreted conditionally on age cohort; that is, they apply specifically to someone who begins the study at age 84. Thus, for that individual, the expected MMSE (in

TABLE 10.1 Model Parameters across Alternative Metrics of Time for Mini-Mental Status Exam (MMSE)

Model parameters	Time since birth		Time to death		Time to dementia	
	Est	SE	Est	SE	Est	SE
Fixed effects:						
Intercept (γ_{00})	52.76**	0.89	55.99**	0.87	52.83**	0.52
Linear slope (γ_{10})	−1.70**	0.24	−1.02**	0.31	−2.16**	0.16
Quadratic slope (γ_{20})	−0.05	0.03	−0.19**	0.04	−0.09*	0.03
Cubic slope (γ_{30})					0.01**	0.00
Cohort on intercept (γ_{01})	1.09**	0.37	0.62*	0.30	0.03	0.13
Cohort on linear slope (γ_{11})	0.08	0.08	0.05	0.08	0.07	0.04
Cohort on quadratic slope (γ_{21})	−0.01*	0.01	−0.02**	0.01	0.00	0.00
Cohort on cubic slope (γ_{31})					−0.00**	0.00
Variance components:						
Residual variance (σ^2_ϵ)	15.38**	1.53	13.61**	1.33	20.16**	1.55
Intercept variance ($\tau^2_{U_0}$)	69.18**	9.00	63.90**	10.11	12.70**	2.44
Linear slope variance ($\tau^2_{U_1}$)	1.09**	0.29	0.97**	0.24		
Intercept slope covariance ($\tau_{U_{01}}$)	1.04	1.15	−2.56*	1.28		
Deviance (−2LL)	3513		3428		3267	
AIC	3533		3448		3287	
BIC	3564		3480		3319	

Note: * $p < .05$, ** $p < .001$. LL, log likelihood; AIC, Akaike information criteria; BIC, Bayesian information criteria. *Cohort* represents the value of each time variable at baseline.

T-score units) at age 84 is 52.76 (the fixed intercept γ_{00}), the instantaneous linear rate of decline as evaluated at age 84 is 1.70 per year (γ_{10}), and that rate of linear decline becomes (nonsignificantly) more negative by 0.10 per year (twice the quadratic slope γ_{20}, whose effect was fixed only). However, significant effects of age cohort (age at baseline) were found, such that for every year older one begins the study than age 84, the intercept at age 84 is expected to be 1.09 higher (γ_{01}), the linear rate of decline at age 84 is expected to be 0.08 less negative (nonsignificant γ_{11}), and the quadratic rate of decline is expected to be 0.01 more negative (γ_{21}). Thus, for MMSE, even after controlling for current age, persons who begin the study older have an age trajectory that begins higher than expected, but with more accelerated decline. For Object Recall, a somewhat different pattern of results was found: The expected value at age 84 was 51.60 for someone who began the study at age 84 (γ_{00}), and persons who begin the study at age 85 instead of age 84 were expected to score 0.53 higher (γ_{01}) at age 84. The form of change in Object Recall was a decelerating negative function, such that the

TABLE 10.2 Model Parameters across Alternative Metrics of Time for Object Recall Memory

Model parameters	Time since birth		Time to death		Time to dementia	
	Est	SE	Est	SE	Est	SE
Fixed effects:						
Intercept (γ_{00})	51.60**	0.82	54.23**	0.87	49.90**	0.68
Linear slope (γ_{10})	−1.76**	0.19	−1.79**	0.15	−1.82**	0.22
Quadratic slope (γ_{20})	0.06*	0.03			−0.11*	0.05
Cubic slope (γ_{30})					0.02**	0.01
Cohort on intercept (γ_{01})	0.53*	0.26			0.04	0.18
Cohort on linear slope (γ_{11})					0.14**	0.05
Cohort on quadratic slope (γ_{21})					−0.02**	0.01
Cohort on cubic slope (γ_{31})					−0.00**	0.00
Variance components:						
Residual variance (σ^2_ϵ)	28.27**	3.00	29.78**	3.27	31.12**	2.59
Intercept variance ($\tau^2_{U_0}$)	62.98**	9.55	63.28**	12.65	15.05**	3.08
Linear slope variance ($\tau^2_{U_1}$)	0.72**	0.31	0.67*	0.35		
Intercept slope covariance ($\tau_{U_{01}}$)	−4.02**	1.33	−5.69**	1.88		
Deviance (−2LL)	2851		2803		2719	
AIC	2867		2815		2737	
BIC	2892		2834		2765	

Note: * $p < .05$, ** $p < .001$. LL, log likelihood; AIC, Akaike information criteria; BIC, Bayesian information criteria. *Cohort* represents the value of each time variable at baseline.

linear rate of decline at age 84 of 1.76 per year (γ_{10}) became less negative by 0.12 (twice the quadratic γ_{20}) per year. For Object Recall, however, the rates of linear and quadratic decline across age did not differ by age cohort (i.e., no γ_{11} or γ_{21} were needed).

The positive effect of age cohort on the intercept shown hy both outcomes could potentially reflect a selection effect, given that the participants who are still alive and capable of agreeing to participate at older ages are not a random subsample of all participants who could have begun the study earlier—they are likely to have comparatively greater cognitive and physical function (but may be more likely to experience greater subsequent decline, at least in MMSE). It is important to note, however, that the cohort effect on the model intercept (at age 84) is necessarily an extrapolation for the older age cohorts, who did not contribute data at age 84. The cohort effect on the intercept reflects the difference predicted by the model *that should have been observed had the data been complete.* Although such an extrapolation may seem somewhat strange, that is exactly what is implied by any acceler-

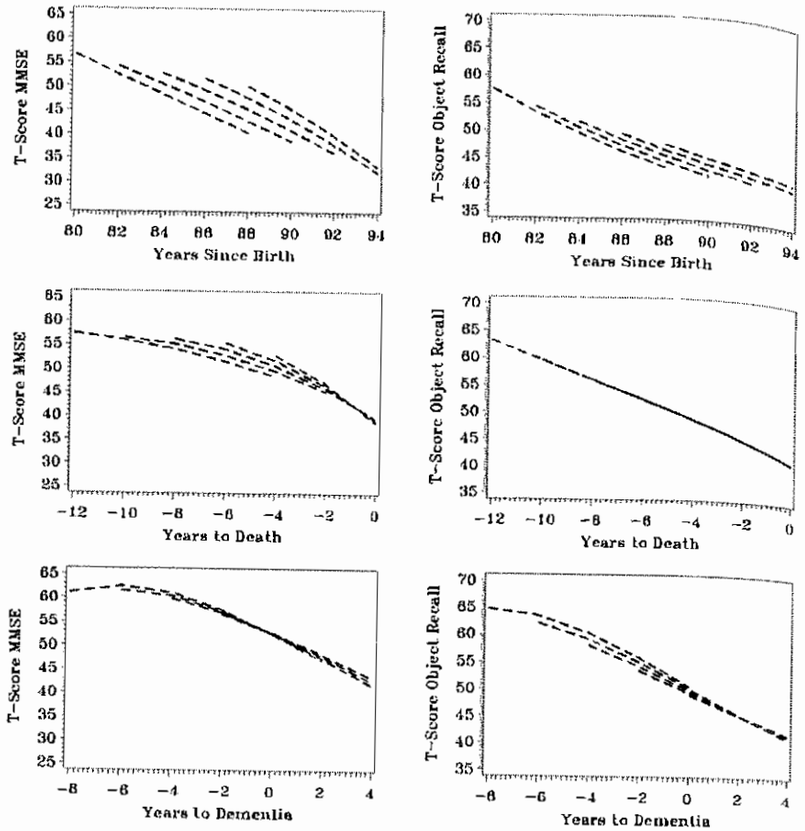

Figure 10.4 Predicted trajectories for Mini-Mental Status Exam (MMSE; left) and Object Recall memory (right) using age as time (top), death as time (middle), and dementia diagnosis as time (bottom).

ated time model—it tries to predict the overall trajectory, even though the resulting trajectory describes no actual observed individual.

Death as Time

Although time since birth is commonly used, an increasingly popular choice is to index change relative to the end of life, rather than relative to the beginning. But because time to death varies across participants at baseline (i.e., its ICC indicated that 24% of its variance was between persons), the same concerns about testing convergence apply to time to death as applied to age. Thus, we can modify the model in equation 10.4 to include time to death as the level-1 time variable (centered at 6 years prior to death)

and time to death at baseline as the level-2 cohort variable. Estimated parameters from the death as time models accounting for death cohort are given in the third columns of Tables 10.1 and 10.2 for MMSE and Object Recall, respectively, and the corresponding model-predicted trajectories are depicted in the middle panels of Figure 10.4. The quadratic slope variance was again not included for either outcome.

The parameters from the death as time model for MMSE followed a very similar pattern as the age as time model. The trajectory toward death for someone who begins the study at 6 years to death (the new reference point) includes an intercept of 55.99, an instantaneous linear rate of decline at that point of 1.02 per year closer to death, and that linear rate of decline becomes more negative by 0.38 per year closer to death (i.e., an accelerating negative function). There were again effects of cohort, such that for every year closer to death one enters the study, the intercept at 6 years prior to death is greater by 0.62, the linear rate of change at 6 years prior to death is less negative by 0.05 (nonsignificant), and the quadratic rate of change is more negative by 0.02. Thus, for MMSE, persons who begin the study closer to death start out higher but have more accelerated decline (even after controlling for current years to death). For Object Recall, however, a significant linear rate of decline of 1.79 per year was predicted across all death cohorts—there was no incremental effect of beginning the study closer to death. Thus, for Object Recall, convergence of the between- and within-person effects of time to death did indeed hold (as shown by the perfectly overlapping lines in the middle right panel of Figure 10.4).

Dementia as Time

Finally, a third potential metric of time (although measured with error) is time to dementia, whereby participants are aligned using their age at diagnosis. Because 71% of its variance was between persons, the same concerns about testing convergence also apply to the effects of time to dementia. Furthermore, the model in equation 10.4 was extended to include a cubic trend as indicated in initial examinations (i.e., it included a β_{3i} placeholder at level 1, defined by just a fixed effect γ_{30} at level 2). Thus, the model included linear, quadratic, and cubic effects of time to dementia as the level-1 time variable (centered at the point of diagnosis), and effects of time to dementia at baseline as the level-2 cohort variable. Estimated parameters from the time to dementia models accounting for dementia cohort are given in the fourth columns of Tables 10.1 and 10.2 for MMSE and Object Recall, respectively, and the corresponding model-predicted trajectories are depicted in the bottom panels of Figure 10.4. None of the slope variances were significant, and so only residual and random intercept variances were retained.

The parameters from the time to dementia model for MMSE indicated an accelerating negative trajectory that eventually began to decelerate (i.e., negative linear and quadratic effects paired with a positive cubic effect), which could reflect range restriction for those scoring very low (i.e., who scored near the floor of MMSE). A significant effect of time to dementia at baseline (dementia cohort) was found only on the cubic trend, such that a lesser amount of reversal of the accelerating negative trend was found for those who began the study further along in the disease progression. For Object Recall, a very similar pattern was observed: an overall accelerating negative trend that abated further into the trajectory, although persons who began the study further along in their dementia progression showed a less negative rate of decline that accelerated more but abated less. As shown in the bottom panels of Figure 10.4, however, these dementia cohort effects do not substantially alter the overall dementia trajectories (i.e., unlike the age cohort effects that do substantially alter the age trajectories in the top panels of Figure 10.4).

Choosing amongst Alternative Metrics of Time

Model Fit

Thus far we have examined three competing variables by which change over time can be indexed: time since birth (age), time to death, and time to dementia. Each of these time metrics implies a different theoretical model and a different means by which different persons can be aligned (or not) onto a single trajectory. The next logical question is, how might one select among these alternative metrics of time? One way to compare alternative models is by using information criteria, such as the Akaike information criteria (AIC) and Bayesian information criteria (BIC) (as estimated under maximum likelihood, given that the models to be compared differ in both their fixed and random effects). Both the AIC and the BIC evaluate the fit of a model relative to the number of parameters estimated, but the BIC also includes a correction based on sample size that rewards greater parsimony. In comparing the AIC and BIC values across models for MMSE (Table 10.1) and Object Recall (Table 10.2), the most preferred model uses dementia as time, followed by death as time, and followed by age as time. This closely matches our earlier intuitions from comparing the mean and individual trajectories across alternative metrics of time (Figures 10.2 and 10.3); in those plots, the mean trajectory through dementia as time seemed to more closely match the individual trajectories.

Variance Components

Another criterion we can utilize to choose between models could be the amount of estimated variance in each model. Up to this point we have focused on the fixed effects from the different time models, but the extent to which the model is a good descriptor of the individual data can also be evaluated by examining the variance left over both between persons and within persons. To facilitate comparison of variance components across models, the models reported in Tables 10.1 and 10.2 were reestimated with only residual and random intercept variances (given that a random slope implies the random intercept variance will change over time, and thus each intercept variance would only apply to the reference point on its time metric). As shown in the top of Figure 10.5, the residual variances

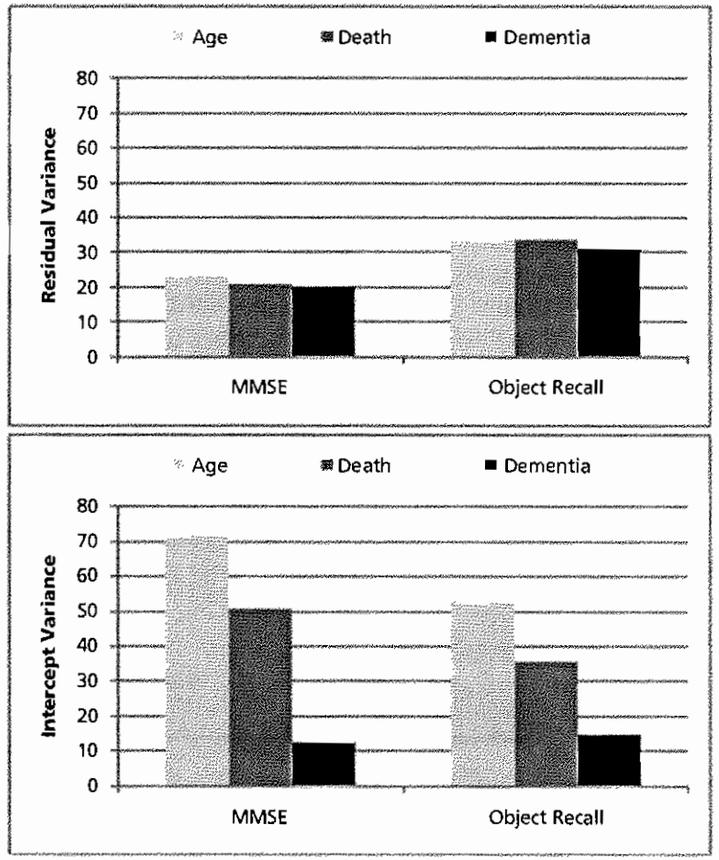

Figure 10.5 Estimated residual variance (top) and random intercept variance (bottom) for MMSE (Mini-Mental Status Exam) and Object Recall memory using age as time, death as time, and dementia diagnosis as time.

differ little between models. This is to be expected given that residual variance represents remaining within-person variation around the predicted trajectory, and within persons, these alternative time metrics are equivalent. Thus greater residual variance could be created by systematic misfit of the form of the trajectory, but these models were selected so as to maximize the fit to the data (i.e., by including higher-order polynomial functions where needed), and thus residual variance should be minimized.

In contrast, as shown in the bottom of Figure 10.5, the random intercept variances did differ markedly between models (even after controlling for cohort) due to differences in how the models aligned different persons onto the same trajectory. Thus, the intercept variance represents remaining individual differences among persons at the same point in "time" (time since birth, time to death, or time to dementia). The intercept variance is greatest for the age model, followed by the death model, followed by the dementia model, for which differences between persons are relatively minimal. Such elimination of individual differences could potentially indicate a well-fitting time metric—for instance, if dementia progression is the relevant causal process, then between-person variability should be minimized once accounting for dementia as time. The substantial age cohort effects can be taken as further evidence that age simply doesn't fit—if it did, age of entry into the study should not matter after controlling for current age (as was more the case in the time to death or time to dementia models in which cohort effects were minimal).

It is also noteworthy that no significant linear or higher-order slope variance was observed in the time to dementia models, indicating that the fixed effects were sufficient to explain the individual variation in change, too. This implies that the individual differences in change that were observed when using time since birth or time to death could have partially resulted from the misalignment of individuals with respect to time. If so, subsequent exploration of those individual differences in change would be misguided at best and misleading at worst.

What about Just Time as "Time"?

Fixed Effects of Time

Thus far we have considered the fit of three alternative metrics of time, and decided that time to dementia appears to have the best relative fit. But what about the fourth possible time metric, simply "time" itself? As discussed earlier, using time in study as time makes no theoretical statement whatsoever about what is responsible for observed change—change is simply specified relative to the baseline observation. Within persons, time is just time. In addition, because time in study contains only longitudinal

information, concerns of testing or assuming convergence do not apply. Thus, using time in study is analogous to group-mean-centering (or person-mean-centering) in multilevel modeling, but here the level-1 effect for time is specified relative to each person's first occasion (rather than to the person's mean time).

To illustrate, consider two equivalent models of linear change across age, in which the first model is specified in terms of time and the second model is specified in terms of age, as shown in equation 10.5:

$$\begin{aligned}
&\text{Time Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}(\text{Age}_{ti} - \text{AgeT1}_i) + e_{ti} \\
&\text{Time Level 2: } \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{AgeT1}_i) + U_{0i} \\
&\qquad\qquad \beta_{1i} = \gamma_{10} \\
&\text{Age Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}(\text{Age}_{ti}) + e_{ti} \\
&\text{Age Level 2: } \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{AgeT1}_i) + U_{0i} \\
&\qquad\qquad \beta_{1i} = \gamma_{10}
\end{aligned} \tag{10.5}$$

in which the level-1 variable for current age has been replaced by current age relative to age at baseline, or simply, "time" in the time-based level-1 model. Age cohort has a main effect on the intercept only in both the time-based and age-based models. The equivalence of these two models can be shown by substituting for the β level-1 placeholders and rearranging common terms into a single-level equation, as shown in equation 10.6:

$$\begin{aligned}
&\text{Time-Based: } y_{ti} = \gamma_{00} + \gamma_{10}(\text{Age}_{ti} - \text{AgeT1}_i) + \gamma_{01}(\text{AgeT1}_i) + U_{0i} + e_{ti} \\
&\text{Time as Age: } y_{ti} = \gamma_{00} + \gamma_{10}(\text{Age}_{ti}) + (\gamma_{01} - \gamma_{10})(\text{AgeT1}_i) + U_{0i} + e_{ti} \\
&\text{Age-Based: } y_{ti} = \gamma_{00} + \gamma_{10}(\text{Age}_{ti}) + \gamma^{*}_{01}(\text{AgeT1}_i) + U_{0i} + e_{ti}
\end{aligned} \tag{10.6}$$

in which the fixed intercept (γ_{00}), the fixed linear within-person effect of age (γ_{10}), the random intercept (U_{0i}), and the level-1 residual (e_{ti}) are the same across models. The effect of age cohort differs predictably across models, in that the incremental between-person effect of age cohort estimated in the age-based model of γ^{*}_{01} (i.e., after controlling for current age) will be produced in the time-based model by subtracting the within-person effect of age from the between-person effect of age cohort ($\gamma_{01} - \gamma_{10}$). This is because the between-person effect of age cohort in the time-based model does not control for current age (i.e., it is the total between-person effect rather than the incremental or contextual between-person effect, as in the age-based model).

Equation 10.6 describes a well-known result in the multilevel modeling literature (Snijders & Bosker, 1999): If the level-1 effect is fixed, the time-based model can be made equivalent to the age-based model, in that it will generate the same predictions and model fit (but with rearranged parameters at level 2). The same is true for any other accelerated time metric, such as time to death or time to dementia. For instance, to include time to death instead, time at level 1 would be the current time to death minus the distance from death at baseline, and distance from death at baseline would be the level-2 cohort variable. The same could be done to include time to dementia instead. Furthermore, although the level-1 age variable was not centered in equation 10.6, centering on any other constant will not change the model fit (it will only change its scale).

Although the equality in equatiou 10.6 is well known, what is less well known is that the fixed-effects model can also be extended to include more complex patterns of change, and yet a time-based version and an accelerated-time version (such as age) can still be made equivalent. Equation 10.7 illustrates this point by adding an interaction of age cohort with the age slope:

$$\begin{aligned}
\text{Time-Based: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti} - \text{AgeT1}_i) + \gamma_{01}(\text{AgeT1}_i) + \\
&\quad \gamma_{02}(\text{AgeT1}_i)^2 + \gamma_{11}(\text{Age}_{ti} - \text{AgeT1}_i)(\text{AgeT1}_i) + U_{0i} + e_{ti} \\
\text{Time as Age: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti}) + (\gamma_{01} - \gamma_{10})(\text{AgeT1}_i) + \\
&\quad (\gamma_{02} - \gamma_{11})(\text{AgeT1}_i)^2 + \gamma_{11}(\text{Age}_{ti})(\text{AgeT1}_i) + U_{0i} + e_{ti} \\
\text{Age-Based: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti}) + \gamma^*_{01}(\text{AgeT1}_i) + \\
&\quad \gamma^*_{02}(\text{AgeT1}_i)^2 + \gamma_{11}(\text{Age}_{ti})(\text{AgeT1}_i) + U_{0i} + e_{ti}
\end{aligned} \tag{10.7}$$

in which a quadratic effect of age cohort is also added in order to maintain equivalency. In equation 10.7 the fixed intercept (γ_{00}), linear within-person effect of age (γ_{10}), age cohort by age interaction (γ_{11}), random intercept (U_{0i}), and level-1 residual (e_{ti}) are the same across models. The linear and quadratic age cohort effects differ across models in the same predictable way as before, in that the age-based model provides direct estimates of the incremental effects of age cohort (after controlling for current age), whereas the time-based model provides direct estimates of the total effects of age cohort instead (not controlling for current age). Finally, these models with a quadratic effect of age at level 1 will also be equivalent, as shown in equation 10.8:

$$\begin{aligned}
\text{Time-Based: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti} - \text{AgeT1}_i) + \gamma_{20}(\text{Age}_{ti} - \text{AgeT1}_i)^2 + \\
&\quad \gamma_{01}(\text{AgeT1}_i) + \gamma_{02}(\text{AgeT1}_i)^2 + \\
&\quad \gamma_{11}(\text{Age}_{ti} - \text{AgeT1}_i)(\text{AgeT1}_i) + \\
&\quad U_{0i} + e_{ti} \\
\text{Time as Age: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti}) + \gamma_{20}(\text{Age}_{ti})^2 + (\gamma_{01} - \gamma_{10})(\text{AgeT1}_i) + \\
&\quad (\gamma_{02} + \gamma_{20} - \gamma_{11})(\text{AgeT1}_i)^2 + (\gamma_{11} - 2\gamma_{20})(\text{Age}_{ti})(\text{AgeT1}_i) + \\
&\quad U_{0i} + e_{ti} \\
\text{Age-Based: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti}) + \gamma_{20}(\text{Age}_{ti})^2 + \gamma^*_{01}(\text{AgeT1}_i) + \\
&\quad \gamma^*_{02}(\text{AgeT1}_i)^2 + \gamma_{11}(\text{Age}_{ti})(\text{AgeT1}_i) + \\
&\quad U_{0i} + e_{ti}
\end{aligned} \tag{10.8}$$

in which the fixed intercept (γ_{00}), linear within-person effect of age (γ_{10}), quadratic within-person effect of age (γ_{20}), random intercept (U_{0i}), and level-1 residual (e_{ti}) are the same across models. The age by age cohort interaction and the linear and quadratic age cohort effects differ across models in the same predictable way as before (i.e., incremental vs. total effects).

Random Effects of Time

The point of showing both the time-based and accelerated-time versions of the same model is this: So long as the level-1 effects are fixed, using time in study (i.e., rather than using direct age in an accelerated-time version) can nevertheless result in equivalent (just slightly rearranged) models. But what if the level-1 effects are random? In this case, we need only examine a simple model to see that the time-based and accelerated-time models cannot be made equivalent, as demonstrated using age as accelerated time in equation 10.9:

$$\begin{aligned}
\text{Time-Based: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti} - \text{AgeT1}_i) + \gamma_{01}(\text{AgeT1}_i) + \\
&\quad U_{0i} + U_{1i}(\text{Age}_{ti} - \text{AgeT1}_i) + e_{ti} \\
\text{Time as Age: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti}) + (\gamma_{01} - \gamma_{10})(\text{AgeT1}_i) + \\
&\quad U_{0i} + U_{1i}(\text{Age}_{ti}) - U_{1i}(\text{AgeT1}_i) + e_{ti} \\
\text{Age-Based: } y_{ti} &= \gamma_{00} + \gamma_{10}(\text{Age}_{ti}) + \gamma^*_{01}(\text{AgeT1}_i) + \\
&\quad U_{0i} + U_{1i}(\text{Age}_{ti}) + e_{ti}
\end{aligned} \tag{10.9}$$

in which a random linear slope for level-1 time/age has been added to allow individuals to differ in their linear rates of change. As shown, in the time-based model, the random slope does not include any baseline variance in time/age, whereas in the age-based model, the random slope does include baseline variance in time/age. Thus, if the level-1 effect of time/age is random, the time-based and accelerated-time models cannot be made equivalent.

So, given the need for a random slope for change over time, which model should be used—a time-based model or an accelerated-time model (e.g., age, time to death, time to dementia)? Previously we used information criteria (ML AIC and BIC) to select amongst the accelerated models (age, death, dementia), and we could do the same to select between the time-based and accelerated-time versions of each index. In this case, though, the AIC and BIC values of the time-based models were largely comparable to their accelerated-time corollaries, which seems to suggest that either version within a given metric (time or accelerated time) would be adequate (although the models for time to dementia fit relatively better than those using age or death).

In discussing this issue for multilevel models more generally, Raudenbush and Bryk (2002, pp. 143–149) present their recommendation to group-mean-center level-1 effects with random slopes, which is analogous to recommending the time-based model variants here. To place their rationale in this context, they note that when substantial variation is observed between persons in the accelerated time metric at baseline, the random intercept variance will likely be estimated with differential precision in the time-based and accelerated-time variants of the same model because of its different interpretation in each. In the time-based model variant, the intercept variance represents individual differences at baseline (or when time = 0), whereas in the accelerated-time model variant, the intercept variance represents individual differences when the accelerated-time metric is 0 (e.g., age 84 or 6 years prior to death). Thus, the intercept in the accelerated-time model variant will require greater extrapolation for those cases in which 0 is not actually observed, resulting in lower reliability and greater shrinkage of the intercept toward the mean. That intercept shrinkage can cause the individual slopes to become homogenized, with the result that the slope variance for the level-1 random effect will be smaller than it should be in the accelerated-time model variants, but accurately estimated in the time-based model variants.

This conjecture was tested in the example data by comparing the slope variance estimates from equation 10.8, in which the time-based versions and accelerated-time versions of the age and death models were specified equivalently in terms of fixed effects, and differed only in their random slopes. The dementia models were not included given that no slope variance

was found. As shown in Figure 10.6, the random slope variance was indeed 33–77% larger in the time-based model variants than in the corresponding accelerated-time variants for both the age and death models, suggesting that the downward bias described by Raudenbush and Bryk (2002) was found in these example data as well. Unless the true model parameters are

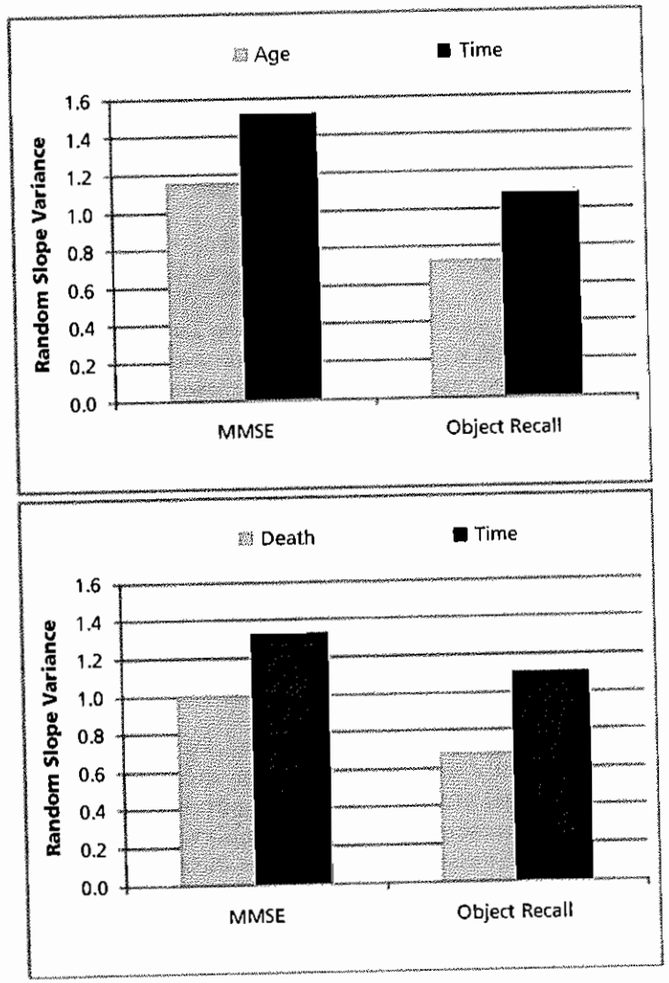

Figure 10.6 Estimated slope variance for MMSE (Mini-Mental Status Exam) and Object Recall memory across the time-based and accelerated-time models using age as time and death as time.

known, however, we cannot be certain if the slope variance in the accelerated-time models was in fact underestimated, or if the greater slope variance reported in the time-based models was actually overestimated instead.

However, these two alternative explanations have also been examined by the author via simulation (Hoffman & Templin, 2008). These simulation results (available upon request) indicated that when data were simulated using a time-based model but analyzed using an accelerated-time model instead, the slope variance was indeed underestimated by up to 50% across conditions, with larger bias found for the conditions with fewer occasions of measurement and with greater variance in the accelerated-time variable at baseline. Given that downward bias in the slope variance should generally limit the power to detect significant predictors of change, this suggests that the time-based versions of the accelerated-time models should be preferred.

Alternative Metrics of Time as Competing Theories

Let us summarize what we have learned so far about using time as "time": if used by itself, time in study is an uninformative time metric, in that change is specified relative to the baseline value (and not relative to a given distance from birth, death, or diagnosis). To make time in study informative for representing differential processes or cause of change, equations 10.5–10.9 paired time in study (at level 1) with age at baseline (age cohort at level 2) in order to reproduce the parameters from the age as time models. Time to death at baseline or time to dementia at baseline could have similarly been included instead at level 2 in order to reproduce those alternative time models. Furthermore, although the time-based and accelerated-time model variants can be made equivalent in their fixed effects, they cannot be made equivalent in their random effects, and it is in accurately quantifying the random slope variance that the time-based model variants seem to have an edge.

But perhaps the most compelling argument for using time as "time" comes from the potential to allow different time metrics (representing different causes of change) to directly compete (or interact) with one another. For instance, why do some people begin the study with lower levels of cognition? Is it their age, distance from death, or their amount of dementia progression? Each time metric's baseline value could be included at level 2 as a competing main effect in order to predict the individual intercepts. Similarly, why do some people decline more rapidly than others? Each time metric's baseline value could also be included as a competing interaction with time in study to also predict the individual slopes. Thus, by using just time as "time" we can obtain a clear and direct estimate of the unique contribution of each alternative temporal process, and we can do so without

any making any assumptions about (or conducting tests for) convergence of its between- and within-person effects.

To illustrate, a final time-as-time model was estimated, as shown in equation 10.10:

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}(\text{Time}_{ti}) + e_{ti} \tag{10.10}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{AgeT1}_i) + \gamma_{02}(\text{DeathT1}_i) + \gamma_{03}(\text{DemT1}_i) + U_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}(\text{AgeT1}_i) + \gamma_{12}(\text{DeathT1}_i) + \gamma_{13}(\text{DemT1}_i) + U_{1i}$$

in which time in study was included at level 1 and in which effects of each alternative time metric (age, death, and dementia) were also included for the intercept and linear time slope (the quadratic time slopes were not significant and were thus not included). Results are given in Table 10.3. As shown in the second column for MMSE, only dementia progression at baseline was significantly related to a lower initial score, and only being closer to death at baseline was significantly related to a greater rate of decline. As shown in the third column for Object Recall, age, time to death, and time

TABLE 10.3 Model Parameters Using Multiple Metrics of Time

	Mini-Mental Status Exam		Object Recall	
Model parameters	**Est**	**SE**	**Est**	**SE**
Fixed effects:				
Intercept (γ_{00})	53.07**	0.50	51.81**	0.73
Linear time slope (γ_{10})	−2.17**	0.14	−1.80**	0.20
Age cohort on intercept (γ_{01})	−0.26	0.15	−0.62**	0.22
Age cohort on slope (γ_{11})	−0.03	0.05	0.13*	0.06
Death cohort on intercept (γ_{02})	−0.06	0.16	−0.56*	0.24
Death cohort on slope (γ_{12})	−0.21**	0.05	−0.06	0.06
Dementia cohort on intercept (γ_{03})	−1.48**	0.11	−1.18**	0.17
Dementia cohort on slope (γ_{13})	0.03	0.03	0.00	0.05
Variance components:				
Residual variance (σ^2_e)	14.50**	1.37	25.96**	2.67
Intercept variance ($\tau^2_{U_0}$)	22.90**	3.89	31.04**	6.42
Linear slope variance ($\tau^2_{U_1}$)	1.18**	0.28	0.87**	0.34
Intercept slope covariance ($\tau_{U_{01}}$)	−2.84**	0.91	−3.53**	1.24
Deviance (−2LL)	3280		2728	
AIC	3304		2752	
BIC	3342		2790	

Note: * $p < .05$, ** $p < .001$. LL, log likelihood; AIC, Akaike information criteria; BIC, Bayesian information criteria. *Cohort* represents the value of each time variable at baseline.

to dementia at baseline were each uniquely related to lower initial memory scores, with the largest effect (in difference per year) found for dementia. Although age significantly predicted rate of decline, older persons were actually predicted to decline less.

CONCLUSIONS

The focus of this chapter was to identify and describe the decisions surrounding the use of multiple potential ways of clocking time in longitudinal studies; that is, alternative metrics of time. Although the decision of what "time" should be is likely to be made on theoretical grounds, in many instances multiple processes may potentially be responsible for observed change, and thus multiple time metrics may need to be considered as a result. Thus, in addition to theoretically motivated choices for time, empirical indices such as differential model fit (ML AIC and BIC) and estimates of between-person heterogeneity may also be useful in comparing amongst multiple plausible metrics of time. In particular, one needs to be aware that greater heterogeneity between persons can be partially due to a misalignment of different individuals with respect to time and, if so, exploring predictors of this heterogeneity may not be informative.

Complicating matters, however, is that persons may differ in "time" at the beginning of a study (i.e., if time is accelerated so that the range of time covered is greater than that observed for any one individual). In this case, one needs to attend to the possibility that the cross-sectional and longitudinal effects of accelerated time may differ, and thus to test for their convergence accordingly. Although models were presented in this chapter to do so (e.g., equation 10.4), these models test for nonconvergence of a particular form (e.g., only a linear effect of age cohort on each growth term), and thus do not preclude the necessity of more complex models to fully disentangle the between- and within-person effects of time (and their possible interactions).

These issues led us to consider a simpler but potentially more useful metric: "time" itself. Although at face value it is the most uninformative choice, specifying time as a function of study duration (i.e., time from baseline) seems to have several advantages. First, because the pattern of change to be approximated by the fixed effects in a time-based model is based solely on within-person variance (and the number of occasions per person), the overall functional form of change can likely be described more parsimoniously (i.e., a linear model may be sufficient given only three or four occasions per person). But when used in combination with person-level predictors representing initial status on other informative metrics of time (e.g., age, time to death, time to dementia), even highly complex trajectories can be represented by a time-based model, given that time-based and accelerated-time versions of the same model can be made equivalent in terms of their fixed effects. In addition, because the random slope variance can he downwardly biased in accelerated-time models, using time-based model variants instead may better recover the true amount of individual differences in change. Furthermore, using time as "time" also permits the inclusion of participants who have not experienced a target event upon which a time metric would otherwise be constructed. For instance, persons without a dementia diagnosis could not be included in a time to dementia model, although they could be included in a time in study model (in which dementia presence and timing could still be included via level-2 predictors).

Finally, perhaps the most compelling support for the use of time as "time" is the fact that it can never be wrong! The use of time in study as time makes no assumptions about why individuals differ a priori, which can be a good thing in absence of strong theory (or when that theory is incorrect). In addition, because time is based solely on longitudinal information, one need not worry about convergence. For instance, in the example data, the between- and within-person effects of age did not converge (i.e., there were effects of age cohort in addition to current age), and the use of age as time produced a pattern of fixed effects that ultimately described the individual data very poorly. This is because using age as time aligned different individuals along a time metric with questionable relevance. In contrast, by using time in study we can instead frame the fit of alternative time metrics as a series of testable hypotheses and easily compare the relative contributions of each. Although the other time metrics could also be added as level-2 predictors in the accelerated time models (e.g., one could estimate an age as time model with time to death as a level-2 predictor), the results would ultimately be less straightforward to interpret, given that the level-2 effects of the accelerated time metric are incremental (i.e., age cohort effects after controlling for current age) whereas the level-2 effects of the other metrics would be total (i.e., total death cohort effects, not after controlling for current years to death), and that age nonconvergence could still be a problem.

In closing, the issues surrounding what time should be and how time should be specified in statistical models for change can be quite complicated. Nevertheless, such deliberations are an important precursor to drawing useful conclusions from longitudinal data, and I hope this chapter will be helpful for those contemplating what time (it) is in their own work.

ACKNOWLEDGMENTS

I would like to thank Scott Hofer, Andrea Piccinin, and Martin Sliwinski for their many helpful discussions and insights on this topic; Jonathan Templin

for his helpful feedback on earlier drafts; and Boo Johansson for the use of the OCTO-Twin data in the example analyses.

REFERENCES

Baltes, P.B., Cornelius, S. W., & Nesselroade, J. R. (1979). Cohort effects in developmental psychology. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 61–87). New York: Academic Press.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1–39). New York: Academic Press.

Bell, R. Q. (1953). Convergence: An accelerated longitudinal approach. *Child Development, 24,* 145–152.

Folstein, M. F, Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State": A practical method for grading the cognitive status of patients for the clinician. *Journal of Psychiatric Research, 17,* 189–198.

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis.* New York: Wiley.

Hofer, S. M., & Sliwinski, M. J. (2006). Design and analysis of longitudinal studies of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (6th ed., pp. 15–37). San Diego, CA: Academic Press.

Hoffman, L., & Stawski, R. (2009). Persons as contexts: Evaluating between-person and within-person effects in longitudinal analysis. *Research in Human Development, 6,* 97–100.

Hoffman, L., & Templin, J. L. (2008, April). *The impact of alternative specifications of time on examining individual differences in change.* Poster presented at the Cognitive Aging Conference, Atlanta, GA.

Johansson, B. (1988). *The MIR-Memory in-Reality Test.* Stockholm, Sweden: Author.

Johansson, B., Hofer, S. M., Allaire, J. C., Maldonado-Molina, M., Piccinin, A. M., Berg, S., et al. (2004). Change in memory and cognitive functioning in the oldest-old: The effects of proximity to death in genetically related individuals over a six-year period. *Psychology and Aging, 19,*145–156.

Johansson, B., Whitfield, K., Pedersen, N. L., Hofer, S. M., Ahern, F., & McClearn, G. E. (1999). Origins of individual differences in episodic memory in the oldest-old: A population-based study of identical and same-sex fraternal twins aged 80 and older. Journal of Gerontology: *Psychological Sciences, 54B,* P173-P179.

McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth curve models for developmental data analysis. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data* (pp. 69–107). Mahwah, NJ: Erlbaum.

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

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin, 64,* 92–107.

Schaie, K. W. (2008). Historical patterns and processes of cognitive aging. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging: Interdisciplinary perspectives* (pp. 368–383). Thousand Oaks, CA: Sage.

Sliwinski, M. J., Hoffman, L., & Hofer, S. M. (2010). Evaluating convergence of within-person change and between-person age differences in age-heterogeneous longitudinal studies. *Research in Human Development, 7,* 45–60.

Snijders, T. A. B., & Bosker, R. (1999). *Multilevel analysis.* Thousand Oaks, CA: Sage.