MULTIVARIATE BEHAVIORAL RESEARCH, *41*(2), 165–187

Cross-Sectional Analysis of Time-Dependent Data: Mean-Induced Association in Age-Heterogeneous Samples and an Alternative Method Based on Sequential Narrow Age-Cohort Samples

Scott M. Hofer, Brian P. Flaherty, and Lesa Hoffman
Department of Human Development and Family Studies
The Pennsylvania State University

The effect of time-related mean differences on estimates of association in cross-sectional studies has not been widely recognized in developmental and aging research. Cross-sectional studies of samples varying in age have found moderate to high levels of shared age-related variance among diverse age-related measures. These findings may be misleading because high levels of association between time-dependent processes can result simply from average population age differences and not necessarily from associations between individual "rates of aging." This is demonstrated both analytically and in a simulation involving cross-sectional sampling of individual trajectories. An alternative cross-sectional narrow age-cohort design is shown to provide a useful alternative for evaluating the interdependence of time-related processes.

> There has been no source more fruitful of fallacious statistical argument than the common influence of the time factor.
>
> Cave and Pearson (1914, p. 354)

The understanding of developmental processes must emphasize change within individuals and with individual differences in rates of change and in the dynamics of such changes over time (e.g., Baltes & Nesselroade, 1979; Wohlwill, 1973). Yet

Correspondence should be sent to Scott M. Hofer, Department of Human Development and Family Sciences, Oregon State University, Corvallis, OR 97331. E-mail: smhofer@gmail.com

the difficulty of following individuals for sufficiently long periods of time in order to observe change, particularly in later adulthood, has led to a large number of cross-sectional studies of aging (with an increasing number of exceptions, see Schaie & Hofer, 2001). As a result, many developmental theories of aging have been based largely on findings from cross-sectional studies, in which direct estimates of individual change are unobserved. Instead, change must be inferred indirectly from analyses of individuals at different ages (e.g., Schneider, Atkinson, & Tardif, 2001; Verhaeghen & Salthouse, 1997), with the result that population average change and any individual differences in change are confounded.

The purpose of this article is to demonstrate both analytically and through simulation work three points regarding analysis of cross-sectional samples of individuals varying in age: (a) that estimates of the associations among age-related differences (as proxy for age-related changes) are biased by mean age differences, (b) that alternative cross-sectional analyses can mitigate this bias and can be performed on data from existing studies, and (c) that statistically controlling for age results in a loss of information regarding associations among age-related changes. In cross-sectional studies of time-dependent variables, covariation between variables can arise from (a) magnitudes and patterns of population average change, or *fixed effects* or *mean trends*, (b) individual differences in rates of change, or *random effects,* and (c) commensurate rates of change within the same individual over time, or *intraindividual covariation*. Because longitudinal studies are needed to examine intraindividual covariation (e.g., Molenaar, 1985; Nesselroade & Schmidt-McCollam, 2000; Wood & Brown, 1994), we focus here on fixed and random effects and their implications for understanding the relation between rates of change within two types of cross-sectional studies: samples varying broadly in age (*age-heterogeneous*), and narrow-age cohort samples (*age-homogeneous*). These issues are applicable to other cross-sectional (i.e., single measurement occasion) designs as well and to covariance analysis of time-dependent processes more generally (quantitative genetic modeling, mediation models, and evaluation of factor structure and invariance; Meredith & Horn, 2001).

CROSS-SECTIONAL ANALYSIS OF AGE-HETEROGENEOUS SAMPLES

Mean Trends and Inferences of Association

That associations between independent processes can be produced by mean trends alone has long been recognized in the time series literature (e.g., Hooker, 1905; Persons, 1917, 1923; Yule, 1921, 1926), as well as in related work demonstrating that random monotone data can produce high values of the squared Pearson correlation coefficient (Parker, Casey, Ziriax, & Silberberg, 1988; Stigler, 1985a,

1985b). To the extent that mean trends are present, estimates of association will be high (the direction of which will depend on the direction of the trend). The essential problem was described by Yule (1903) as "illusory correlation" resulting from the inappropriate aggregation of groups that differ in the means of the correlated attributes, and is closely related to the ecological fallacy (Goodman, 1953; Robinson, 1950), Simpson's paradox (Simpson, 1951), and Lord's paradox (Lord, 1967).

However, that mean trends can also lead to spurious associations in the cross-sectional analysis of time-dependent variables is not as well recognized by developmental researchers. Although several authors have identified the associated problems this poses for single-occasion studies of development and aging (e.g., Hertzog, 1985; Hofer, Berg, & Era, 2003; Hofer & Sliwinski, 2001; Hofer, Sliwinski, & Flaherty, 2002; Kelley, 1928; Kraemer, Yesavage, Taylor, & Kupfer, 2000; Lindenberger & Potter, 1998; Storandt & Hudson, 1975; Wohlwill, 1973; Yule, 1903), cross-sectional analyses of individuals differing in age continue to be the rule rather than the exception. One analysis frequently used within cross-sectional studies of aging concerns the extent to which age-related differences appear common or specific across variables (i.e., decomposition of age-related variance). A recurrent finding is that variance among chronological age and measures of cognitive processing is highly shared, and that a common factor, such as processing speed (i.e., the general slowing hypothesis; Salthouse, 1992; Verhaeghen & Salthouse, 1997) or sensory functioning (i.e., the "common cause" hypothesis; Anstey & Smith, 1999; Baltes & Lindenberger, 1997; Lindenberger & Baltes, 1994; Marsiske, Klumb, & Baltes, 1997; Salthouse, Hambrick, & McGuthry, 1998) is sufficient to account for individual differences in age-related decline across numerous outcomes (Salthouse & Czaja, 2000).

Given that many cognitive and sensory processes decrease substantially with age, however, it is more likely that shared age-related variance arises from average, age-related mean trends than from true associations between rates of change. This poses a significant problem for evaluating developmental theories with cross-sectional data, and perhaps explains why cross-sectional studies with a broad age range usually find large proportions of shared age-related variance and good fit of simple common factor models for various age-related outcomes. Unfortunately, these common factors may not reflect anything substantive about the interdependence of rates of change or the causal dimensionality of such changes, but rather, may simply indicate that change, on average, has occurred (Hofer et al., 2003; Hofer & Sliwinski, 2001; Hofer et al., 2002; Kraemer et al., 2000).

A simple example of how associations may arise solely from mean trends is provided in Figure 1a, which displays two variables, X and Y, that exhibit mean differences across age (to simplify the diagram, only two groups of out a possible continuum of age groups are shown). Assume that the association between X and Y within age groups is zero, and that there is no association between rates of

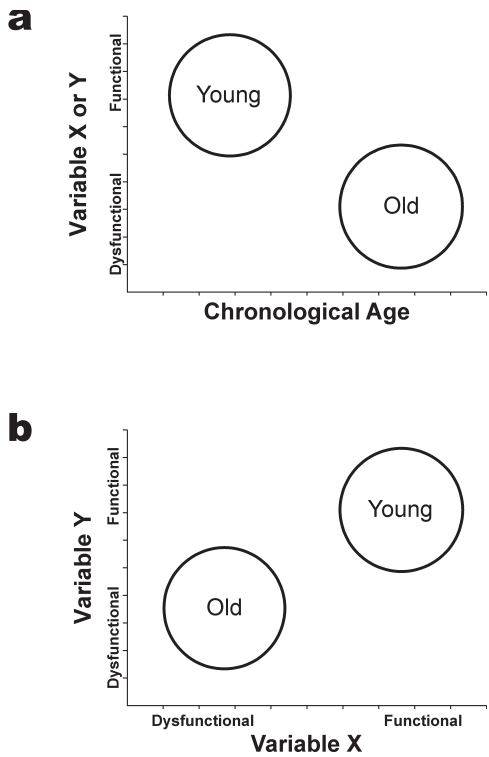

FIGURE 1 Effect of population-level average differences (fixed effects) on association between two age-dependent variables.

change (random effects) within the population. The only similarity between the two variables is that both exhibit average population differences across individuals differing in chronological age. Under such circumstances, Figure 1b shows that the high association between variables X and Y may result exclusively from the mean trends.

Cross-Sectional Sampling of Longitudinal Trajectories: Analytical Results

The following analytical results show how associations between time-dependent variables are at least partly and may be entirely due to mean trends in cross-sectional data. In this example we are concerned with cross-sectional sampling of time-dependent processes and use a simple longitudinal model of change (of which cross-sectional observations are essentially sampling draws of observations

at particular points in time from individual trajectories) to demonstrate how these associations may arise. We use the distinction between fixed and random effects, terminology common to analysis of multilevel data structures, to denote population averages and systematic individual-level deviations from the population averages, respectively (e.g., Laird & Ware, 1982). Suppose that we have a sample of individuals followed continuously over their lifespan such that each individual has a complete realization of scores on two processes, *X* and *Y*, as a function of their initial level (*L*) and rate of change (slope or *S*) over time. Parameters subscripted by *i* refer to an individual's deviation from the population mean (i.e., the random effect). Equation 1 shows that an individual's scores on *X* and *Y* at a time *t* are a function of both fixed and random effects,

$$\begin{aligned} x_{it} &= L_x + L_{xi} + S_x t + S_{xi} t + e_{xi} \\ y_{it} &= L_y + L_{yi} + S_y t + S_{yi} t + e_{yi}\,, \end{aligned} \tag{1}$$

where L_x, L_y are population average intercepts or levels, L_{xi}, L_{yi} are individual *i*s deviations from the population average intercepts, *t* denotes time or age, S_x, S_y are the population average rates of change or slopes, S_{xi}, S_{yi} are individual *i*s deviations from the population average rates of change, and e_{xi}, e_{yi} denote random errors, assumed to be normally distributed and independent.

In an age-heterogeneous, cross-sectional design, a broad range of ages is potentially represented in a sample of individuals differing in age, *t*. Equation 2 results from substituting the population expectation of Equation 1 (where t_i denotes individual *i*s age at time *t*) into the formula for a covariance (see Appendix A for complete derivation). Equation 2 shows that the complexity of the covariation between two time dependent processes, *X* and *Y* is, in part, a function of the fixed effects for population average change in addition to other systematic sources of covariance related to initial individual differences (i.e., intercepts), covariance between rates of change (i.e., slopes), and covariance between intercepts and slopes,

$$\begin{aligned} Cov(X,Y) = E(&L_{xi}L_{yi} + L_{xi}S_y t_i + L_{xi}S_{yi}t_i \\ &+ S_x t_i L_{yi} + S_x t_i S_y t_i + S_x t_i S_{yi} t_i - S_x t_i S_y \overline{t} \\ &+ S_{xi} t_i L_{yi} + S_{xi} t_i S_y t_i + S_{xi} t_i S_{yi} t_i - S_{xi} t_i S_y \overline{t} \\ &- S_x \overline{t} S_y t_i - S_x \overline{t} S_{yi} t_i + S_x \overline{t} S_y \overline{t}\,). \end{aligned} \tag{2}$$

To demonstrate further the potential influence of mean trends on covariance between processes, consider a situation in which processes are independent, but in which both exhibit systematic average change over time. Because there are no associations between *X* and *Y*, all terms involving random effects (terms subscripted by *i*) are zero and can be dropped. This condition is shown in Figure 1, where *X* and

Y are not associated within age groups but where both *X* and *Y* exhibit mean differences across age. With only fixed effects present, Equation 2 simplifies to Equation 3, and the resulting covariance simplifies to the product of the variance of age in the sample multiplied by the average rates of change in processes *X* and *Y*,

$$Cov(X,Y) = S_x S_y Var(t_i). \tag{3}$$

Therefore, as the age range increases, the variance of *t* will increase, as will the effect that mean trends have on the covariance. Therefore, covariation will result from mean trends in cross-sectional samples varying in age even when individual differences in rates of changes are completely independent. Additionally, we expect that stronger covariances will obtain in samples that vary more broadly in age, when the rates of change are larger, or both. It is important to recognize that although Equations 2 and 3 are based on a simple linear model of change for both processes, a more complex model of change would simply introduce more terms and would not alter the main point of this derivation, namely, that fixed effects are a component of covariance in age-heterogeneous, cross-sectional designs. Indeed, the addition of more complex models of change to the derivation of the cross-sectional covariance does not diminish the general problem but rather leads to even greater difficulty in understanding individual differences in change.

Cross-Sectional Sampling of Longitudinal Trajectories: Simulation Results

To provide a further exemplar of the analytical expectations from age-heterogeneous cross-sectional designs, a simulation was performed in which cross-sectional estimates of the correlation between two time-dependent processes were compared with the correlation between random effects specified in the simulated data. Repeated measures data were generated that conform to patterns of data typically analyzed in studies of aging based on the linear model shown in Equation 1 (further details of the simulation are provided in Appendix B). A six-way analysis of variance was used to evaluate the salient simulation conditions that contributed to differences in the size of the correlation between *X* and *Y* in age-heterogeneous samples. No four-way or higher-order interactions were significant. Although many effects were significant, the two major contributing factors were magnitude of slope correlation and mean slope, which accounted for 46% and 30% of the total sums of squares, respectively. Other contributing factors accounting for 2% to 6% of the total sums of squares included initial level correlation, variance in slope, error, and two interaction terms of slope correlation with mean slope and with slope variance.

To illustrate we describe the outcome from one simulation condition in which there was a low positive correlation between individual slopes (.30), a high amount of change over time (mean slope = –8 per unit time), low error variance (1), low interindividual variation in slopes (5 and 4, for *X* and *Y*, respectively), and no association between initial levels or between level and slope. Under age-heterogeneous, cross-sectional sampling (drawing one observation at random from each trajectory), the estimated *X*-*Y* correlation was .76, a substantial overestimate of the specified value of .30. Across conditions, we observed moderate to highly positive correlations and covariances even when the generated correlation was zero or negative (e.g., observed .43, .52, .62, and .74 instead of –.30, .00, .30, and .70, respectively). Standardized biases (the standard deviation from the true parameter; see Appendix B) further quantify the substantial bias in the estimates from the age-heterogeneous, cross-sectional sample. When the slope correlation was –30, the standardized bias was 2121. For the slope correlation conditions .00, .30, and .70, standardized bias estimates were 1745, 1304, and 363, respectively. Only when the generated correlation was high and positive (.70) was the observed correlation a close approximation. It is also interesting to note that greater error in age-heterogeneous cross-sectional sampling of time-dependent data lead to less-biased estimates of the slope correlation. Further results of the simulation are available from Scott Hofer at http://www.smhofer.net.

Summary

These analytical and simulation results demonstrate how associations between time-dependent processes may arise from trends in mean level even in the absence of correlated rates of change and correlated initial individual differences. This presents a fundamental problem for the evaluation of interdependency between age-related processes within age-heterogeneous, cross-sectional designs, as there is no way to disentangle mean trends from other possible sources of covariation. Thus, associations between variables that change with age on average should not be taken as evidence for a common causal aging mechanism, given that such associations could have been produced solely by simple mean age trends in the population.

CROSS-SECTIONAL ANALYSIS OF AGE-HOMOGENEOUS SAMPLES

Mean Trends and Inferences of Association

An alternative type of cross-sectional design, the analysis of narrow-age cohorts, examines associations between variables within and across age-homogeneous

samples (i.e., individuals of the same or nearly the same chronological age; e.g., Era et al., 1996; Heikkinen, Berg, Schroll, Steen, & Viidik, 1997; Wohlwill, 1973). Ideally, in age-homogeneous samples, age is a constant in order to minimize the influence of mean trends on the within-group covariance. If associations are calculated within groups with little to no variance in age, fixed effects or mean age trends will not bias estimates of association between processes (in narrow-age cohorts with some within-group age variance, the influence of mean trends on estimates of association can be eliminated by partialing for age within groups).

The particular utility of sequential narrow-age cohort designs (i.e., multiple age-homogeneous groups) for examining interdependency of age-related processes relies on the existence of individual differences in rates of change. As time elapses, the rank ordering of individuals at a given cross-section (e.g., age 8 or age 70) will increasingly reflect the rank ordering of individual rates of change. Correlated rates of change on several different processes will be observed as moderate and increasing covariances across narrow age-cohort samples of increasing age. As a result, sequential narrow-age cohort designs are more suitable for examining the interdependence of aging-related change than are age-heterogeneous designs. Sequential narrow-age cohort designs have already been used to evaluate differentiation of cognitive functioning in childhood and dedifferentiation in late life (e.g., Balinsky, 1941; Garrett, Bryan, & Perl, 1935; Garrett, 1946; Lindenberger & Baltes, 1997; Reinert, 1970; Salthouse, Hancock, Meinz, & Hambrick, 1996), the finding of decreasing or increasing covariance between age-related processes as a function of correlated rates of development or aging, respectively. In addition, this design has formed the basis for several gerontological studies (e.g., Era, 1987; Era et al., 1996; Hofer, Berg, & Era, 1998, 2003).

Cross-Sectional Sampling of Longitudinal Trajectories: Analytical Results

Beginning with Equation 2, we derive the covariance for a narrow age-cohort sample by omitting terms for fixed effects. Equation 4 shows the general case for a covariance within a sample of individuals of exactly the same age (t is no longer subscripted because it is constant within an age-group),

$$Cov(X,Y) = Cov(L_{xi}L_{yi}) + [t]Cov(L_{xi}S_{yi}) + [t]Cov(S_{xi}L_{yi}) + [t^2]Cov(S_{xi}S_{yi}). \quad (4)$$

As shown, the covariance of any single narrow-age cohort will be a function of covariance between levels and slopes, $[t]Cov(L_{xi}S_{yi})$, $[t]Cov(S_{xi}L_{yi})$, initial level covariance, $Cov(L_{xi}L_{yi})$, and covariance related to rates of change, $[t^2]Cov(S_{xi}S_{yi})$. Average population change (fixed effects) will not enter into the estimate of association.

In a narrow-age cohort sample, associations between variables may arise from initial individual differences as well as common rates of aging. We must assume that intraindividual change due to development or aging overwhelms any initial (i.e., early childhood or adulthood) individual differences in functioning, and that the rank order across individuals of same age (i.e., within a narrow-age cohort) will become more and more informative regarding the associations between aging-related rates of change with the passage of time. Equation 5 shows how as time increases, the effect of the covariance between the rates of change in the processes increases as a function of t^2, whereas the other sources of covariance increase only as a function of t,

$$t \rightarrow Cov(X,Y) = Cov(L_{xi}L_{yi}) + [t]Cov(L_{xi}S_{yi}) + [t]Cov(S_{xi}L_{yi}) + [t^2]Cov(S_{xi}S_{yi})$$

$$\begin{aligned}
t = 1 &\rightarrow Cov(X,Y) = Cov(L_{xi}L_{yi}) + [1]Cov(L_{xi}S_{yi}) + [1]Cov(S_{xi}L_{yi}) + [1]Cov(S_{xi}S_{yi}) \\
t = 2 &\rightarrow Cov(X,Y) = Cov(L_{xi}L_{yi}) + [2]Cov(L_{xi}S_{yi}) + [2]Cov(S_{xi}L_{yi}) + [4]Cov(S_{xi}S_{yi}) \\
t = 3 &\rightarrow Cov(X,Y) = Cov(L_{xi}L_{yi}) + [3]Cov(L_{xi}S_{yi}) + [3]Cov(S_{xi}L_{yi}) + [9]Cov(S_{xi}S_{yi}) \\
t = 4 &\rightarrow Cov(X,Y) = Cov(L_{xi}L_{yi}) + [4]Cov(L_{xi}S_{yi}) + [4]Cov(S_{xi}L_{yi}) + [16]Cov(S_{xi}S_{yi}). \quad (5)
\end{aligned}$$

If the covariation among the rates of change is nonzero and moderate, it will quickly overwhelm the contributions of the covariances involving initial level. The observed association between processes will increasingly reflect individual differences in the rates of change, and will only be zero when initial individual differences are cancelled by the covariance among the rates of change, or when the true covariance among the rates of change is indeed close to zero. Based on this reasoning, the within-group correlations across sequential narrow-age samples would be expected to increase if the rates of change in the outcomes are correlated.

Cross-Sectional Sampling of Longitudinal Trajectories: Simulation Results

A seven-way analysis of variance (the six previous factors, plus time) was used to evaluate the salient simulation conditions that contributed to differences in the size of the correlation between X and Y in age-homogeneous samples. The main effect of time alone accounted for 92% of the total sums of squares, and interactions of time with mean slope and with slope variance accounted for an additional 7%. Faster convergence across time towards the specified X-Y correlations resulted from greater magnitude and variability of change, and thus faster movement away from initial patterns of covariation that are present earlier in the time sequence.

Age-homogeneous cross-sectional samples were selected from times 2 and 8 from the longitudinal trajectories generated in the same simulation condition as described previously (i.e., correlation between slopes = .30, mean slope = –8 per

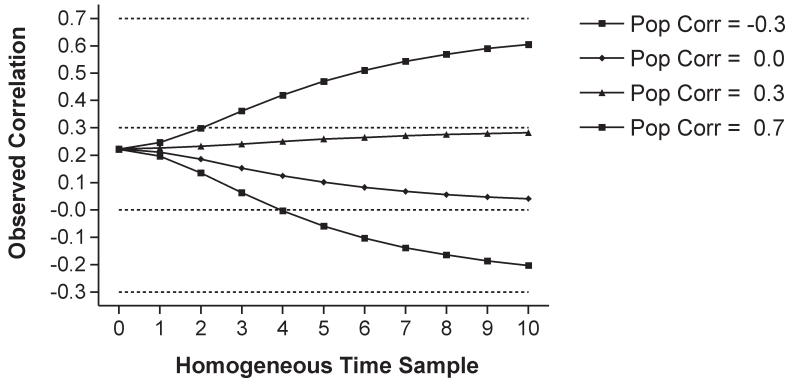

FIGURE 2 Observed correlations between rates of change (random slope effects) across age homogeneous samples at four population values.

unit time, error variance = 1, slope variance for $X = 5$, for $Y = 4$; see Appendix B). The obtained *X*-*Y* correlation at time 2 was near zero as expected, given the initial level correlation of zero. With time, however, the *X*-*Y* correlation increasingly reflected the covariance between rates of change, as observed at time 8, where the *X*-*Y* correlation was .30 as specified in the data. As shown in Figure 2, convergence towards the generated slope correlations was attained at later points in time across all conditions.

Another important feature of sequential narrow-age cohort designs is that they permit evaluation of increases or decreases in covariation arising from cumulative effects of correlated rates of change, as shown graphically in Figure 3. Thus, sequential narrow-age cohort designs can distinguish covariance arising from initial individual differences and common or unique rates of aging within individuals from covariance arising from between-person differences, and may lead to different conclusions than those from age-heterogeneous analyses. For example, in narrow-age cohort population samples of individuals aged 75 years from three countries, Hofer, Berg, and Era (1998, 2003) found no evidence for the common-cause hypothesis—few associations were found between sensory acuity, balance, and cognitive processing, in contrast to previous research. Similarly, Sliwinski and Buschke (1999) reported that while statistical control of processing speed greatly attenuated cross-sectional age effects, it did not attenuate longitudinal effects. If strong associations found between variables in age-heterogeneous cross-sectional samples are not found within age-homogeneous samples where mean trends can be controlled, or in longitudinal analyses in which correlated rates of change can be estimated separately from mean trends, then one must entertain the notion that the association between age-related processes merely reflects the fact that both variables change, on average, over time.

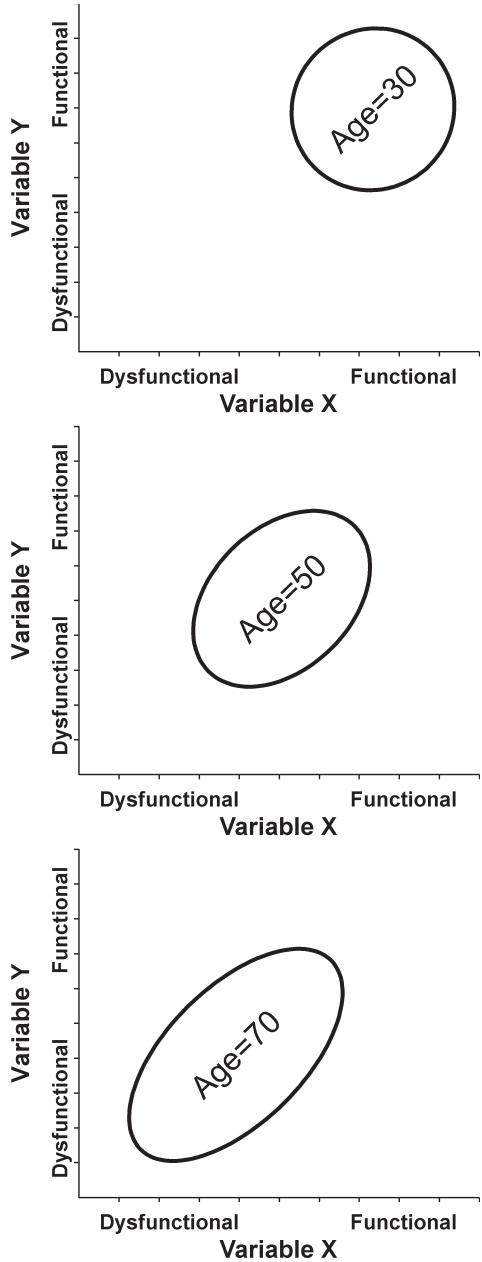

FIGURE 3 Sequential narrow age-cohort (SNAC) design demonstrating effect of common rates of aging.

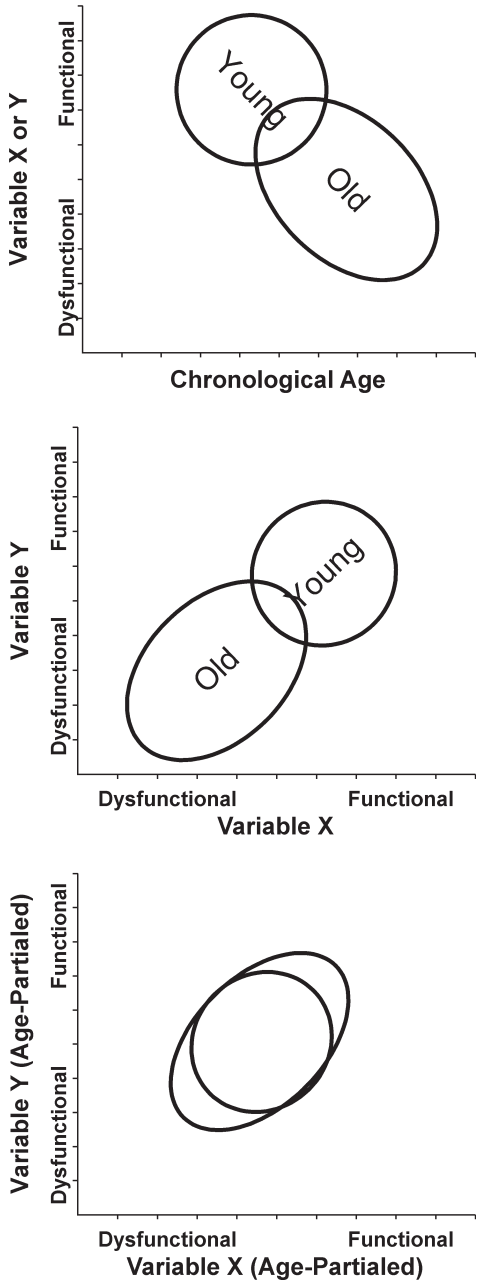

FIGURE 4 Effect of chronological age regression partial on association between two age-dependent variables.

THE UTILITY OF STATISTICALLY PARTIALING FOR CHRONOLOGICAL AGE

An approach often taken to reduce the influence of mean trends within age-heterogeneous, cross-sectional studies is the statistical partialing of the effect of age from the association between processes (i.e., the removal of age-related effects resulting from covariance due to mean differences across individuals varying in age). However, the partial correlation between *X* and *Y* controlling for age, $r_{xy.age}$, is the same as the zero-order correlation for a narrow-age cohort sample at the mean age of the age-heterogeneous sample (Lord, 1963). A similar result would be obtained by aggregating the zero-order covariances across the multiple narrow-age cohorts that comprise the full age-heterogeneous sample. The cumulative influences of correlated rates of change (i.e., increasing or decreasing covariation across time) still remain, however, and are simply averaged over in the age-partial analysis of an age-heterogeneous sample to produce an estimate of what the association would be at the average age of the sample, as shown in Figure 4. Thus, the age-partialing approach is not making full use of the available data, in that information regarding the changing patterns of covariation between processes across time is lost.

Variance decomposition approaches for estimating the proportion of *common* or *shared age-related variance* are also problematic in the age-heterogeneous, cross-sectional design. Indeed, what is estimated as shared age-related variance is confounded with any mean trends present in the data. Sources of common variance, observed in terms of shared age-related covariance and in common factors associated with age, provide little information regarding individual differences in rates of change when age-related mean trends are present. For this reason, when only cross-sectional data are available, we view the analysis of sequential narrow-age cohort samples as more useful in examining covariances among age-related variables, given that direct estimates of the extent to which covariance between processes increases or decreases with time can be represented explicitly (i.e., with product terms for Age Group × Process interactions).

DISCUSSION

Although cross-sectional studies of development and aging necessarily focus on age-related individual differences and so confound age differences with age changes, they remain one of the most utilized designs for understanding developmental processes, particularly that of aging (Schaie & Hofer, 2001). Variance decomposition has played a central role in the analysis of cross-sectional samples varying broadly in age, for cognitive aging research in particular. For example, Salthouse and Czaja (2000) reported high magnitudes of shared age-related covariance in diverse cognitive processes and that the "current results strongly

suggest that broad explanatory mechanisms play an important role in the age-related effects found in many cognitive variables" (p. 54). The problem is that many researchers assume that they are studying aging-related changes in individuals, and that cross-sectional estimates of the association between processes are useful in this regard. In this article we sought to clarify the analytical basis for estimating the interdependence between time-dependent processes, as well as the interpretational basis when associations are estimated within age-heterogeneous and age-homogeneous cross-sectional designs. We offer the following conclusions on the basis of our analytical and simulation results:

1. *Analysis of cross-sectional samples varying broadly in age will result in upwardly biased estimates of association between processes exhibiting mean trends over time.* This bias in estimates of association results from the contribution of fixed effects or mean trends, in addition to random effects, to an individual's score at a given time. The population mean trend will enter into the covariance in age-heterogeneous samples as a function of the variability of the age range sampled and the magnitude of average change. Because of this, we regard the evidence based on age-heterogeneous cross-sectional samples, when used for understanding covariance of aging-related change, to be highly misleading. Cross-sectional analyses can be useful, however, in examining mean trends per se (i.e., experimental designs comparing young and old individuals), although Faust, Balota, Spieler, and Ferraro (1999) have also demonstrated how group differences in response time can often overshadow individual differences and produce spurious overadditive interactions. Finally, variables that change little over time within age-heterogeneous samples may not have mean trends as a major source of their covariances, and thus estimates of association between such processes may be less biased.

2. *In analyses of cross-sectional samples of narrow-age cohorts, the influence of mean trends is removed from estimates of covariance for a given cohort, resulting in substantially less bias in the estimates of correlations between rates of change than in age-heterogeneous samples. Further, changing patterns of covariance across age (i.e., dedifferentiation) arising from correlated rates of change can be observed directly across cohorts.* For these reasons, we regard the analysis of narrow-age cohorts to be more informative than the analysis of age-heterogeneous samples in evaluating correlations between rates of aging and the generality or specificity of aging-related changes. An additional strength of the narrow-age cohort design is that it permits analysis of existing age-heterogeneous, cross-sectional data, either by analyzing a single narrow-age cohort or by carving a sample varying broadly in age into sequential narrow-age cohorts. For analysis of narrow-age cohorts in which age is not exactly constant, one could partial for age within subsamples, and then analyze associations of age-partialed estimates across narrow-age cohorts. Note that this is *not* the same as partialing for age overall in age-heterogeneous cross-sectional samples, which we do not recommend.

3. *Regression analyses that partial for chronological age in age-heterogeneous samples estimate associations at the average age of the sample and ignore age-related changes in patterns of association, are thus, are not highly informative regarding the interdependence between rates of change.* We have argued that the most informative information regarding cross-sectional associations among age-related outcomes is the interaction of age (or age-group) on the magnitude of correlations among the outcomes of interest (Hofer, Sliwinski, & Flaherty, 2002). The estimation of shared age-related variance from regression decomposition analysis may be primarily informative about the mean trends across outcome variables and this is a poor basis for understanding individual differences in aging. Certainly, the description and comparison of mean trends is a fundamental aspect of understanding general change in the population but this is separate from an analysis of individual differences. Additionally, because there is only one measure for each individual in a cross-sectional study, we must assume that population trends can appropriately describe the level of each individual at a given age.

Finally, we offer the following observations about analysis of longitudinal samples that begin with individuals varying broadly in age. Despite providing direct information on within-individual change, variation, and covariation of aging-related processes, the same problem of mean-induced associations can also affect associations between rates of change in longitudinal samples as in cross-sectional samples. Rates of change may correlate due to age-based periods of relatively greater change. Indeed, longitudinal designs that begin with a relatively homogeneous age sample at the first occasion may permit a clearer interpretation of the associations in much the same way as the cross-sectional age-homogeneous designs permit. Researchers might consider employing multiple cohorts that are relatively homogeneous in age (e.g., cohorts of 20, 35, 50, 65, and 80 years in age), which would provide ample sample size for age-homogeneous cross-sectional analyses and could also serve as a foundation for longitudinal investigations. Other important sampling issues, such as cohort differences and population mortality, will influence results from age-heterogeneous designs, including the inspection of cross-group levels of association.

CONCLUSION

In addition to the often cited statement that "correlation does not imply *causation*," our main message is that "correlation does not always imply *association*." Finding associations of scores in samples that vary in age does not imply association between rates of change *at the level of the individual*. The analytic and simulation results reported here have demonstrated that analysis of variables that exhibit mean change over time — typical outcomes in developmental and aging studies — will usually yield an overestimate of the association between processes unless the true

association is positive and high. High magnitudes of association can obtain due to mean trends even when the population association among rates of change is zero or negative.

The potential for the confounding of fixed and random effects in estimates of association appears to be profound in developmental and aging research. The interpretations from many of these investigations have led researchers to hypothesize that general, common causal effects produce many of the observed age-related changes across a variety of functions. It is our opinion that many of the associations between age-related variables reported in the literature may be at least upwardly biased in terms of their importance, and some may be completely spurious.

We present an alternative solution, based on analysis of age-homogeneous cohorts, which provides what we consider to be a superior cross-sectional approach for understanding the relative interdependence of aging-related changes within individuals. The analysis of narrow-age cohorts offers certain strengths over the age-heterogeneous analysis, mainly that it eliminates covariance arising from mean trends present in developmental data. While we do not regard the age-homogeneous sampling design as a strong alternative to longitudinal designs and analysis, it does permit the reanalysis of new and existing cross-sectional samples and permits an alternative evaluation of developmental and aging hypotheses.

ACKNOWLEDGMENTS

Please note that the complete simulation results are available at http://www.personal.psu.edu/faculty/s/m/smh21/

Scott M. Hofer is now at the Department of Human Development and Family Sciences, Oregon State University. Brian P. Flaherty is now at the Department of Psychology, University of Washington. Lesa Hoffman is now at the Department of Psychology, University of Nebraska.

This research was initially reported at the 1998 Gerontological Society of America Conference, Philadelphia, PA, and subsequently at the 2000 Cognitive Aging Conference, Atlanta, GA and Gerontological Society of America Conference, Washington, DC.

We thank John J. McArdle for suggestions regarding the analytical formulation and also Fumiaki Hamagami, Jennie Noll, Andrea Piccinin, and Martin Sliwinski for helpful comments on previous versions of this article.

REFERENCES

Anstey, K. J., & Smith, G. A. (1999). Interrelationships among biological markers of aging, health, activity, acculturation, and cognitive performance in late adulthood. *Psychology and Aging*, *14*, 605–618.

Balinsky, B. (1941). An analysis of the mental factors of various age groups from nine to sixty. *Genetic Psychology Monographs*, *23*, 191-234.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life-span: A new window to the study of cognitive aging? *Psychology and Aging*, *12*, 12–21.

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.

Cave, B. M., & Pearson, K. (1914). Numerical illustrations of the variate difference correlation method. *Biometrika*, *10*, 340–355.

Era, P. (1987). Sensory, psychomotor, and motor functions in men of different ages. *Scandinavian Journal of Social Medicine: Supplementum, 39*, 9–67.

Era, P., Schroll, M., Ytting, H., Gause-Nilsson, I., Heikkinen, E., & Steen, B. (1996). Postural balance and its sensory-motor correlates in 75-year-old men and women: A cross-national comparative study. *Journal of Gerontology: Medical Sciences*, *51A*, M53–M63.

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*, 777–779.

Garrett, H. E. (1946). A developmental theory of intelligence. *American Psychologist*, *1*, 372–378.

Garrett, H. E., Bryan, A. I., & Perl, R. E. (1935). The age factor in mental organization. *Archives of Psychology*, *176*, 5–31.

Goodman, L. A. (1953). Ecological regressions and behavior of individuals. *American Sociological Review*, *18*, 663–664.

Heikkinen, E., Berg, S., Schroll, M., Steen, B., Viidik, A. (1997). Functional status, health, and aging: The NORA Study. *Facts*, *Research*, *and Intervention in Geriatrics*, New York: Springer Publishing.

Hertzog, C. (1985). An individual differences perspective: Implications for cognitive research in gerontology. *Research on Aging*, *7*, 7–45.

Hofer, S. M., Berg, S., & Era, P. (1998, November). *Evidence for independent aging effects on perceptual acuity*, *balance*, *and cognitive capabilities: The NORA study.* Paper presented at the annual meeting of the Gerontological Society of America, Philadelphia.

Hofer, S. M., Berg, S., & Era, P. (2003). Evaluating the interdependence of aging-related changes in visual and auditory acuity, balance, and cognitive functioning. *Psychology and Aging*, *18*, 285–305.

Hofer, S. M., & Sliwinski, M. J. (2001). Understanding ageing: An evaluation of research designs for assessing the interdependence of ageing-related changes. *Gerontology*, *47*, 341–352.

Hofer, S. M., Sliwinski, M. J., & Flaherty, B. P. (2002). Understanding aging: Further commentary on the limitations of cross-sectional designs for aging research. *Gerontology*, *48*, 22–29.

Hooker, R. H. (1905). On the correlation of successive observations. *Journal of the Royal Statistical Society*, *68*, 696–703.

Kelley, T. L. (1928). *Crossroads in the mind of man*. Stanford: Stanford University Press.

Kraemer, H. C., Yesavage, J. A., Taylor, J. L., & Kupfer, D. (2000). How can we learn about developmental processes from cross-sectional studies, or can we? *American Journal of Psychiatry*, *157*, 163–171.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*, 963–974.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging*, *9*, 339–355.

Lindenberger, U., & Baltes, P. B. (1997). Intellectual functioning in old and very old age: Cross-sectional results from the Berlin Aging Study. *Psychology and Aging*, *12*, 410–432.

Lindenberger, U., & Potter, U. (1998). The complex nature of unique and shard effects in hierarchical linear regression: Implications for developmental psychology. *Psychological Methods*, *3*, 218–230.

Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 21–38). Madison: The University of Wisconsin Press.

Lord, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, *68*, 304–305.

Marsiske, M., Klumb, P., & Baltes, M. M. (1997). Everyday activity patterns and sensory functioning in old age. *Psychology and Aging*, *12*, 444–457.

Mathsoft (1997). *S-PLUS user's guide*, *version 4.0*. Seattle, WA: Author.

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

Meredith, W., & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins (Ed.), *New methods for the analysis of change* (pp. 203–240). Washington, DC: American Psychological Association.

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

Nesselroade, J. R., & Schmidt-McCollam, K. M. (2000). Putting the process in developmental processes. *International Journal of Behavioral Development*, *24*, 295–300.

Parker, S., Casey, J., Ziriax, J. M., & Silberberg, A. (1988). Random monotone data fit simple algebraic models: Correlation is not confirmation. *Psychological Bulletin*, *104*, 417–423.

Persons, W. M. (1917). On the variate difference correlation method and curve-fitting. *Publications of the American Statistical Association*, *15*, 602–642.

Persons, W. M. (1923). Correlation of time series. *Journal of the American Statistical Association*, *18*, 713–726.

Reinert, G. (1970). Comparative factor analytic studies of intelligence through the human life-span. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory* (pp. 468–485). New York: Academic.

Robinson, W. S. (1950). Ecological regressions and behavior of individuals. *American Sociological Review*, *15*, 351–357.

Salthouse, T. A. (1992). Influence of processing speed on adult age differences in working memory. *Acta Psychologia*, *79*, 155–170.

Salthouse, T. A., & Czaja, S. J. (2000). Structural constraints on process explanations in cognitive aging. *Psychology and Aging*, *15*, 44–55.

Salthouse, T. A., Hambrick, D. Z., & McGuthry, K. E. (1998). Shared age-related influences on cognitive and noncognitive variables. *Psychology and Aging*, *13*, 486–500.

Salthouse, T. A., Hancock, H. E., Meinz, E. J., & Hambrick, D. Z. (1996). Interrelations of age, visual acuity, and cognitive functioning. *Journal of Gerontology: Psychological Sciences and Social Sciences*, *51B*, P317–P330.

Schaie, K. W., & Hofer, S. M. (2001). Longitudinal studies in aging research. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the Psychology of Aging* (pp. 53–77). San Diego: Academic.

Schneider, B. H., Atkinson, L., & Tardif, C. (2001). Child-parent attachment and children's peer relations: A quantitative review. *Developmental Psychology*, *37*, 86–100.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *American Statistician*, *13*, 238–241.

Sliwinski, M. J., & Buschke, H. (1999). Cross-sectional and longitudinal relationships among age, cognition, and processing speed. *Psychology and Aging*, *14,* 18–33.

Stigler, S. M. (1985a). Mass extinction times and correlations. *Nature*, *314*, 586.

Stigler, S. M. (1985b). Terrestrial mass extinctions and galactic plane crossings. *Nature*, *313*, 159.

Storandt, M., & Hudson, W. (1975). Misuse of analysis of covariance in aging research and some partial solutions. *Experimental Aging Research*, *1*, 121–125.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analysis of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin*, *122*, 231–249.

Wohlwill, J. B. (1973). *The study of behavioral development*. New York: Academic.

Wood, P., & Brown, D. (1994). The study of intraindividual differences by means of dynamic factor models: Rationale, implementation, and interpretation. *Psychological Bulletin*, *116*, 166–186.

Yule, G. U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika*, *2*, 121–134.
Yule, G. U. (1921). On the time-correlation problem, with especial reference to the variate-difference correlation method. *Journal of the Royal Statistical Society*, *84*, 497–537.
Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, *89*, 1–63.

Accepted May 2005

APPENDIX A
Covariance Expectation in Age-Heterogeneous and Age-Homogeneous Cross-Sectional Designs based on Simple Linear Model of Change

Covariance Expectation in Age-Homogeneous Cross-Sectional Designs

Equation A1 shows a simple linear model of change such that a particular individual's scores on X and Y at a time t are a function of both fixed and random effects (parameters subscripted by i refer to an individual's deviation from the population mean),

$$\begin{aligned} x_{it} &= L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi} \\ y_{it} &= L_y + L_{yi} + S_y t_i + S_{yi} t_i + e_{yi}. \end{aligned} \tag{A1}$$

Where L_x, L_y are population average intercepts or levels, L_{xi}, L_{yi} are individual is deviations from the population average intercepts, t denotes time or age, S_x, S_y are the population average rates of change or slopes, S_{xi}, S_{yi} are individual is deviations from the population average rates of change, and e_{xi}, e_{yi} denote random errors, assumed to be normally distributed and independent. Equation A3 shows the covariance between two time dependent processes, X and Y, which is the result of substituting the population expectation of Equation A1 (where t_i denotes an individual is age at time t) into the formula for a covariance shown as Equation A2,

$$Cov(X,Y) = E[(X - \mu_x)(Y - \mu_y)], \tag{A2}$$

and

$$\begin{aligned} Cov(X,Y) = E([(L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi}) - E(L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi})] \\ [(L_y + L_{yi} + S_y t_i + S_{yi} t_i + e_{yi}) - E(L_y + L_{yi} + S_y t_i + S_{yi} t_i + e_{yi})]). \end{aligned} \tag{A3}$$

Equation A4 derives the mean of the X process (with an equivalent representation for the expected mean for process Y). According to the algebra of expectation, the expectation of a constant is the constant; therefore, $E(L_x) = L_x$. The random effects are deviations from the population mean; therefore $E(L_{xi}) = 0$. If we have a variable multiplied by a constant, the constant can be pulled out of the expectation; therefore, $E(S_x t_i) = S_x E(t_i) = S_x \bar{t}$. To compute the expectation of $S_{xi}t_i$, we make an assumption that an individual's slope deviation, S_{xi}, is independent of time t_i. This assumption is not required, but it makes the following derivation more straightforward. With this assumption, $E(S_{xi}t_i) = E(S_{xi})E(t_i)$. Because S_{xi} are deviations from a mean, $E(S_{xi}) = 0$, and therefore $E(S_{xi}t_i) = 0$. Last, because the error distribution is assumed to have a mean of zero, $E(e_{xi}) = 0$, leaving

$$E(L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi}) = L_x + S_x \bar{t}. \tag{A4}$$

Substituting the expectation for the mean of X and Y processes into Equation A3 gives,

$$\begin{aligned} Cov(X,Y) = E([(L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi}) - L_x - S_x \bar{t}\,] \\ [(L_y + L_{yi} + S_y t_i + S_{yi} t_i + e_{yi}) - L_y - S_y \bar{t}\,]). \end{aligned} \tag{A5}$$

The fixed effects for level cancel giving Equation A6,

$$\begin{aligned} Cov(X,Y) = E([L_{xi} + S_x t_i + S_{xi} t_i + e_{xi} - S_x \bar{t}\,] \\ [L_{yi} + S_y t_i + S_{yi} t_i + e_{yi} - S_y \bar{t}\,]). \end{aligned} \tag{A6}$$

Expanding Equation A6 produces Equation A7, a general formula for a cross-sectional covariance between two processes based on a linear model of change. However, in Equation A7 we have dropped the error terms. Because errors are always assumed to be uncorrelated with other factors, any expectations involving an error term will factor into a product of expectations, one of which is $E(e) = 0$. Furthermore, we can drop ($L_{xi}S_y\bar{t}, L_{yi}S_x\bar{t}$) because they are the expectation of a constant multiplied by a variable with an expectation of zero (L_{xi}, L_{yi}),

$$\begin{aligned} Cov(X,Y) = E(&L_{xi}L_{yi} + L_{xi}S_y t_i + L_{xi}S_{yi}t_i \\ &+ S_x t_i L_{yi} + S_x t_i S_y t_i + S_x t_i S_{yi} t_i - S_x t_i S_y \bar{t} \\ &+ S_{xi} t_i L_{yi} + S_{xi} t_i S_y t_i + S_{xi} t_i S_{yi} t_i - S_{xi} t_i S_y \bar{t} \\ &- S_x \bar{t} S_y t_i - S_x \bar{t} S_{yi} t_i + S_x \bar{t} S_y \bar{t}). \end{aligned} \tag{A7}$$

The covariance involving fixed effects only can be shown by dropping all terms involving random effects. With only fixed effects present, Equation A7 simplifies to Equation A8,

$$Cov(X,Y) = E(S_x t_i S_y t_i - S_x t_i S_y \bar{t} - S_x \bar{t} S_y t_i + S_x \bar{t} S_y \bar{t}). \tag{A8}$$

Distributing the expectation over the terms on the right-hand side of in Equation A8 and dropping terms that cancel leads to Equation A9,

$$Cov(X,Y) = S_x S_y Var(t_i). \tag{A9}$$

To obtain Equation A9, we recognize that $E(S_x S_y t_i^2) = S_x S_y E(t_i^2)$ and that $E(t_i^2) = Var(t_i) - \mu_{t_i}^2$.

Covariance Expectation in Age-Homogeneous Cross-Sectional Designs

Beginning with Equation A7, we derive the covariance for a narrow-age cohort sample. Because age (t) is the same for all members of a particular group, the fixed effects are constant and do not contribute to the covariance between processes, and thus can be omitted. Equation A10 shows the covariance within a sample of individuals of exactly the same age [VAR(t) = 0],

$$Cov(X,Y) = E(L_{xi}L_{yi} + L_{xi}S_{yi}t + S_{xi}tL_{yi} + S_{xi}tS_{yi}t). \tag{A10}$$

Note that t is no longer subscripted because it is assumed to be constant within age-group ($t_i = c$, where c is a constant age for a particular group). Also, in Equation A10, the processes of change in X and Y are not assumed to be independent; it is the general case for a single age group.

By manipulating Equation A2 to obtain $Cov(X,Y) = E(X,Y) - \mu_x\mu_y$, Equation A10 becomes

$$\begin{aligned} Cov(X,Y) = {} & Cov(L_{xi}L_{yi}) + \mu_{L_{xi}}\mu_{L_{yi}} \\ & + [t]Cov(L_{xi}S_{yi}) + \mu_{L_{xi}}\mu_{S_{yi}} \\ & + [t]Cov(S_{xi}L_{yi}) + \mu_{S_{xi}}\mu_{L_{yi}} \\ & + [t^2]Cov(S_{xi}S_{yi}) + \mu_{S_{xi}}\mu_{S_{yi}}. \end{aligned} \tag{A11}$$

Because the means of the random effects terms are zero, all the means on the right of Equation A11 drop out, leaving,

$$Cov(X,Y) = Cov(L_{xi}L_{yi}) + [t]Cov(L_{xi}S_{yi}) + [t]Cov(S_{xi}L_{yi}) + [t^2]Cov(S_{xi}S_{yi}). \tag{A12}$$

APPENDIX B

Effects of Age-Heterogeneity on Associations in Cross-Sectional Samples of Simulated Longitudinal Data

Each process, *X* and *Y*, is a simple linear function of time, as shown in Equation 1. Fixed and random intercepts and slopes were specified for each process, as well as correlations between the random components of the intercepts and slopes. The values of these functions were chosen with expected features of empirical data to form a broad range of possible relations between two linear processes. This permitted examination of a range of plausible values with which to evaluate differences between age-heterogeneous and age-homogeneous estimates of association. The data were generated by using a bivariate normal mixed model of the form:

$$x_{it} = L_x + L_{xi} + [t]S_x + [t]S_{xi} + e_{xi}$$
$$y_{it} = L_y + L_{yi} + [t]S_y + [t]S_{yi} + e_{yi}.$$

Parameters subscripted by *i* refer to the random effects. An individual's score at a given time is a function of the population average intercept or level (L_X, L_Y), random effects as deviation from population average intercept (L_{Xi}, L_{Yi}), population average rate of change or slope ($[t]S_X$, $[t]S_Y$), random effects as deviation from population average rate of change ($[t]S_{Xi}$, $[t]S_{Yi}$), and combined sources of systematic and random variance (e_X, e_Y).

In addition, data were generated to permit covariance between random effects of level and slope under multivariate normal distributions with expected population variance and covariance structure. All of the simulation factors were completely crossed forming a simulation study with 192 cells. As seen in Table B1, manipulated factors included the following: *mean slope for rate of change per time interval* (–5 and –8, with the same values specified for both coefficients within each condition, either –5, –5, or –8, –8), *individual variability in the slopes* (either variances of 5 for *X* and 4 for *Y*, or 10 for *X* and 8 for *Y*), *correlation between the intercepts* (i.e., amount of association at the beginning of each trajectory, .0, .3, or .6,), *correlation between the slopes* (i.e., associated between rates of change, –.3, 0 , .3, or .7), *correlation between intercept and slope* (i.e., how initial status is related to change, .0 or –.3) , and *random error* (variance of 1 or 10, where errors were drawn from a normal distribution with a mean of zero in both conditions). Although the correlation between the intercepts strongly influences the correlations observed between the *X* and *Y* processes, particularly at early time points, the correlation between slopes will quickly overwhelm any such initial association with the passage of time. Random error was included to reflect the combination of measurement error and nonsystematic within person variation commonly encountered in social science research.

TABLE B1
Simulation Factor Levels

Condition	*No. of Levels*	*Values of Levels*
Mean level	1	100
Variance level	1	100
Level correlation (Lc)	3	0.0, 0.3, 0.6
Mean slope (Sm)	2	–5, –8
Variance slope (Sv)	2	(5, 4), (10, 8)
Slope correlation (Sc)	4	–0.3, 0.0, 0.3, 0.7
Slope-level correlation (SLc)	2	–0.3, 0.0
Error (E)	2	1, 10
Time (age-homogeneous only)	11	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Data Generation

The simulation was run using a series of scripts written in the S-Plus programming language (Mathsoft, 1997). Each of the 192 cells of the simulation contained 500 replications of a dataset of 500 individuals with 100 generated values for time, ranging from 1 to 10 by increments of .1. Variances, covariances, and correlations were computed and stored as output. Analyses were performed on both correlation and covariance metrics to evaluate the degree to which increasing variance associated with linear models of this type (e.g., Mehta & West, 2000) influenced the results. Standardized bias of the estimates was calculated for each condition as the difference between the average estimate (of the correlation between *X* and *Y*) and the true parameter value, divided by the SD of the estimates, and multiplied by 100 to give the percentage standard deviation for a particular simulation condition.

To estimate the *X*-*Y* correlation from an age-heterogeneous, cross-sectional sample, a randomly drawn single "time-point" was chosen for each person and the correlation between the processes was calculated from that random sample of 500 observations. This process was repeated 10 times for each data set, producing 10 correlation estimates of the correlation between processes based on random, cross-sectional, age-heterogeneous samples. Because we generated longitudinal data, we could draw multiple cross-sectional samples from the data to make these multiple estimates. Therefore, each cell of the simulation contained 5000 estimates of the age-heterogeneous correlation between the processes. Eleven estimates of the *X*-*Y* correlation were also obtained for age-homogeneous, cross-sectional samples by computing the correlation coefficient at each time point from times 0 to 10. Age-homogeneous correlations reflect the combined effect of the correlation of the intercepts and the slopes of the processes.