

The Complex Nature of Unique and Shared Effects in Hierarchical Linear Regression: Implications for Developmental Psychology

Ulman Lindenberger

Max Planck Institute for Human Development

Ulrich Pötter

Johann-Wolfgang Goethe Universität

Hierarchical linear regression and related techniques, such as commonality analysis, path analysis, and linear structural equation models with mediator variables, are often used to determine the extent to which the influence of an exogenous variable on a dependent variable, *A*, is “unique” to this exogenous variable, or “shared with” another predictor variable, *B*. The authors formally show that shared and unique effects are related to the partial correlation between *A* and *B* controlling for the exogenous variable. We discuss the implications of this property of hierarchical linear regression with a special consideration of the role of chronological age in developmental psychology and warn against the uncritical use of hierarchical linear regression procedures.

A central goal in science is to identify connections among variables and to determine the directions of causality. The systematic manipulation of antecedent conditions in well-controlled experimental settings is a powerful tool to achieve this end. However, in the life sciences, many variables of fundamental theoretical importance resist direct experimental control. For instance, age, gender, and general intelligence cannot be randomly assigned to people. As a consequence, experimental methods need to be complemented by potentially less powerful research tools, such as quasi-experiments, experimental simulations, and correlational techniques (Baltes, Reese, & Nesselrode,

1988). Paradoxically, this necessary retreat to weaker (or at least more complex) methods does not weaken our desire to identify connections and determine causes. Often, the resulting mismatch between weak methods and strong causal intentions tends to produce an undue tendency to attribute more causal force to the results of statistical analyses than they actually deserve. A good example in this context concerns the confirmation bias associated with the interpretation of structural models (cf. Breckler, 1990; MacCallum, Wegener, Uchino, & Fabrigar, 1993). Here, a satisfactory model fit is often taken as strong support for the theoretically assumed causal structure, in part because competing models, which, for empirical or mathematical reasons, fit the data equally well, are not given sufficient attention.

Within the general linear model, the goal of this article is to highlight the risks of overinterpretation associated with hierarchical linear regression procedures. Hierarchical linear regression is often used to determine the extent to which the predictive link of the exogenous variable (i.e., the more distal predictor) to the dependent variable is reduced after statistical control of individual differences in a more proximal predictor, or mediator variable. For instance, researchers interested in relations among chronological age (exogenous variable), perceptual speed (mediator variable), and reasoning (dependent variable) may use hierarchical linear regression analysis to conclude that a certain percentage—say, 70%—of the age-related

Ulman Lindenberger, Max Planck Institute for Human Development, Berlin, Germany; Ulrich Pötter, Department of Economics, Johann-Wolfgang Goethe Universität, Frankfurt, Germany.

Ulman Lindenberger expresses gratitude to Paul B. Baltes for sharing his knowledge and views about multivariate methodology. In addition, we thank Will Fleeson, Chris Hertzog, Shu-Chen Li, Todd Little, John R. Nesselrode, and Tim Salthouse for helpful comments on earlier versions of this article, and we acknowledge discussions with Reinhold Kliegl, Michael Marsiske, and Ulrich Mayr.

Correspondence concerning this article should be addressed to Ulman Lindenberger, Max Planck Institute for Human Development, Center for Lifespan Psychology, Lentzeallee 94, D-14195 Berlin, Germany. Electronic mail may be sent to lindenberger@mpib-berlin.mpg.de.

variance in reasoning is “shared with” perceptual speed. Conversely, they may conclude that 30% of the variance in reasoning predicted by chronological age is “unique” to age. But what does such a result actually mean? Does it mean that age differences in reasoning are predicted to 70% by age differences in perceptual speed? Or does it mean that age differences in reasoning are predicted by individual (and not solely age-related) differences in perceptual speed? If only the latter is true, how can it be that a quantity that is age-related—age differences in reasoning—is shared with a quantity that encompasses all of the variance in perceptual speed related to reasoning, including its age-partial component?

The goal of this article is to answer these questions. Through formal analysis, we demonstrate that the magnitude of shared and unique effects in hierarchical linear regression depends on the partial correlation between the mediator variable and the dependent variable controlling for the exogenous variable. In deriving, illustrating, and discussing this finding, we focus on chronological age as an exogenous variable. Albeit admittedly arbitrary, this emphasis seems justified, because chronological age is an important exogenous variable that escapes direct experimental control (Baltes et al., 1988). Furthermore, given our substantive research interest in life-span developmental psychology (Baltes, Lindenberger, & Staudinger, 1997; Lindenberger & Baltes, 1995, 1997; Lindenberger et al., in press), we hope that our decision to discuss statistical issues in terms of age will help overcome the current overreliance of developmental psychology on cross-sectional data sets analyzed with hierarchical linear regression.

Overview

The remainder of this article is divided into five major sections. In the Method section we first provide a description of hierarchical linear regression. To this end, we distinguish three variables: the exogenous variable, age; the mediator variable, B ; and the dependent variable, A . Then we define the ratio of shared over simple effects (i.e., the proportion of age-related variance in A “shared with” B), termed *SOS*, as the critical parameter, introduce an analytic model to decompose *SOS* into age-related and age-orthogonal variance components (i.e., into variance components represented by variables that are either related or unrelated to age), and formally derive *SOS* as a function of the age relations of A and B as well as

the partial correlation between A and B controlling for age. In the subsequent section we illustrate the implications of our formal analysis with some numerical examples and demonstrate its generalizability to related statistical procedures such as path analysis, structural models with mediator variables, commonality analysis, and the quasi-partial correlation coefficient introduced by Salthouse (1994). After that we summarize and further discuss our findings and end with a plea for methodological pluralism.

Method

Problem Definition: The Hierarchical Linear Regression Procedure

In hierarchical linear regression, a series of linear regression analyses is performed to determine the extent to which a given predictor variable uniquely accounts for individual differences in the dependent variable. The Appendix illustrates the two-predictor case using the linear regression procedure of a widely distributed statistics package (i.e., SPSS-X, 1988). As before, we distinguish the exogenous variable, age; the mediator variable, B ; and the dependent variable, A .

In the first analysis, age alone is entered into the linear regression equation to determine its *simple* effect on A .¹ In a second analysis, B is entered alone to determine its simple effect on A . In a third analysis, both age and B are entered into the equation to determine the *total* effect of age and B on A . After that, the simple effect of B is subtracted from the total effect to determine the *unique* effect of age, that is, the component of the effect of age on A that is unrelated to Variable B . (Note that most statistical packages directly deliver unique effects when variables are entered in a stepwise fashion. Thus, the unique effect of age on A is identical to the R^2 increment observed when age is entered into the linear regression equation after B .)

The unique effect of age is then subtracted from the simple effect of age to yield the shared effect of age and B on A . Finally, *SOS* of age on A represent the proportion of age-related variance in A that is shared

¹ Throughout this article, the term *effect* is used to refer to variance components (e.g., the amount of variance in Variable A predicted by Variable B) as well as to expression derived on the basis of such components, such as differences between variance components and variance ratios.

with B . For instance, if age accounts for 12% of the variance in A (simple effect of age), B for 15% of the variance (simple effect of B), and age and B together for 20% (total effect of age and B), then the unique effect of age is 5% (i.e., $20 - 15$), and the effect of age shared with B is 7% (i.e., $12 - 5$). Thus, the computation of SOS of age on A indicates that 58% (i.e., $7/12 \times 100$) of the age-related variance in A is "shared with" (e.g., collinear with, predicted by, correlationally connected to) B .

An advantage of SOS is that they directly indicate the degree to which a given variable is, according to the logic of hierarchical linear regression, independently related to the dependent variable: The greater SOS, the smaller the independent contribution of a given variable to individual differences in the dependent variable. For instance, an SOS of 1.0 indicates that shared and simple effects are of equal magnitude, which means that all of the predictive variance of the critical variable is shared with other predictors and that the variable has no unique effects on A . In contrast, an SOS of 0.0 indicates that unique and simple effects are of equal magnitude, which means that none of the predictive variance of the critical variable is shared with other predictors. Finally, an SOS in the negative range indicates that unique effects are larger than simple effects and thereby points to the presence of suppressor effects (cf. Cohen & Cohen, 1983, pp. 90, 145; Pedhazur, 1982, pp. 208–211).

Statistical Background

We now express SOS as a function of the age relation of A , α , the age relation of B , β ; and the partial correlation between A and B controlling for age, $r_{AB.age}$ or *partial*. First, we derive the formula for the spuriousness model, in which A and B are exclusively related to each other through age (i.e., *partial* = 0). Afterward, we release this constraint as we move to the general model.

Our derivation of SOS as a function of α , β , and *partial* is informed by the work of Dawid on conditional independence (Dawid, 1979a, 1979b; cf. Cox & Wermuth, 1996; Dempster, 1972; for the general theory of linear regression, cf. Kendall & Stuart, 1973). Dawid (1979a) also introduced the notation \perp to refer to the orthogonality, or stochastic independence, of two variables.

In the following, variables are assumed to obey to multivariate normal law. Under such conditions, linear independence and orthogonality are equivalent. Note, however, that our conclusions do not depend on

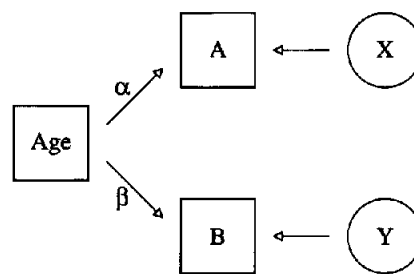


Figure 1. The spuriousness model.

multivariate normality but are also correct under the weaker assumption of linearity.²

Spuriousness Model

Consider three variables— A , B , and age—with a mean of 0 and a standard deviation of 1. Figure 1 displays the interrelations among the three variables according to the spuriousness model. The magnitude of the influence of age on A and B is expressed by two parameters, α and β , respectively. Because all three variables are normed, α and β can be interpreted as standardized linear regression coefficients, ranging from -1 to 1 . In addition, A and B both have specificities represented by two unobserved variables, X and Y , respectively. These specificities reflect error variance, specific variance, or any possible combination thereof.

The spuriousness model meets the following conditions of orthogonality, or independence:

$$\text{age} \perp X, \text{age} \perp Y, X \perp Y \quad (1)$$

Thus, it is assumed that: (a) age is independent of X , (b) age is independent of Y , and (c) X is independent of Y . Because independence is a symmetric property (cf. Dawid, 1979a), the reverse is also true, that is: (a') X is independent of age, (b') Y is independent of age, and (c') Y is independent of X . The name of the model is meant to highlight the fact that the relation between A and B is spurious in the sense that the two variables are related to each other through age alone.

Note that the three orthogonality conditions systematically differ in status: The two orthogonality conditions $\text{age} \perp X$ and $\text{age} \perp Y$ represent definitional

² The equations derived in the following two sections were also programmed using the AlgebraicRules procedure of Mathematica (Wolfram, 1994). A copy of the program is available from the authors (Pötter & Lindenberg, 1996).

properties of the model that cannot be falsified by data. In contrast, the orthogonality constraint $X \perp Y$ is falsifiable. For instance, in the linear case, $X \perp Y$ is violated if the partial correlation between A and B is different from 0 (i.e., if $partial \neq 0$).

The variances of A and B can be decomposed as follows:

$$\sigma_A^2 = \alpha^2 + \sigma_X^2 = \alpha^2 + 1 - \alpha^2 \quad (2)$$

$$\sigma_B^2 = \beta^2 + \sigma_Y^2 = \beta^2 + 1 - \beta^2 \quad (3)$$

The simple effect of age on A corresponds to the square of the standardized linear regression coefficient:

$$r_{age}^2 = \alpha^2. \quad (4)$$

To exclude the discussion of irrelevant special cases, we will assume that α is always unequal to 0.

The simple effect of B on A corresponds to the square of the correlation between A and B . Because $X \perp Y$, the general formula (see Equation 9) reduces to the square of the product of the path coefficients relating age to A and B :

$$r_B^2 = \alpha^2 \beta^2. \quad (5)$$

When looking at Figure 1, this expression makes intuitive sense because Variables A and B are linked to each other exclusively through age.

Next, we need to express the total effects of age and B on A . Because A and B are conditionally independent given age (e.g., A and B are related through age alone), we have:

$$r_{age, B}^2 = \alpha^2. \quad (6)$$

Equations 6 and 4 are identical, reflecting the fact that the total effects of age and B on A are identical to the simple effect of age if $X \perp Y$. The unique effect of age on B is obtained by subtracting Equation 5 from Equation 6:

$$r_{unique\ age}^2 = \alpha^2 - \alpha^2 \beta^2 = \alpha^2 (1 - \beta^2). \quad (7)$$

Finally, SOS of age on A under conditions of spuriousness, termed $SOS_{X \perp Y}$, result as the difference between simple and unique over simple effects:

$$SOS_{X \perp Y} = [\alpha^2 - (\alpha^2 (1 - \beta^2))]/\alpha^2 = \beta^2. \quad (8)$$

Another way to arrive at the same result proceeds from the observation that the correlation between A and B is identical with the shared effect of age and B on A if $X \perp Y$. Therefore, when $\alpha^2 \beta^2$ represents the shared effects of variables age and B on A , and when

α^2 represents the simple effect of age on A , then $\alpha^2 \beta^2 / \alpha^2 = \beta^2$ represents SOS.

Thus, we can conclude that SOS of age on A are equal to the square of the standardized linear regression coefficient of B on age, if the partial correlation between A and B controlling for age is 0. To provide an example, if $\alpha = .20$, $\beta = .90$, and $r_{A, B, age} = partial = 0$, the proportion of age-related variance in A shared with B equals .81. Likewise, given that the structure of our model is symmetric with respect to A and B (cf. Dawid, 1979a), the proportion of age-related variance in B shared with A equals .04. Therefore, 81% of the age-related variance in A is shared with B , and 4% of the age-related variance in B is shared with A .

The results obtained with the spuriousness model are encouraging in two ways. First, the square of a standardized linear regression coefficient is a parameter with well-known statistical properties. For instance, this parameter is known to vary between 0 and 1. Thus, SOS are easily interpretable whenever correlations among variables on which an exogenous variable (e.g., age) exerts its influence are reduced to 0 after controlling for the influence of this exogenous variable. In this context, it is important to note that $SOS_{X \perp Y}$ is entirely based on exogenous-variable variance (e.g., variance related to age); variance components of variables X and Y , which are defined to be orthogonal to the exogenous variable, are not implicated in this expression. This observation may seem tautological, but we will soon see that it is not.

General Model

In this section we derive the formula for SOS for the general case, that is, we release the orthogonality constraint $X \perp Y$. The resulting general model is displayed in Figure 2. The only difference between this model and spuriousness model consists in the introduction of a bidirectional path between X and Y and the corresponding coefficient $partial$ to represent the

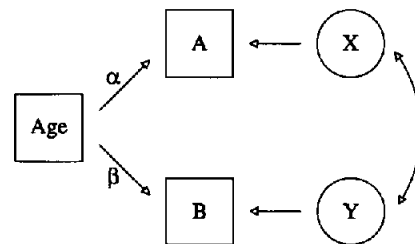


Figure 2. The general model.

partial correlation between A and B after controlling for age.

As before, the simple effect of age is α^2 . The simple effect of B is:

$$r_B^2 = (\alpha \beta + \text{cov}_{XY})^2, \quad (9)$$

where cov_{XY} is the covariance between X and Y . This covariance can also be written as a function of α , β , and the partial correlation between A and B controlling for age (i.e., $partial$):

$$\text{cov}_{XY} = \sqrt{(1 - \alpha^2)(1 - \beta^2)} partial. \quad (10)$$

Following standard results³ (cf. Rao, 1965, Chapter 4g), the total effect of age and B on A is:

$$r_{\text{age},B}^2 = \alpha^2 + \text{cov}_{XY}^2 / (1 - \beta^2). \quad (11)$$

The unique effect of age is obtained by subtracting Equation 9 from Equation 11. Finally, SOS of Age on A for the general case, termed $SOS_{\text{general}B}$ result as the difference between simple and unique over simple effects:

$$SOS_{\text{general}} = 1/\alpha^2 [\alpha^2\beta^2 + 2\alpha\beta \text{cov}_{XY} - (\beta^2/(1 - \beta^2)) \text{cov}_{XY}^2]. \quad (12)$$

Using Equation 10, we obtain:

$$SOS_{\text{general}} = \beta^2 + [2\beta \sqrt{(1 - \alpha^2)(1 - \beta^2)} partial] / \alpha + \beta^2 partial^2 - (\beta^2 partial^2) / \alpha^2, \quad (13)$$

which can also be expressed as follows:

$$SOS_{\text{general}} = 1/r_{\text{age}}^2 [r_B^2 - (1 - r_{\text{age}}^2) partial^2]. \quad (14)$$

As becomes apparent by inspection of Equation 13, SOS_{general} , the expression for the computation of SOS derived on the basis of the general model, consists of two parts. The first part, β^2 , is identical to the expression obtained for the spuriousness model. However, the second part of the expression is more complex and cannot be reduced to age-related variance components of A or B (i.e., α or β), because it contains a quadratic function of $partial$ (i.e., the partial correlation between A and B controlling for age). Thus, in contrast to $SOS_{X|Y}$, SOS_{general} is influenced by variance components represented by variables that are orthogonal to the exogenous variable. In the following section we illustrate the implications of this finding.

Illustration and Application to Related Statistical Procedures

Influence of Partial on SOS

The effect of $partial$ on the magnitude of SOS_{general} is substantial. This is seen in Figure 3, in which we

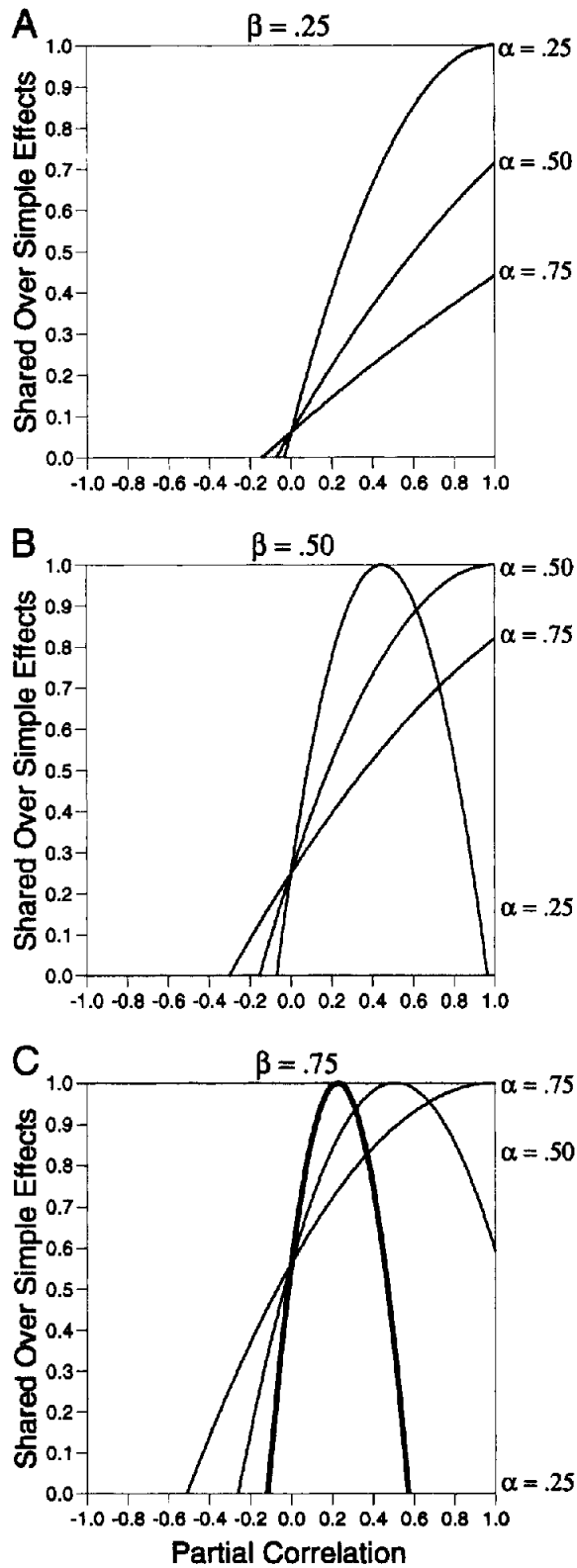
continuously varied $partial$ from -1 to 1 across nine different combinations of α and β . In each of the three panels, β was fixed at one of three levels (.25, .50, or .75), and α assumed values of either .25, .50, or .75. The nine resulting functions depict the magnitude of SOS_{general} as a function of $partial$.

As illustrated by Figure 3, the statistical properties of SOS_{general} as a function of α , β , and $partial$ can be summarized in five points: (a) SOS_{general} reduces to $SOS_{X|Y}$ when $partial$ is zero; (b) with increasingly negative values of $partial$, SOS_{general} quickly approaches 0 and eventually becomes negative; (c) when α and β are equal in magnitude, SOS_{general} increases with increasingly positive values of $partial$ and reaches unity when $partial$ equals 1; (d) when α is greater than β , SOS_{general} increases with increasingly positive values of $partial$ but does not reach unity; (e) when β is greater than α , then SOS_{general} rapidly approaches, and eventually reaches, unity with increasingly positive values of $partial$ and then decreases as a function of further increments in $partial$.

Connection to Related Statistical Procedures

To further illustrate the influence of $partial$ on both hierarchical linear regression in the narrow sense and on other statistical techniques that use the logic of hierarchical linear regression, we again consider the influence of $partial$ when α equals .25 and β equals .75 (see the thick line in Panel C of Figure 3). This condition is not uncommon, because it corresponds to an empirical situation in which we want to know whether a variable with a strong age relation (i.e., B) mediates most or all of the age-related variance of a variable with a weaker age relation (i.e., A). With $partial = 0$, we obtain $SOS = SOS_{X|Y} = \beta^2 = .56$ and conclude that 56% (i.e., $.75^2 \times 100$) of the age-related variance in A is shared with B . In the range in which $partial$ is larger than 0 but smaller than .23, small increments in $partial$ lead to drastic increments in SOS. SOS is equal to 1.0 when $partial$ equals .23. After having reached this maximum, further increments in $partial$ lead to decrements in SOS. SOS is again at .56 when $partial$ equals .46. Finally, with $partial = .57$, SOS is 0, and with even larger values of $partial$, SOS becomes negative.

³ Specifically, r^2 can be expressed as $1 - (1/\rho^{AA})$, where ρ^{AA} is the element of the inverse of the correlation matrix of A , B , and age pertaining to A (cf. Rao, 1965).



The four panels of Figure 4 further illustrate the dependency of SOS on *partial* and demonstrate how this dependency affects the outcome of hierarchical linear regression analyses and related statistical procedures. As for the thick line in Panel C of Figure 3, the age relations of *A* and *B* are held constant at $r = \alpha = .250$ and $r = \beta = .750$, respectively. *Partial* (i.e., the partial correlation between *A* and *B*, controlling for age) is varied at the four levels discussed before: Panel A, *partial* = .000; Panel B, *partial* = .228; Panel C, *partial* = .455; Panel D, *partial* = .572. As a consequence, the simple correlation between *A* and *B* also increases from left to right: Panel A, $r = .188$; Panel B, $r = .335$; Panel C, $r = .479$; Panel D, $r = .554$. Because the age relations of *A* and *B* remain constant over the four cases, the simple effects of age on *A* are always equal to the square of $\alpha = .250$, that is, $\text{simple}_{\text{age}} = .063$. In contrast, the simple effects of *B* on *A*, which are identical to the square of the correlation between *A* and *B*, vary as a function of *partial*.

Hierarchical linear regression and path analysis. Panel A of Figure 4 exemplifies a situation in accordance with the spuriousness model, because *partial* is 0. This fact expresses itself in two ways: (a) The unique effect of *B* on *A* is 0 (hierarchical linear regression proper) and (b) the path from *B* to *A* is 0 (path analysis). Note that the shared effect of *B* and age on *A* is *not* 0, because the two variables are related to each other through their respective age relations. In fact, on the basis of the spuriousness model, we already know that the magnitude of the shared effect is equal to the square of the multiplied age paths, that is, $(.250 \times .750)^2 = .035$. Accordingly, SOS of age on *A* is $.035/.063 = .563$, which, as we know from Equation 8, is equal to β^2 . (Throughout, slight numeric deviations are due to rounding error.)

In Panel B of Figure 4, *partial* is increased from .000 to .228—the point at which, according to Equations 12–14, unique effects of age on *A* are 0. As can be expected, both the unique effects of *B* and the

Figure 3. $\text{SOS}_{\text{general}}$ as a function of α , β , and *partial*. $\text{SOS}_{\text{general}}$ refers to the proportion of age-related variance in *A* shared with *B*. Within each of the three panels, β is fixed at a given level (Panel A: $\beta = .25$, Panel B: $\beta = .50$, Panel C: $\beta = .75$), α is varied across the same three levels, and *partial* varies continuously from -1 to 1 . The figure demonstrates the substantial influence of *partial* on the magnitude of $\text{SOS}_{\text{general}}$. The figures display only positive values of $\text{SOS}_{\text{general}}$.

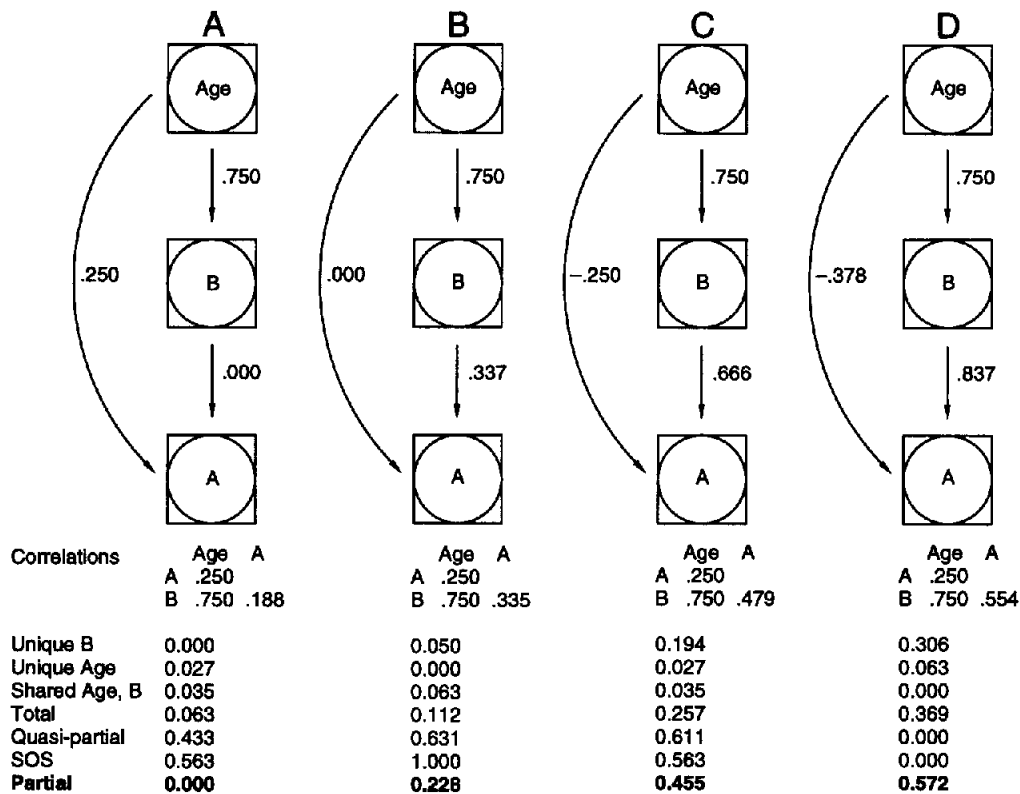


Figure 4. Illustration of the quadratic influence of the partial correlation between *A* and *B* controlling for age on hierarchical regression and related statistical procedures. Path coefficients were freely estimated using generalized least squares estimation. Correlations = first-order correlations among age, *B*, and *A*; Unique *B* = unique effect of *B* on *A*; Unique Age = unique effect of age on *A*; Total = effect of *B* and age on *A*; Quasi-partial = quasi-partial correlation coefficient (cf. Salthouse, 1994); SOS = shared over simple effects of age on *A*; Partial = partial correlation between *B* and *A* controlling for age.

shared effects of age and *B* are now larger than with *partial* = .000. In the corresponding path model, the absence of unique age effects expresses itself as a path coefficient of 0 for the direct path from age to *A*. The shared effect of age and *B* on *A* now corresponds to the square of the age → *B* and *B* → *A* paths, that is, $(.750 \times .337)^2 = .250$. Given that the unique effect of age on *A* is 0, shared and simple effects of age are equal in magnitude, and SOS is 1.000. If researchers using hierarchical linear regression or path analysis had collected data of this kind, they would conclude that all of the age-related variance in *A* is predicted by *B*.

After *partial* = .228, further increments in *partial* lead to decrements in SOS (see also the thick line in Panel C of Figure 3). This is illustrated in Panel C of Figure 4. With *partial* = .00 (Panel A), SOS is .563.

With *partial* = .455 (Panel C), SOS is again .563, which illustrates the nonlinear (i.e., quadratic) relation of SOS to *partial*. The same symmetry is reflected in the path diagram: The path from age to *A* is now equal in magnitude but opposite in sign to the path observed with *partial* = .000 (which, in turn, was equal to the simple correlation between age and *A*).

In Panel D, *partial* is increased to a value of .572. As a consequence, the shared effect of age and *B* on *A* is reduced to 0, which means that SOS is 0 as well. When we look at the path model, we notice that the two paths on *A* diverge even more than in Panel C. Both in Panels C and D, the hierarchical linear regression procedure reveals a suppressor constellation in the sense that *B* suppresses variance in *A* (negative suppression; cf. Conger, 1974).

Linear structural equation models with mediator

variables. In all panels of Figure 4, the three variables age, *B*, and *A* are depicted with squares and circles. This notation follows McArdle (1994) and is meant to indicate that the analysis of SOS presented in this article is equally valid for observed and latent variables. That is, to the extent that structural models based on latent variables embody the logic of hierarchical linear regression, they are subject to the same statistical laws as hierarchical linear regression procedures based on observed variables.

We agree with others that structural models based on latent variables often lead to more generalizable (e.g., unbiased) representations of psychological constructs and construct relations than methods based on observed variables (Hertzog, 1996; Meredith, 1993). In the present context, however, the latent character of construct representations in linear structural equation models is less crucial than another property of structural models: In contrast to a predominantly exploratory use of hierarchical linear regression (as in commonality analysis; see below), a structural model forces researchers to explicitly represent a theory-based expectation about the causal structure of the data.

Of course, there is nothing fundamentally wrong with estimating standardized direct and indirect effects on the basis of linear structural equation models with mediator variables of the kind represented in Figure 4. However, when doing so, it has to be kept in mind that all interpretations based on such models are *conditional on the truth of the mediation assumption*. Thus, the interpretability of the absolute and relative magnitude of direct and indirect effects in such a model rests entirely on the basic assumption of the model—namely, that the mediator variable does in fact act as a mediator of the causal effect of the exogenous variable on the dependent variable. The degree of statistical model fit does not help in this regard, because the model may fit the data perfectly, even after constraining direct effects of the exogenous variable on the dependent variable to 0, but the mediation assumption may still be wrong. To back up the plausibility of the mediation assumption, one has to provide converging evidence about the mechanisms by which the mediator is alleged in theory to have a causal effect on the dependent variable. Such converging evidence can be gathered in many different ways and includes experimental manipulation of the relevant mechanisms, more precise (e.g., mechanism-specific or process-pure) measurement, as well as the use of statistical methods and research designs that are

better suited to test assumptions about causality (see below).⁴

Commonality analysis. In contrast to linear structural equation models with mediator variables, commonality analysis is an exploratory data-analytic procedure. It consists of a coordinated series of hierarchical linear regression analyses to achieve a complete effect decomposition into unique and shared components (cf. Pedhazur, 1982). In the two-predictor case, this decomposition involves three effects: the unique effect of age, the shared effect of *B* and age, and the unique effect of *B*. So far, we did not consider the unique effect of *B*, because it was not needed to compute SOS of age on *A*. As is seen in Figure 4, the increase in *partial* leads to an increase in the unique effect of *B* on *A*. Figure 4 also illustrates that the three effects sum to yield the total proportion of explained variance in *A*, which also increases with increasing *partial*. Obviously, a larger partial correlation between *A* and *B* controlling for age implies a stronger association between these two variables and results in a larger proportion of explained variance.

In our view, the scientific value of commonality analysis depends on its use. It is clearly problematic to carry out a commonality analysis, inspect the magnitude of shared and unique variance components, and then design a structural model in which variables with relatively large unique effects occupy the mediator position and variables with relatively large shared effects occupy the exogenous-variable position. In doing so, researchers surrender themselves to the complexities of SOS, and the resulting structural model lacks a theoretical basis. On the more positive side, however, commonality analysis provides a good general overview of the degree of collinearity among a set of predictors and the magnitude of suppressor effects. If many of the shared variance components are negative, this implies that many of the unique effects are larger than the corresponding simple effects, which is always indicative of suppression. (Note, however, that suppression can already be present when unique effects are smaller than or equal to simple effects; see Panels C and D of Figure 4.) However, if all shared variance components are positive, and some of them are very large, and if unique effects tend to be small throughout, this suggests that the predictors act as a positive manifold in relation to the dependent vari-

⁴ We thank an anonymous reviewer for clarifying this issue.

able. Such an empirical situation is consistent with models that assume that most or all predictors are expressions of a common cause (for examples, see Lindenberg & Baltes, 1994, Table 6 and Footnote 7; Baltes & Lindenberg, 1997, Table 3).

Quasi-partial correlation coefficient. Recently, Salthouse (1994) introduced the quasi-partial correlation method to “evaluate the extent to which the age-related influences in a set of variables are independent of each other” (Salthouse, 1994, p. 413). In Figure 4 of his article and elsewhere, Salthouse described how to compute quasi-partial correlations by means of hierarchical linear regression. On the basis of these descriptions, the quasi-partial correlation coefficient is equal to the square root of the geometric mean of two SOS: the SOS of age on *A* (with *B* as the other predictor), on the one hand, and the SOS of age on *B* (with *A* as the other predictor), on the other. In a sense, then, the quasi-partial correlation coefficient is the symmetrical version of SOS transformed into a correlational metric by taking its square root.

Salthouse (1994) suggested that “the quasi-partial correlation can be thought of as the converse of a partial correlation because . . . the partial correlation procedure *excludes* age-related variance . . . whereas the quasi-partial correlation procedure is *restricted to* age-related variance” (Salthouse, 1994, p. 419). Given that the quasi-partial correlation coefficient is the square root of the geometric mean of two complementary SOS, it can be expressed as a function of α , β , and *partial* (see Equation 15 below).

The first part of Equation 15 is identical with Equation 13, whereas the second part is its mirror image, that is, all α s are replaced with β s and vice versa. Attempts to transform this expression into a more elegant form through expanding or factoring were unsuccessful. More important, however, this expression again contains *partial*, that is, it includes variance components in *A* and *B* that can be represented by variables that are orthogonal to age. As was true for SOS, this influence of age-orthogonal variance components disappears only when *partial* = 0. In that case, the quasi-partial correlation coefficient reduces to the square root of the geometric mean of the two age relations, α and β :

$$\text{quasi-partial}_{XLY} = (\alpha^2\beta^2)^{1/4}. \quad (16)$$

The influence of *partial* on *quasi-partial*_{general} is readily apparent by looking at the four panels of Figure 4, in which the quasi-partial correlation changes as a function of *partial* despite the fact that the age relations of the two variables remain unchanged. Thus, our arguments regarding SOS also apply to the quasi-partial correlation coefficient. Specifically, the quasi-partial correlation coefficient, in the general case, is not restricted to age-related variance.

Summary and Implications

The formal analysis of hierarchical linear regression presented in this article has revealed that unique and shared effects of the exogenous variable on the dependent variable are quadratically related to the partial correlation between the mediator variable and the dependent variable controlling for the exogenous variable. The examples provided in Figure 4 have demonstrated the ubiquity of this dependency for all methods that adhere, in one form or another, to the logic of hierarchical linear regression, such as path analysis, linear structural equation models with mediator variables, commonality analysis, and the quasi-partial correlation coefficient introduced by Salthouse (1994).

How is it possible that variance components representable by variables that are orthogonal to the exogenous variable, such as *X* and *Y* in Figure 2, influence a variance ratio deemed to decompose the variance of that exogenous variable into SOS? The answer to this question lies in the hierarchical linear regression procedure itself (see the Appendix). When we subtract the simple effect of the mediator variable from the total effect of both predictors to arrive at the unique effect of the exogenous variable, we are making use of all variance components in the mediator variable related to the dependent variable, including those components that can be represented by variables that are orthogonal to the exogenous variable. As a consequence, the magnitude of the unique effect of the exogenous variable is influenced by variance components in the mediator variable that are orthogonal to the exogenous variable if these variance components are related to the dependent variable. In other words, the magnitude of the unique effect of the exogenous variable is influenced by exogenous-variable extraneous variance.

$$\text{quasi-partial}_{\text{general}} = \left\{ \left[\beta^2 + \frac{2\beta\sqrt{(1-\alpha^2)(1-\beta^2)}\text{partial}}{\alpha} + \beta^2\text{partial}^2 - \frac{\beta^2\text{partial}^2}{\alpha^2} \right] \left[\alpha^2 + \frac{2\alpha\sqrt{(1-\beta^2)(1-\alpha^2)}\text{partial}}{\beta} + \alpha^2\text{partial}^2 - \frac{\alpha^2\text{partial}^2}{\beta^2} \right] \right\}^{1/4} \quad (15)$$

In our view, our formal exposition of the hierarchical linear regression procedure has two implications, one technical and the other substantive. We first discuss the technical implication. Although our discussion is couched in terms of variables *A*, *B*, and age, most points generalize to any field of psychology in which hierarchical linear regression has been used to decompose the influence of an exogenous variable (e.g., socioeconomic status, gender, grade level, etc.) on a dependent variable of interest.

Technical Implication

The technical consequence concerns the choice of words when reporting results obtained with hierarchical linear regression procedures. Whenever the partial correlation between *B* (mediator variable) and *A* (dependent variable) controlling for age (exogenous variable) differs from 0, statements such as “56% of the age-related variance in *A* is predicted by age differences in *B*,” are technically wrong because some of the relevant variance in *B* is orthogonal to age. The technically correct statement is “56% of the age-related variance in *A* is predicted by *individual* differences in *B*.” Paradoxical as it may sound (i.e., something age-orthogonal is used to predict something age-related), this statement correctly summarizes the hierarchical linear regression procedure. With respect to Panel A of Figure 4, both statements would be correct because the partial correlation between *A* and *B* controlling for age is 0. With respect to Panel C of Figure 4, however, only the second statement corresponds to statistical reality.

Substantive Implications

The substantive consequence of our analysis concerns the usefulness of variance-partitioning procedures as a statistical tool for obtaining knowledge about the causal influence of the exogenous variable on the dependent variable. As demonstrated, the capacity of a mediator variable to statistically account for exogenous-variable variance in the dependent variable is influenced not only by the relations of the mediator variable and the dependent variable to the exogenous variable but also by the partial correlation between the mediator and the dependent variable controlling for the exogenous variable.

A striking example of this dependency can be observed when one compares the path coefficients of Panels A, B, and C in Figure 4. The path coefficients of Panel B would be consistent with the hypothesis that *B* is a perfect mediator of age differences in *A*. In

contrast, Panels A and C would indicate that a substantial portion (i.e., 44%) of the age-related variance in *A* is *not* mediated by *B* but unique to age. Thus, depending on variations in the magnitude of the partial correlation between *A* and *B* controlling for age, researchers most likely will come to very different conclusions regarding the role of *B* as a mediator of age differences in *A*. If the data correspond to Panels A or C, they would probably conclude that *B* is a moderately powerful mediator of age differences in *A*. As a consequence, some researchers would invest efforts into the identification of additional variables that are able to mediate some of the remaining unique effect of age on *A*. If the data correspond to Panel B, however, researchers would conclude that Variable *B* is a perfect mediator of age differences in *A*. As a consequence, some may decide to closely examine the mechanisms underlying age differences in *B* with the hope of understanding age differences in *A*. In making this decision, they may not be aware that it was influenced not only by the age relations of *A* and *B* but also by the age-partialled correlation between these two variables.

In contrast to this practice, we recommend that the decision to entertain the hypothesis that a certain variable mediates the causal effect of another should be based on theoretical considerations and not on the outcome of hierarchical linear regression analyses. Of course, such theoretical considerations are informed by empirical findings, but the present analysis indicates that the outcome of hierarchical linear regression analyses should not serve as the primary or sole source of relevant empirical information. Specifically, it is important to realize that hierarchical linear regression does not offer a test of the basic mediation assumption. All it does is tell us how the world may look if that assumption were true.

The scientific legitimacy of hierarchical linear regression, then, critically depends on the theoretical status of the mediator variable *B* and the exogenous variable age. For instance, if theory posits that perceptual speed (mediator variable *B*) is a major cause of individual differences in reasoning (dependent variable *A*) in general (cf. Salthouse, 1996), then the quadratic influence of the partial correlation between *A* and *B* on the magnitude of SOS does not constitute a major problem of interpretation. In a way, the purpose of hierarchical linear regression analysis, in this case, is to test whether the exclusion of an exogenous variable from the linear regression equation, be it age, gender, socioeconomic status, or any other source of

individual differences, is associated with a major decrement in explained variance, with the expectation that this is not the case.

However, the use of hierarchical linear regression rests on weak ground if the mediator variable is assumed to be causal with respect to exogenous-variable-related variance alone, or if assumptions about causality are absent to begin with. For instance, if we assume that only age-associated individual differences in perceptual speed, but not individual differences in perceptual speed in general, predict reasoning, or if we have no strong a priori assumptions about the role of perceptual speed altogether, then the results of hierarchical linear regression are difficult to interpret in theoretical terms.

Conclusions and Outlook

Hierarchical linear regression is a statistical tool that reorganizes information contained in the covariance matrix (or, if standardized, the correlation matrix). As such, this method cannot escape the limitations of correlation research, and its results must not be overinterpreted. Even statistically sophisticated variants of hierarchical linear regression, such as linear structural equation models with mediator variables, do not provide more than a consistency check between the theory underlying the model and the empirical data (cf. Breckler, 1990; MacCallum et al., 1993). And, of course, all these methods assume that relations among variables are linear.

When developmentalists analyze cross-sectional data sets with hierarchical linear regression procedures, strong causal intentions collide with a relatively weak method. Quite often, the ultimate goal of the analysis is to identify basic components or mechanisms acting as pacemakers of age-associated change in more complex forms of behavior (e.g., "developables"; cf. Flavell, 1992). However, what hierarchical linear regression does is identify variables that excel as mediators in the limited statistical sense of the hierarchical linear regression procedure. Whether these variables represent, in fact, basic components that drive development in more complex forms of behavior cannot be decided on the basis of the hierarchical linear regression procedure itself.

What is needed, then, is the use of additional statistical tools and research designs, coupled with sound theorizing and precise measurement, to move toward the identification of the directionality and structure of causal connections (Baltes et al., 1988; Hertzog,

1996). Particular attention should be given to methods that are nonexperimental (i.e., that do not presuppose random assignment of participants to the exogenous variable) but nevertheless are better suited to test assumptions about causality than hierarchical linear regression. Admittedly, such methods are difficult to implement because they often require repeated measurements on the same participants and a high degree of statistical expertise. Nevertheless, we believe that a stronger reliance and further elaboration of these methods is worth the effort. A few developmentally oriented examples follow.

Chronological age has the rather unique property that individuals assume different levels on this variable in a perfectly predictable manner until they die. The typical cross-sectional research design does not take advantage of this property, but longitudinal designs do, and therefore they allow for the analysis of interindividual differences in intraindividual change (Baltes et al., 1988; Nesselroade, 1991; cf. Reinert, Baltes, & Schmidt, 1966). By observing the same individuals on multiple measures at different ages, it is possible to examine intercorrelations among change scores, relations between level and slope on multiple measures, and lead-lag relations among different variables (Hultsch, Hertzog, Dixon, & Small, in press; McArdle & Nesselroade, 1994; McArdle & Woodcock, 1997). These methods do not rely on the logic of hierarchical linear regression with age as the exogenous variable because age changes are observed directly and do not need to be inferred on the basis of age differences.

Another promising method is the analysis of dynamic patterns of interrelations among variables over time within subjects (Molenaar, 1994). Presumably, many causal connections can be identified more precisely by looking at covariance structures representing time-ordered multivariate patterns within one organism, rather than at covariance structures representing differences between organisms.

Other methods are of broader applicability because they do not depend on the special properties of the age variable. For instance, quasi-experiments can be used to study predictors of individual differences in learning in a multivariate framework (Rogers, Fisk, & Hertzog, 1994), to identify predictors of individual differences in asymptotic levels of performance (Baltes & Kliegl, 1992), to identify process dissociations between groups (Kliegl, Mayr, & Krampe, 1994), or to empirically simulate the effects of exogenous variables (Lindenberger & Baltes, 1995). Fi-

nally, mathematical simulations similar to the one presented in this article can be used to test the discriminatory power of theoretical models and data-analytic tools (Molenaar & van der Molen, 1994; Perfect, 1994) or to promote theorizing in a given domain of research (Li & Lindenberger, in press). The search for causes in (developmental) psychology will profit from an increased reliance on these methods and some skepticism where hierarchical linear regression is concerned.

References

- Baltes, P. B., & Kliegl, R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology, 28*, 121–125.
- Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the life span: A new window to the study of cognitive aging? *Psychology and Aging, 12*, 12–21.
- Baltes, P. B., Lindenberger, U., & Staudinger, U. (1997). Life-span theory in developmental psychology. In W. Damon (Ed.) & R. M. Lerner (Vol. Ed.), *Theoretical models of human development: Handbook of child psychology* (5th ed., pp. 1029–1143). New York: Wiley.
- Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1988). *Life-span developmental psychology: An introduction to research methods* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Breckler, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin, 107*, 260–273.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Conger, A. J. (1974). A revised definition for suppressor variables: A guide to their identification and interpretation. *Educational Psychological Measurement, 34*, 35–46.
- Cox, D. R., & Wermuth, N. (1996). *Multivariate dependencies: Models, analysis, and interpretation*. London: Chapman and Hall.
- Dawid, A. P. (1979a). Conditional independence in statistical theory. *Journal of the Royal Statistical Society B, 41*, 1–31.
- Dawid, A. P. (1979b). Some misleading arguments involving conditional independence. *Journal of the Royal Statistical Society B, 41*, 249–252.
- Dempster, A. P. (1972). Covariance selection. *Biometrics, 28*, 157–175.
- Flavell, J. H. (1992). Cognitive development: Past, present, and future. *Developmental Psychology, 28*, 998–1005.
- Hertzog, C. (1996). Research design in studies of aging and cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 24–37). New York: Academic Press.
- Hultsch, D. F., Hertzog, C., Dixon, R. A., & Small, B. J. (in press). *Memory change in the aged*. New York: Cambridge University Press.
- Kendall, M. G., & Stuart, A. (1973). *The advanced theory of statistics*. London: Griffin.
- Kliegl, R., Mayr, U., & Krampe, R. T. (1994). Time-accuracy functions for determining process and person differences: An application to cognitive aging. *Cognitive Psychology, 26*, 134–164.
- Li, S.-C., Lindenberger, U. (in press). Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and dedifferentiation of cognitive abilities in old age. In L.-G. Nilsson & H. Markowitsch (Eds.), *Cognitive neuroscience and memory*. Toronto: Hogrefe.
- 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. (1995). Testing-the-limits and experimental simulation: Two methods to explicate the role of learning in development. *Human Development, 38*, 349–360.
- 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., Gilberg, R., Little, T. L., Nuthmann, R., Pötter, U., & Baltes, P. B. (in press). Sample selectivity and generalizability of the results from the Berlin Aging Study. In P. B. Baltes & K. U. Mayer (Eds.), *The Berlin Aging Study: Aging from 70 to 100*. New York: Cambridge University Press.
- MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114*, 185–199.
- McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research, 29*, 409–454.
- McArdle, J. J., & Nesselroade, J. R. (1994). Using multivariate data to structure developmental change. In S. H. Cohen & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological contributions* (pp. 223–267). Hillsdale, NJ: Erlbaum.

- McArdle, J. J., & Woodcock, R. W. (1997). Expanding test-retest designs to include developmental time-lag components. *Psychological Methods*, 2, 403-435.
- Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, 58, 525-543.
- Molenaar, P. C. M. (1994). Dynamic latent variable models in developmental psychology. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 155-180). Thousand Oaks, CA: Sage.
- Molenaar, P. C. M., & van der Molen, M. W. (1994). On the discrimination between global and local trend hypotheses of life-span changes in processing speed. *Acta Psychologica*, 86, 273-293.
- Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change. Recent advances, unanswered questions, future directions* (pp. 92-105). Washington, DC: American Psychological Association.
- Pedhazur, E. J. (1982). *Multiple regression in behavioral research*. New York: Holt.
- Perfect, T. J. (1994). What can Brinley plots tell us about cognitive aging? *Journal of Gerontology: Psychological Sciences*, 49, P60-P64.
- Pötter, U., & Lindenberger, U. (1996). *The formal analysis of variance decomposition in hierarchical linear regression using the "AlgebraicRules" procedure of "Mathematica"*. (Technical Report) Berlin: Max Planck Institute for Human Development and Education.
- Rao, C. R. (1965). *Linear statistical influence and its applications*. New York: Wiley.
- Reinert, G., Baltes, P. B., & Schmidt, L. R. (1966). Kritik einer Kritik der Differenzierungshypothese der Intelligenz [Critique of a critique of the differentiation hypothesis of intelligence]. *Zeitschrift für experimentelle und angewandte Psychologie* 13, 602-610.
- Rogers, W. A., Fisk, A. D., & Hertzog, C. (1994). Do ability-performance relationships differentiate age and practice effects in visual search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 710-738.
- Salthouse, T. A. (1994). How many causes are there of aging-related decrements in cognitive functioning? *Developmental Review*, 14, 413-437.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403-428.
- SPSS-X. (1988). *User's guide*. Chicago, IL: SPSS Inc.
- Wolfram, S. (1994). *Mathematica: Ein System für Mathematik auf dem Computer* [Mathematica: A system for doing mathematics by computer]. Bonn, Germany: Addison-Wesley.

Appendix

The Linear Hierarchical Regression Procedure

Regression variables: *A*, *B*, age

| | |
|----------------------|--------------------------------------|
| /Dependent <i>A</i> | |
| /Enter age | ————> Simple _{age} |
| /Dependent <i>A</i> | |
| /Enter <i>B</i> | ————> Simple _{<i>B</i>} |
| /Dependent <i>A</i> | |
| /Enter age, <i>B</i> | ————> Total _{age, <i>B</i>} |

$$\begin{aligned} \text{Unique}_{\text{age}} &= \text{Total}_{\text{age}, B} - \text{Simple}_B \\ \text{Shared}_{\text{age}, B} &= \text{Simple}_{\text{age}} - \text{Unique}_{\text{age}} \\ \text{SOS}^a &= \text{Shared}_{\text{age}, B} / \text{Simple}_{\text{age}} \end{aligned}$$

Note. The top half of the illustration shows how the simple effect of age on *A*, the simple effect of *B* on *A*, and the total effect of age and *B* on *A* can be computed using a widely distributed statistics package, SPSS-X (1988). The bottom half of the illustration displays the computation of unique and shared variance components.

^a Shared over simple effects of age on *A*.

Received December 23, 1996

Revision received August 14, 1997

Accepted November 16, 1997 ■