# The Impact of Methodological Changes in Gerontology

K. Warner Schaie

*The Pennsylvania State University*

Paper Presented at a Symposium on

Generations of Gerontology: Aging, Continuity and Change

M. Seltzer, Chair

at the Annual Meeting of the

Gerontological Society of America

Boston, November 17, 1990

# The Impact of Methodological Changes on Gerontology[1]

## Introduction

My task in this presentation is to delineate how changes in research methodology have impacted our field over the three decades or so that I have been privileged to me a member of the gerontological research community. Although I wish that I could be more inclusive, I will confine my remarks to progress in research methods within the context of behavioral and social gerontology. I hope to point out that gerontologists in these areas have been in the forefront in both the development of novel research methodology and the early application of new methods particularly suited to the developmental sciences.

The methodological advances that I wish to focus on may be seen particularly clearly if one subscribes to a more or less dialectic view of scientific inquiry. This view specifies the dynamic interplay of empirically derived inductive models, that are then expanded or differentiated deductively, and applied to new data sets. Revised models are then once again generated in an inductive fashion (cf. Riegel, 1976).

I will begin with some brief remarks on how methodological innovation leads to different views of extant data bases and their subsequent impact in the revision of gerontological theorizing. As concrete illustrations of this process I then provide two rather different examples that illustrate the impact of methodological changes on our field. My first example involves the shift from a primary emphasis on cross-sectional data to more complex longitudinal inquiry. The second shows how the introduction of a new methodology, confirmatory factor analysis, influences work on adult development.

## Methodological Innovation in Gerontology

Important recent innovations in research methodology in Gerontology have been of three major types: advances in instrumentation, research design, and techniques of analysis. The first involves the level of instrumentation and sophistication of measurement devices or scales. For example, introduction of computer assisted tomography and direct measurements of blood flow led to the obsolescence of earlier indirect methods for assessing the integrity of cortical structures. Likewise, moving from an electro-encephalographic description of the resting brain to the analysis of evoked potentials, has lead to paradigmatic shift in conceptualizing the role of the cortex and its structures as they interact with an active environment.

Advances in instrumentation also involve a shift from mere categorical description of the presence or absence of a phenomenon to the development of scales that have ordinal or interval properties. An advance of this kind was the realization that psychological scales could be developed for subjective properties, for example by the method of comparative judgment, and the more modern techniques of multi-dimensional scaling (Cliff, 1982).

A second class of methodological innovations that has impacted gerontological research involves the specification of novel schemes for collecting data and evaluating the validity of theoretical constructs. For example, rules for determining the internal and external validity of experiments and quasi-experiments (Campbell & Stanley, 1963; Cook & Campbell, 1979) and the application of these concepts to gerontological studies (Schaie, 1977, 1978, 1988b) require researchers to define side conditions under which propositions derived from theory can be expected to hold. Methodological discussions that specify alternate models for the structuring of data collections in aging studies, moreover, will determine properties of any data set employed to test theory or provide bases for public policy formulations (cf. Riley, Johnson, & Foner, 1972; Schaie, 1965).

Innovations in methods of analysis, finally, make it possible to conduct new tests of theory (e.g., estimation of complex causal models a la Jöreskog [1979]), and lead to major paradigmatic shifts. An example of the latter impact is the shift from theories based upon directly observable variables to theories that account for relationships among latent (unobserved) variables. Likewise, introduction of methods of survival and event history analysis (cf. Allison, 1984) have made it possible to examine the impact of time-varying predictors on the occurrences of physiological, social, and behavioral morbidity (e.g., Schaie, 1979).

In each of the above instances, methodological advances permitted the testing of theoretical propositions that were previously not amenable to empirical test. In addition, constraints were identified that required expansion of theoretical formulations, and led to the derivation of new models that explain and include the attributes of a phenomenon uncovered by methodological innovations. Nevertheless, there is no uniform path by which methodological innovation and modification of research practice interact. To obtain a better understanding, two specific instances on how methodological advances have impacted gerontological research, will next be examined.

## Innovations in Research Design: The Case of the Age-Cohort-Period Problem

One of the prominent examples of the impact of research methodology in Gerontology has been the introduction of appropriate methodologies for the study of change over time, commonly referred to as the age-cohort-period problem (cf. Mason, Mason, & Winsborough, 1973; Palmore 1978; Ryder, 1965; Schaie, 1965, 1973, 1977, 1984, 1986, 1988). This case presents an illustration of the dialectic interaction between puzzling data sets, the examination of the appropriateness of standard research design, and the specification of alternate designs leading to the collection of new data sets that would fit the new paradigms.

My own work in this area began with the realization that data on the adult development of mental abilities showed wide discrepancies between cross-sectional and longitudinal data collected on the same subject population over a wide age range. In particular, it became evident that for some dependent variables substantial age differences obtained in cross-sectional studies could not be replicated in the longitudinal data while for other dependent variables, longitudinal age changes reflected more profound decrement than was shown in the comparable cross-sectional age difference patterns (Schaie & Strother, 1968).

In an inductive effort to explain these discrepancies I described a general model for the study of developmental change that explicated the relationships between the cross-sectional and longitudinal methods. This model showed that cross-sectional data involve the description of age differences at a single point in time; representing a separate samples design a la Campbell and Stanley (1963). Designs of this type suffer from the problem that maturational change (age) is confounded with cohort acting as a selection factor (Schaie, 1984). Longitudinal data, by contrast, represent a time series assessing the same individuals at two or more points in time. Here maturational change (age) is confounded with historical (secular) trends. The general model also specified a third approach to developmental data collections for which the term "time-lag," was coined (cf. Palmore 1978). The latter approach compares two samples at the same chronological age but at different calendar times, as would be the case, for example, in the comparison of SAT scores for successive classes of high school graduates. In this design cohort differences are confounded with historical trends. Consideration was then given to the possibility of deducing more complex designs, termed "sequential methods," that permit estimates of the magnitude of specific components of developmental change by controlling for the confounds mentioned above (Schaie, 1965).

The sequential designs were applied to empirical data sets obtained as part of the Seattle Longitudinal Study (Schaie, 1983, 1989). From these applications it became apparent that specific patterns of data acquisition lend themselves most readily to optimal utilization of the sequential analysis strategies (Schaie & Willis, 1991). It also became obvious that different sequential designs were appropriate for different developmental questions (Schaie, 1973), and that there was design complications needed to be specified to control for some of the validity threats listed by Campbell and Stanley (1963). Design variations were therefore explicated that permit controls for reactivity, practice and experimental mortality (Schaie, 1977, 1988).

Implications for Gerontological Research. At least three different implications of these designs have had important implications: Their use has influenced our understanding of the normative course of aging; second, they have impacted the structuring of data collections; and third they have affected the manner in which researchers explain phenomena involving change over time and age.

Many theories of aging, whether involving wear and tear, cumulations of waste products, or successive loss of neurons, implicitly or explicitly include the assumption of irreversible decrement. The design specifications resulting from the general developmental model, however, clearly indicate that in behavioral data, irreversible decrement represents only one of a number of observable patterns of aging. The plausibility of the irreversible decrement model, as any other model, can be tested only under certain conditions; a multi-cohort longitudinal study is needed to protect against the possibility that adverse historical effects could either inflate or suppress maturationally determined change. A given theoretical model must therefore indicate under what conditions the predicted age effects could validly be observed (cf. also Baltes, Cornelius, & Nesselroade, 1979).

In examining the contributions of aging research methodology it is important to distinguish between models that serve to help interpret the results of data acquisitions from models that explicitly specify how data should be acquired (cf. Schaie & Baltes, 1975). If alternative models of aging are to be tested, it becomes necessary to embed theory into data acquisition plans. Given the confounds described above, few theory-based questions are likely to be answerable by studying a single cross-sectional data set or even a single sample followed longitudinally over time. A logical consequence then is to conceptualize data collection plans such that extensions of the initial acquisition can be suitable extended. This involves data acquisitions that are structured as cross-sectional or longitudinal sequences (cf. Baltes, 1968; Schaie & Baltes, 1975).

One important derivative from the theoretical analysis of the general developmental model has been recommendations for an optimal data collection approach that permits flexible application of sequential data analysis strategies, as well as the provision of controls to protect against major threats to the internal validity of developmental studies. Noting the fact that all longitudinal studies must begin somewhere with a single first measurement occasion, I have long been convinced that it is always prudent to commence with an age-comparative cross-sectional design. However, in those instances where such a design cannot answer the questions of interest additional data must then be collected across time. A hypothetical data collection scheme which I have previously identified as the "most efficient design" is depicted in Figure 1 (Schaie, 1965; Schaie & Willis, 1991).

---------------------------

Insert Figure 1 about here

---------------------------

Recent analyses of the age-cohort-period problem have also provided a better understanding of what needs to be done, if the dependency of chronological age upon calendar time is to be broken in a meaningful manner. Models that contain parameters, one of which is wholly determined by the others are well known in science (e.g. the attributes of volume, pressure and temperature in physics). There are many reasons why one would want to examine the relative contributions of any of the three possible combinations of two developmental parameters. The recent literature on the analysis of sequential data matrices, moreover, would finesse the problem of invalid parametric assumptions by promoting regression models that estimate simultaneously the effects of age, period and cohort under an additivity assumption that allows for no interaction among the factors (e.g. Buss, 1979/80; George, Siegler, & Okun, 1981; Horn & McArdle, 1980; Mason, Mason, Winsborough, & Poole, 1973; Winsborough, Duncan, & Read, 1983).

The concern with methodologies designed to separate age, cohort and period effects, however, arose primarily from our preoccupation with the role of age as the independent variable of prime interest to students of development (Featherman, 1985; Slife, 1981). It would seem though that cohort and period may have more interesting explanatory properties than age. When cohort effects and historical time are conceptually separated from calendar time, cohort becomes a selection variable (Nesselroade, 1988) that characterizes the common point of entry for a group of individuals into a given environment, and period becomes a measure of event density (for further elaboration see Schaie, 1984, 1986). Similarly, chronological age can be reconceptualized as a

This work was made possible as a direct consequence of the methodological developments in restricted factor analysis most notably represented by the work of K. G. Jöreskog and his associates. Jöreskog (1971) and other methodologists such as Bentler (1980) demonstrated that hypotheses about the relationships among latent constructs that underlie empirically observed variables can be formulated as structural models. The original objective of confirmatory factor analysis was to permit the fitting of data to specific measurement models depicting the relationship between latent constructs and observed variables. It became soon apparent, however, that confirmatory factor analysis could also be used to test the equivalence of factor structure across and within groups (Jöreskog, 1979).

Implications for Gerontological Research. An important role of confirmatory factor analysis in theory-guided research is the applicability of structural equation models to the demonstration of measurement equivalence (Hertzog, 1986; Schaie & Hertzog, 1985). Structural equation models are of particular utility in aging studies because the uni-directionality of time permits sounder guides for the specification of causal paths then is possible in studies using single observation points only. Longitudinal factor analysis is particularly useful for the modeling of individual differences in intra-individual change, the central focus of any individual differences approach to aging (cf. Hertzog & Schaie, 1986; Jöreskog, 1979; Jöreskog & Sörbom, functional age dimension. In this approach, chronological age, or other age functions related to calendar time, rather than serving as explanatory concepts, emerge as useful scalars, that measure the amount of time elapsed within the life of individuals over which developmental phenomena have occurred. The dependencies implicit in the age-cohort-model, moreover, can thus be resolved, and all three effects can be directly estimated, whenever one of the dimensions is redefined in terms other than calendar time[2].

## Methodological Innovations that Inform Theory:
## The Case of Confirmatory Factor Analysis

Most direct observations conducted by social and behavioral scientists are of interest only to the extent that such observations represent reliably measured markers of latent (unobserved) psychological constructs. Theory-guided research requires the specification of the psychological constructs of interest, as well as the observations that will be used to estimate the constructs. Early concerns in the study of aging were directed primarily to the estimation of change in performance level on directly observable measures (cf. Dixon, Kramer, & Baltes, 1985). More recent work has extended these concerns to the comparison of structure (i.e., the regression of observables upon latent constructs) across different age groups and within cohorts over time (Cunningham & Birren, 1980; Hertzog & Schaie, 1986; Schaie & Hertzog, 1985, 1986; Schaie, Willis, Hertzog, & Schulenberg, 1987; Schaie, Willis, Jay, & Chipuer, 1989).

1977). Finally, methods of confirmatory factor analysis turn out to be the method of choice to test the stability of latent constructs under conditions of serendipitous or planned interventions in the aging process (cf. Schaie & Willis, 1986; Schaie, Willis, Hertzog, & Schulenberg, 1987).

Utilization of the same questionnaire or test apparatus does not guarantee measurement equivalence over time or different subject populations. Two fundamentally different aspects are at issue: The first is the traditional problem of reliability of measures across occasion and regression to the mean when using fallible measures (cf. Nesselroade, Stigler, & Baltes, 1980). The second aspect concerns the fact that measurement equivalence would not be guaranteed, even if only perfectly reliable measures were used because of systematic but nonuniform changes occurring among individuals over time.

In cross-sectional studies it is legitimate to ask whether a task that may be a good estimate of one construct in young adulthood remains so in late life or in fact becomes a measure of some other construct. What must be demonstrated then is the invariance of factor structure across multiple groups or sub-populations (cf. Alwin & Jackson, 1981). When samples are followed longitudinally, we are must then demonstrate factorial invariance for the same individuals measured longitudinally. Structural equation analysis consequently is the approach of choice to assess measurement equivalence issues involving multiple groups and occasions (cf. Rock, Werts, & Flaugher, 1978; Schaie & Hertzog, 1985).

Availability of explicit statistical models for testing construct equivalence has had profound effects for theory building and testing in aging research. For example, any discussion of a model that argues for the successive differentiation and dedifferentiation of human abilities (cf. Reinert, 1970), must not only specify the constructs that are involved in such a process, but also must specify hypotheses on the manner in which the constructs may be represented by different abilities (or the same abilities weighted differentially) at different life stages. The issue of equivalence of constructs will also impact future work in experimental gerontology. The experimental literature on age differences (cf. Salthouse, 1982) is largely an account of manipulations of single observable measures in the laboratory context. For such data to become meaningful in understanding age-related behavior in real life contexts complex structural models will need to be formulated and manipulations must be attempted at the latent variable level.

## Conclusions

In this presentation I have examined some ways in which developments in research methodology have impacted the behavior of the gerontological research community. Methodological innovations require major changes in testing theories and demand the addition of new corollaries and boundary conditions to retain the viability of existant theories. Other methodological innovations can directly produce paradigm shifts either by permitting the direct investigation of phenomena that were not previously specifiable by a theoretical model or by providing methods that allow investigation at different levels of conceptual specification. The impact of changing research methodology can be seen most clearly when investigators have been forced to abandon previous methods of measuring, designing or analyzing data in favor of conceptually superior innovations. I hope the illustrations I presented have made that point.

**References**

Allison, P. D. (1984). Event history analysis: Regression for longitudinal event data. Beverly Hills, CA: Sage.

Alwin, D. F., & Jackson, D. J. (1981). Applications of simultaneous factor analysis to issues of factorial invariance. In D. J. Jackson & E. F. Borgatta (Eds.), Factor analysis and measurement (pp. 249-278). London: Sage.

Baltes, P. B. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. Human Development, 11, 145-171.

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

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. Annual Review of Psychology, 31, 332-456.

Buss, A. R. (1979/1980). Methodological issues in life-span developmental psychology from a dialectical perspective. Journal of Aging and Human Development, 10, 121-163.

Campbell, D.T., & Stanley, J.C. (1963). Experimental and quasi-experimental designs for research in teaching. In N. L. Gage (Ed.), Handbook of research on teaching (pp. 171-246). Chicago: Rand McNally.

Cliff, N. (1982). What is and isn't measurement. In G. Keren (Ed.), Statistical and methodological issues in psychology and social science research. Hillsdale, NJ: Erlbaum.

Cook, T. C., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand-McNally.

Cunningham, W. R., & Birren, J. E. (1980). Age changes in the factor structure of intellectual abilities in adulthood and old age. Educational and Psychological Measurement, 40, 271-290.

Dixon, R. A., Kramer, D. A., & Baltes, P. B. (1985). Intelligence: A life-span developmental perspective. In B. B. Wolman (Ed.), Handbook of intelligence: Theories , measurements, and applications (pp. 301-350). New York: Wiley.

Featherman, D. L. (1985). Individual development and aging as a population process. In J. R. Nesselroade & A. von Eye (Eds.), Individual development and social change: Explanatory analysis. New York: Academic Press.

George, L. K., Siegler, I. C., & Okun, M. A. (1981). Separating age, cohort, and time of measurement: Analysis of variance and multiple regression. Experimental Aging Research, 7, 297-314.

Hertzog, C. (1986). On the utility of structural regression models for developmental research. In P. B. Baltes, D. Featherman, & R. M. Lerner (Eds.), Life-span development and behavior (Vol. 8). Hillsdale, NJ: Erlbaum.

Hertzog, C., & Schaie, K. W. (1986). Stability and change in adult intelligence: I. Analysis of longituinal covariance structures. Psychology and Aging, 1, 159-171.

Horn, J. L., & McArdle, J. J. (1980). Perspectives on mathematical-statistical model building (MASMOB) in research on aging. In L. F. Poon (Ed.), Aging in the 1980s (pp. 503-541). Washington: American Psychological Association.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. Psychometrika, 36, 409-426.

Jöreskog, K. G. (1979). Statistical estimation of structural models in longitudinal developmental investigations. In J. R. Nesselroade & P. B. Baltes (Eds.), Longitudinal research in the study of behavior and development (pp. 303-351). New York: Academic Press.

Jöreskog, K. G., & Sörbom, D. (1977). Statistical models and methods for analysis of longitudinal data. In D. J. Aigner & A. S. Goldberger (Eds.), Latent variables in socioeconomic models (pp. 285-325). Amsterdam: North Holland Publishers.

Mason, K. G., Mason, W. H., Winsborough, H. H., & Poole, W. K. (1973). Some methodological problems in cohort analyses of archival data. American Sociological Review, 38, 242-258.

Nesselroade, J. R. (1988). Sampling and generalizability: Adult development and aging research examined within the general methodological framework of selection. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. C. Rawlings (Eds.), Methodological issues in aging research (pp. 13-42). New York: Springer.

Nesselroade, J. R., Stigler, S. M., & Baltes, P. B. (1980). Regression towards the mean and the study of change. Psychological Bulletin, 88, 622-637.

Palmore, E. B. (1978). When can age, period, and cohort be separated? Social Forces, 57, 282-295.

Reinert, G. (1970). Comparative factor analytic studies of intelligence throughout the human life-span. In L. R. Goulet & P. B. Baltes (Eds.), Life-span developmental psychology: Research and theory (pp. 467-484). New York: Academic Press.

Riegel, K. F. (1976). Psychology of development and history. New York: Plenum.

Riley, M. W., Johnson, M. J., & Foner, A. (1972). Aging and society: Vol. 3: A sociology of age stratification. New York: Russell Sage.

Rock, D. A., Werts, C. E., & Flaugher, R. L. (1978). The use of analysis of covariance structures for comparing the psychometric properties of multiple variables across populations. Multivariate Behavioral Research, 13, 403-418.

Ryder, N. B. (1965). The cohort as a concept in the stdy of social change. American Review of Sociology, 30, 843-861.

Salthouse, T. A. (1982). Adult cognition: An experimental psychology of human aging. New York: Springer-Verlag.

Schaie, K. W. (1965). A general model for the study of developmental problems. Psychological Bulletin, 64, 91-107.

Schaie, K. W. (1973). Methodological problems in descriptive developmental research on adulthood and aging. In J. R. Nesselroade & H. W. Reese (Eds.), Life-span developmental psychology: Methodological issues (pp. 253-280). New York: Academic Press.

Schaie, K. W. (1977). Quasi-experimental designs in the psychology of aging. In J. E. Birren & K. W. Schaie (Eds.), Handbook of the psychology of aging (pp. 39-58). New York: Van Nostrand Reinhold.

Schaie, K. W. (1978). External validity in the assessment of intellectual development in adulthood. Journal of Gerontology, 33, 695-701.

Schaie, K. W. (1983). The Seattle Longitudinal Study: A twenty-one year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), Longitudinal studies of adult psychological development (pp. 64-135). New York: Guilford Press.

Schaie, K. W. (1984). Historical time and cohort effects. In K. A. McCloskey & H. W. Reese (Eds.), Life-span developmental psychology: Historical and generational effects (pp. 1-15). New York: Academic Press.

Schaie, K. W. (1986). Beyond calendar definitions of age, time and cohort: The general developmental model revisited. Developmental Review, 6, 252-277.

Schaie, K. W. (1988a). The impact of research methodology on theory-building in the developmental sciences. In J. E. Birren & V. L. Bengtson (Eds.), Emergent theories of aging (pp. 41-58). New York: Springer.

Schaie, K. W. (1988b). Internal validity threats in studies of adult cognitive development. In M. L. Howe & C. J. Brainerd (Eds.), Cognitive development in adulthood: Progress in cognitive development research (pp. 241-272). New York: Springer-Verlag.

Schaie, K. W. (1989). The hazards of cognitive aging. Gerontologist, 29, 484-493.

Schaie, K. W., & Baltes, P. B. (1975). On sequential strategies in developmental research: Description or explanation? Human Development, 18, 384-390.

Schaie, K. W., & Hertzog, C. (1985). Measurement in the psychology of adulthood and aging. In J. E. Birren & K. W. Schaie (Eds.), Handbook of the psychology of aging (2nd ed., pp. 61-92). New York: Van Nostrand Reinhold.

Schaie, K. W., & Hertzog, C. (1986). Toward a comprehensive model of adult intellectual development: Contributions of the Seattle Longitudinal Study. In R. J. Sternberg (Ed.), Advances in human intelligence (Vol. 3, pp. 79-118). New York: Academic Press.

Schaie, K. W., & Strother, C. R. (1968). The cross-sequential study of age changes in cognitive behavior. Psychological Bulletin, 70, 671-680.

Schaie, K. W., & Willis, S. L. (1986). Can intellectual decline in the elderly be reversed? Developmental Psychology, 22, 223-232.

Schaie, K. W., & Willis, S. L. (1991). Adult development and aging. 3rd ed. New York: HarperCollins.

Schaie, K. W., Willis, S. L., Hertzog, C., & Schulenberg, J. E. (1987). Effects of cognitive training upon primary mental ability structure. Psychology and Aging, 2, 233-242.

Schaie, K. W., Willis, S. L., Jay, G., & Chipuer, H. (1989). Structural invariance of cognitive abilities across the adult life span: A cross-sectional study. Developmental Psychology, 25, 652-662.

Slife, B. D. (1981). Psychology's reliance on linear time: A reformulation. Journal of Mind and Behavior, 2, 27-46.

Winsborough, H. H., Duncan, O. D., & Read, P. B. (1983). Cohort analysis in social research. New York: Academic Press.

## Footnotes

[1]This presentation is a condensation and update of a previously published chapter on "The Impact of Research Methodology on Theory Building in the Developmental Sciences," (Schaie, 1988a). The present version was prepared while the author was a fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, CA. Preparation of this work was supported in part by research grant R037 AG-08055 from the National Institute on Aging.

[2]Empirical examples of data sets that would allow direct estimation of age, cohort and period effects, given the proposed reformulation of the general developmental model are provided in Schaie (1986).

FIGURE 1

*Schaie's most efficient design. In 1960, four groups are tested (a cross-sectional study). They are retested in 1970 and 1980 (a longitudinal sequence with repeated measures). New groups from the same cohorts are first-tested in 1970 and in 1980 (cross-sectional sequences, independent samples). These new groups are later retested to form new longitudinal sequences.*