ACCESS

Aging Curriculum Content for Education in the Social-Behavioral Sciences

Iris Parham
Leonard Poon
Ilene Siegler
Editors

MODULE II

Methodological Constraints:

So You Want To Do Research on Aging

K. Warner Schaie

Springer Publishing Company, Inc.
536 Broadway
New York, NY 10012

90 91 92 93 94 / 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Aging curriculum content for education in the social sciences / Iris Parham, Leonard Poon, Ilene Siegler, editors.
p. cm.
Bibliography: p.
ISBN 0-8261-6070-0
1. Aging—Psychological aspects—Outlines, syllabi, etc. 2. Aged—Psychology—Outlines, syllabi, etc. I. Parham, Iris A. II. Poon, Leonard W., 1942- . III. Siegler, Ilene C.
BF724.8.A35 1989
155.67—dc19 89-11351
CIP

Printed in the United States of America

2
Methodological Constraints— So You Want To Do Research on Aging

K. Warner Schaie

INTRODUCTION

The purpose of this module is to present material on research methodology in the psychology of aging that would not ordinarily have been presented in the traditional courses in research design and statistics encountered in psychology at both undergraduate and graduate levels. Coverage of most of the materials presented here, however, assumes that such general courses have been taken. I am thus addressing graduate students, instructors, and researchers who have good methodology grounding, who have become interested in the study of aging phenomena, or who wish to extend their instructional repertory by developing aging research methodology components or courses. We will begin by focusing on some special research problems in the developmental sciences that researchers on aging need to consider, will consider some basic designs commonly used for descriptive quasi-experimental aging studies as well as experiments involving older participants, and will address particular implications and threats to the internal and external validity of aging studies. We will then examine sampling issues and discuss some major data analysis paradigms of interest to researchers on aging, and will end with some special concerns on the ethical reporting of research results.

Special Methodological Problems in Aging Research

Age and Time

Gerontological researchers usually employ calendar age as their primary independent variable. Frequently ignored, however, is the fact that while calendar age is a time-ordered process, time of onset or duration of a given behavior is not necessarily to be correlated with age. Executive burnout, for example, could be a function of length of time on the job or of the currency of professional knowledge. Since executive burnout would usually be measured in terms of time elapsed since initial training, this phenomenon may be spuriously correlated with calendar age in those career trajectories where initial training is customary at a specific age. The same phenomenon, however, would be uncorrelated with age when observed in a population where career entry occurs over a wide age range. Executive burnout therefore is time dependent but not necessarily age related (Schaie & Hertzog, 1985).

Models of Aging

A measurement complication in the developmental sciences is that data collected over time and age are not normally distributed but, instead, represent ordered sequences whose form may vary depending upon life stage or substantive content. The most common form hypothesized to fit age-related phenomena is the well-known Gompertz curve which assumes steeply decelerating increment until a young adult asymptote followed by slowly accelerating linear decline. The implied irreversible decrement model is unduly restrictive. It does not allow for the recursive or lagged phenomena that are so characteristic of development and aging. Other models that may fit frequently observed data sets include decrement with compensation and stability throughout adulthood (cf. Salthouse, 1987; Schaie, 1973, 1977; Schaie & Hertzog, 1985). The unidirectionality of time-ordered observations, however, does present a sound basis for causal inference, in that most consequents occurring later in time cannot be said to have "caused" antecedent behaviors and events. The unidimensionality of time-ordered behaviors, therefore, provides the major rationale for applying cross-lagged correlation and linear structural analysis to developmental data or for choosing many of the fixed parameters required for model estimation.

Experiments and Quasi-Experiments

It is not possible to assign study participants randomly to different ages, birth cohorts, or measurement occasions. Age-comparative and longitudinal studies are therefore in the nature of quasi-experiments (Campbell & Stanley, 1967; Schaie, 1973, 1977). True experiments in aging research are therefore typically either within-group or age-by-treatment

types of studies. Hypotheses tested by comparing nonrandom samples are subject to a variety of internal and external validity threats (to be discussed below).

Inter- and Intraindividual Differences

A major distinction among aging studies is between those that assess differences between age groups (interindividual differences) at a particular point in time and those that are concerned with changes within the same individuals across age (intraindividual differences). It should be noted that the major source of interindividual differences of interest to developmentalists are the patterns of individual differences in change within individuals. Results from studies involving comparisons of different age groups and those studying change over time within individuals are not interchangeable (cf. Labouvie, 1982; Nesselroade & Labouvie, 1985; Schaie, 1983).

Basic Design of Quasi-Experimental Aging Studies

Cross-Sectional Versus Longitudinal Methodology

Description of developmental status always involves the designation of an individual's age, the point of entry into the environment (cohort), and the point in time (period) when the dependent variable is measured. Two of these three components are ordinarily confounded, and the choice of data collection and analysis strategies should be directed by which components are of major concern (Schaie, 1965, 1977). It is widely held that the highest quality of aging data arises from longitudinal studies. However, there are many circumstances where a cross-sectional approach may indeed be the method of choice. On the other hand, if the question to be asked requires knowledge of the magnitude of intraindividual change, it cannot be answered without longitudinal data. Sometimes planned design mis-specifications may be useful when assumptions that might otherwise be unduly strong can be justified in the particular instance. For example, cohort effects are likely to be trivial when the range of ages compared is small, while period effects may be relatively trivial in many laboratory paradigms, except as expressions of reactivity or instrumentation effects (cf. Campbell & Stanley, 1967).

Data Collection Strategies

The major choice is, of course, whether one follows the same individuals over time or compares different-age individuals at the same point in time. A longitudinal sequence would follow two or more birth cohorts over a specified period of time. A cross-sectional sequence, on the other

hand, would assess samples of different ages at two or more points in time. A successive accretion of data suitable for a variety of developmental analysis schemes is provided by Schaie's "most efficient design" (Schaie & Willis, 1986a, pp. 28–34).

Data Analysis Strategies

Data on different age groups collected at one point in time are *cross-sectional* in nature. We do not know whether the older group was equivalent to the younger group at an earlier age; hence, group differences may be due to either age or cohort differences or both. Single cohort *longitudinal* studies follow individuals over a specific time interval. The observed change may be a period-specific effect and thus might not be age related. Longitudinal data therefore confound effects of age and history. Another analysis method is the *time-lag* approach commonly used in the study of educational performance. Here samples of the same age are measured at different points in time. In the case of a single time-lag comparison, the difference may be due either to specific period effects, or due to cohort differences occurring prior to the observational period. If two or more appropriately selected sets of observations are available, it is possible to estimate two of three components of development, while assuming the third to be trivial or of no interest. *Cohort-sequential* studies that cross age and cohort effects and assume period to be trivial may be most useful in circumstances where an irreversible decrement model applies. *Cross-sequential* studies cross age and period and are appropriate particularly in studies of midlife or for those variables where an adult stability model appears reasonable. *Time-sequential* studies cross age and period, and may be particularly relevant in those instances where an underlying aging-with-compensation model applies. The occurrence of significant interaction effects in each instance casts some doubt on the model assumption made (Schaie, 1965, 1977).

Experimental Aging Studies

Within-Age-Group Studies

Experimental within-group designs in aging research are frequently concerned with some type of intervention paradigm designed to remediate age-related deficits in performance (cf. Krauss, 1980). In such instances a young comparison group would often be inappropriate because it may not be possible to select younger participants who have suffered such decline. Random assignment to treatment and control groups will often be successful only if the within-sample age range is reasonably limited. Otherwise, because of increasing variability with age, pretest data are

needed to assure comparability of groups in small-sample studies. In any event, within-group designs should properly address the effects of intraindividual change. A powerful within-group approach involves intervention studies with groups of known developmental histories (e.g., Schaie & Willis, 1986b).

Age-Comparative Studies

The traditional paradigm in the experimental psychology of aging is to expose a young and an old group of subjects to a manipulation and then test the hypothesis that such a manipulation would be more effective for the old group and thus account for the observed age difference at base level. Main effects are of little interest in such studies since the treatment is selected because it is known to modify performance on the dependent variable, which is also known to differ by age; hence it is the interaction that is of primary interest (Baltes, Reese, & Nesselroade, 1977; Birren & Cunningham, 1985; Kausler, 1982). It is important, of course, to ascertain that the age difference in the dependent variable is not due to a failure to match on demographic or other base status variables. On the other hand, one should not inappropriately match away age differences. This could be avoided, for example, by matching relative to cohort averages. An alternative design would be to equate age groups for levels of prior practice, or to match young and old subjects on some complex behavior and then use Salthouse's (1987) molar equivalence–molecular decomposition approach to study the nature of possible underlying age effects and their compensatory mechanisms.

Many experimental aging studies suffer from inadequate cell sizes that result in inadequate power of significance tests. Minimum cell sizes should always be determined by appropriate power analyses (cf. Cohen, 1977).

Threats to Internal and External Validity

Internal Validity

Eight different threats to the internal validity of quasi-experiments have been described by Campbell and Stanley (1967) and their implications for aging research have been examined by Schaie (1977, 1988c). Of these, maturation is no threat in aging studies; it is typically the effect about which hypotheses are to be tested. However, the remaining seven threats represent rival hypotheses to the presence of maturational or age-specific effects, as follows: (1) *History* represents the impact of environmental events that could account for behavioral changes. For example, Schaie, Orshowsky, and Parham (1982) demonstrated different

temporal patterns of age changes for black and white elderly. (2) *Reactivity* involves the effect of practice or testing in changing the performance of individuals at a second or subsequent assessment, independent of any maturational change (Schaie, 1988c). (3) *Instrumentation* involves the use of different equipment or test protocols across time or for the different comparison groups. This threat also includes lack of structural invariance of measures across measurement occasions (cf. Schaie, Willis, Hertzog, & Schulenberg, 1987). (4) *Statistical regression* implies that because of measurement errors, means across time will move toward the population average, leading to spurious findings of lowered performance for above-average groups or raised performance for below-average groups (Furby, 1973; Nesselroade, Stigler, & Baltes, 1980; Rogosa, 1988). Examples of a time reversal technique used to differentiate between changes due to regression effects or actual age may be found in Baltes, Nesselroade, Schaie, and Labouvie (1972) and in Schaie and Willis (1986b). (5) *Experimental mortality* becomes a problem when not all subjects originally examined are available for retest. It has been shown that dropouts typically score lower at base test than do retest survivors (Cooney, Schaie, & Willis, 1988; Riegel & Riegel, 1972; Siegler & Botwinick, 1979). Methods for assessing and adjusting for experimental mortality may be found in Schaie (1977, 1988c). (6) Differential *selection* is a problem in cross-sectional and other nonequivalent control group designs (see Baltes, Cornelius, & Nesselroade, 1979; Labouvie, 1982; Nesselroade, 1988). (7) All of the above internal validity threats can occur in *interaction*. For example, dropout effects may differ by age; for young-old individuals they are greatest due to illness, while for the old-old they are greatest due to death (Cooney, Schaie, & Willis, 1988).

External Validity

The validity of aging studies may also be threatened by the selection of samples or procedures so specific that generalizability is severely restricted (cf. Baltes, Reese, & Nesselroade, 1977; Cook & Campbell, 1979; Nesselroade, 1988; Schaie, 1978). Five different threats to the external validity of aging studies have been identified: (1) *Measurement variables* may be selected that are differentially appropriate at different age levels or different historical periods. (2) *Treatment conditions* may be differentially supportive or aversive for the young or the elderly. (3) It may not be possible to generalize across *settings* unless laboratory paradigms are indeed representative of field settings (cf. Willis, 1987). (4) Studies of *subgroups of individuals* may not be generalizable to other older groups; for example, senior center visitors or nursing home residents may not well represent the older population. (5) Generalization may be difficult across *occasions of measurement* when a study is done under unusual conditions of secular change.

Sampling Issues in Aging Studies

Frequent design criticisms are concerned with the selection of age/cohort boundaries. While it is easy to dismiss out of hand research paradigms that compare young adult college students with senior citizen center visitors, it may be much more difficult to come up with positive prescriptions. Barring a better rationale, and without knowing at what age changes are to be expected, it would make sense to place the number of selected intervals at random ages. To elucidate age-related change free of cohort confounds, it would probably be best to choose narrow age ranges at that point where such change is observed in substantial numbers of persons. Sampling requirements may differ markedly across disciplines. Certain of the secondary data analyses proposed by economists, political scientists, and sociologists involve national probability samples. By contrast, very few psychological studies are concerned with the provision of national or regional population parameters. Representative samples for the psychologist therefore may mean no more than study populations that are broad enough to cover the full range of occurrence of the behavior of interest. But even within the more limited sampling objective there are different options. One may, for example, wish to sample the occurrence of behaviors roughly proportional to their incidence within a specified parent population. Alternatively, when the primary research interest is in the mechanism of a phenomenon rather than its population distribution, it may be quite legitimate to select subjects in terms of their suitability for the study of that mechanism (cf. Keith, 1988; Nesselroade, 1988; Schaie, 1973).

Health Status and Aging Research

One major sampling issue concerns the health status of study populations. Health status may either be of direct interest as an independent variable or simply be of concern as a control variable that informs us as to the possible generalizability of the results of the findings. Under the first rubric, many studies are concerned with identifying health status consequences of demographic factors or behavioral interventions. Likewise, other studies are concerned with health status as causal influences leading to behavioral events. In these instances the question arises, however, whether the critical health status indicators should refer to the subjects' physiological state, the subjects' subjective perception of their health, or the behavioral consequences of disease (possibly unrelated to specific diagnostic entities), such as those obtained, for example, by sickness impact measures (Elinson, 1988). It would seem reasonable to expect that subjects should routinely be characterized as to their health status in at least as gross a manner as would be done for other characteristics such as age, education, and occupational status.

Measurement Development

Selection of Age-Appropriate Materials

There has been considerable discussion of the need for ecological validity of assessment instruments for the elderly (cf. Schaie, 1978). However, it is probably wise to determine whether existing instrumentation can be adapted for work with the elderly before considerable effort is invested in new methodologies. Materials specifically designed for the elderly should generally be preceded by explorative studies into the world of older people, and should involve the use of a participant-observer approach (Keith, 1988) or the participation of older persons as sources of data and judges for scaling stimuli (e.g., Scheidt & Schaie, 1978).

Equivalence of Measures Across Age and Time

Substantial methodological advances during recent years have provided formal techniques that allow us to examine the equivalence of constructs across age and time, and the efficiency with which observed variables measure the theoretical constructs in different populations and at different ages. The relevant methods are discussed by Alwin (1988) and by Schaie and Hertzog (1985). An example of a measurement equivalence study in an older population may be found in Schaie, Willis, Hertzog, and Schulenberg (1987).

Major Analysis Techniques of Particular Interest to Aging Researchers

Analysis of Variance

This technique is widely used in aging research and assumes that age groups and/or cohorts are defined as discrete intervals. Such an assumption may be valid in age-comparative studies which typically use *factorial* designs (Kausler, 1982). In descriptive and longitudinal studies direct regression approaches may be preferred, unless discrete manipulations or controls for practice or experimental mortality are involved. It should be noted that two-point *repeated measurement* ANOVAs are the exact equivalent of factorial designs employing difference scores. A variety of ANOVA schemes for various aging studies are provided by Schaie (1977), with empirical examples given in Schaie (1988c). Analysis of covariance methods are often used to adjust for inability to provide experimental controls in aging studies. In addition to grave statistical problems in this procedure, such use is problematic also because it leads to estimates under conditions not prevalent in any circumstances to which findings might be generalized.

Multivariate Analysis of Variance

Whenever multiple measures are used with the same samples that are correlated, or whenever more than two measurement occasions are to be compared, MANOVA designs would be preferred. In such analyses it is then possible to estimate effects due to nonlinear components in change over time. When significant multivariate effects are found, univariate tests are appropriate to assess the magnitude of specific contrasts (cf. Schaie & Hertzog, 1985).

Multiple Regression Analyses

In many instances multiple regression analysis will be used to determine the proportion of variance in individual differences accounted for by effects isolated in the laboratory that are represented in behaviors in real life situations (cf. Willis, 1987). Other uses involve the determination to what extent age will account for differences in the dependent variable after the effects of other competing independent variables (e.g., education, occupation, health status) have been removed. Because of high colinearity between many measures in elderly populations it is generally inappropriate to use hierarchical regression procedures. The use of simultaneous (canonical) regression analysis or newer techniques such as ridge regression is to be preferred (cf. Rogosa, 1988). Another variant of regression analysis, path analysis, has long been used by sociologists to disaggregate variance components into those associated with direct, as against mediated, effects. Current practice would prefer the use of structural equation models in these instances (see Alwin, 1988; Schaie & Hertzog, 1985; and below).

Linear Structural Analysis

If our primary interest is in studying changes or differences on theoretical constructs, then methods of linear structural analysis will become of great importance in aging research. Factor analysis can be used to reduce the large number of observable variables and to explore the structure of variable domains. More powerful than exploratory analyses, however, are the methods of confirmatory factor analysis that are used to test hypotheses about the relation between observables and theoretical constructs. Confirmatory analyses are a powerful tool used to demonstrate measurement equivalence (construct validity) over time or different subject populations (see Cunningham, 1978; Schaie & Hertzog, 1985; Schaie, Willis, Hertzog, & Schulenberg, 1987).

Structural equation models are of particular utility in aging studies because the unidirectionality of time permits sounder guides for the specification of causal paths than is possible in studies using single ob-

servation points only. Longitudinal factor analysis is a particularly useful approach to the modeling of individual differences in intraindividual change, the central focus of any individual differences approach to aging (cf. Alwin, 1988; Campbell, 1988; Joreskog, 1979; Schaie & Hertzog, 1985).

Assessment of Individual Change

While structural equation models seem to be the appropriate avenue to the assessment of measurement equivalence and the modeling of structural changes, including mean structures, there remain some very practical reasons why many investigators may wish to rely on more simplistic and direct descriptors of observed change (e.g., Schaie, 1988e). The furor over problems in interpreting change scores in developmental research started by the influential Cronbach and Furby paper (see Furby, 1973) seems to have led quantitatively unsophisticated investigators to abandon the use of change scores and for some journal editors to resist publication of change score results. This is unfortunate, since more recent work has shown that the direct assessment of change is useful in many circumstances, and that difference scores are not necessarily less reliable than other approaches (for detailed discussions see Nesselroade, Stigler, & Baltes, 1980; Rogosa, 1988.)

Event History Analysis

The study of aging involves the occurrence of discrete events that make a difference (often irreversible) to the life of individuals. It may be rather important to know how such events are timed, and what the relative importance of possible predictors might be in determining the time of occurrence of developmental transitions or disfunctions. Epidemiologists and quality control experts have long used methods of accelerated failure analysis and proportional hazard analysis to deal with similar problems in their fields. Under the name of event history analysis these methods have recently become accessible to the social and behavioral sciences (Allison, 1984).

Ethical Reporting of Research Results in Aging Studies

The Problem of Ageism in Psychological Research

Researchers interested in aging tend to examine those variables that distinguish the young and the old; differences being typically in favor of younger subjects. Only rarely are variables examined where the older person may be at advantage because of experience. Research findings may therefore provide support for stereotypes regarding the elderly or become the scientific basis for policy decisions that may disadvantage

older persons (Schaie, 1988d). Researchers have an obligation, therefore, to be cautious in reporting results; specifically, they need to state the magnitude (and thus practical significance of their findings) in addition to the statistical significance (reliability) of observed differences. It should also be noted to what extent there is overlap between young and old comparison groups just as one would expect to find in studies of sex differences. Finally, it would seem appropriate to interpret the ecological relevance of small age differences (their practical significance) to make sure our findings are not misunderstood by the general public.

Collateral Data That Should Be Reported

The objectives just stated can best be achieved if research reports routinely report measures of variability (standard deviations and ranges) in addition to significance levels. In ANOVA designs it would further be appropriate to report measures of effect size, such as Omega Squared or Cohen's *d* (1977). In comparing performance levels of different age groups it would also be important to include measures of population overlap, such as the percent included in a joint distribution; and effect sizes in other relevant demographic categories (e.g., gender, education) would help to put the magnitude of age differences and changes in their proper perspective.

MODULE OUTLINE

I. Special research problems in the developmental sciences
  A. The study of age and time (Schaie & Hertzog, 1985)
    1. Time is unidirectional; antecedents and consequents must therefore be temporally distinct.
    2. Processes that have their onset and/or asymptote at a particular age are age dependent.
    3. Processes that vary in age of onset are time dependent.
  B. Models of aging (Schaie, 1973, 1988a)
    1. *Irreversible decrement* involves systematic change from a young-adult or middle-aged peak.
      (a) The Gompertz curve is a representative form of this model.
    2. *Decrement with compensation* involves age-related decline in a basic process that is compensated for by intervention or reorganization of complex behaviors.
      (a) Experienced typists maintain speed of typing even though reaction time goes up (Salthouse, 1987).
    3. *Stability* in the absence of pathology involves maintenance of optimal levels of functioning throughout adulthood.

C. Experiments and quasi-experiments: A preview (Schaie, 1977)
   1. Experimental subjects cannot be assigned to different ages or times of measurement.
   2. Age-comparative or longitudinal studies are therefore typically quasi-experiments.
   3. Experiments in aging research are typically within-group or age-by-treatment studies (Kausler, 1982; Krauss, 1980).

D. Inter- and intraindividual differences (Baltes, Reese, & Nesselroade, 1977; Schaie, 1983)
   1. Aging research differentiates between studies that assess differences between groups of individuals and those that study change within individuals.
   2. Studies of change within individuals involve both intraindividual changes and changes in interindividual differences over time.
   3. Studies of change must be longitudinal in nature.

E. Characteristics of observations (Schaie & Hertzog, 1985)
   1. Except for demographic indicators, observed variables serve as markers of theoretically interesting constructs.
   2. Measurement equivalence of inferred constructs is usually more stable over age and time than is true for directly observed variables.
   3. A choice must be made whether developmental changes or differences are to be measured as absolute values or as relative positions.

II. Basic design of quasi-experimental aging studies (Campbell & Stanley, 1967; Cook & Campbell, 1979; Nesselroade & Labouvie, 1985; Schaie, 1977)

A. Cross-sectional versus longitudinal methodology
   1. Designs of data collection must be distinguished from data analysis strategies (Schaie, 1983; Schaie & Hertzog, 1982).
   2. The age-cohort-period problem (Baltes, Cornelius, & Nesselroade, 1979; Schaie, 1965, 1986).
      (a) Definition: Development is indexed by an organism's age and cohort membership and the time-of-measurement (period) of a given response.
      (b) Similar to temperature, pressure, and volume in physics, the third factor in development is a known once the other two have been determined (two factors are always confounded).

B. Data collection strategies (Labouvie, 1982; Schaie, 1983)
   1. Longitudinal sequences follow several cohorts over time (and age) (also cf. Campbell, 1988).
   2. Cross-sectional sequences assess samples of different ages at two or more points in time.

   3. Schaie's "most efficient design" starts out with a cross-sectional sample and converts it into a series of short-term longitudinal studies (Schaie & Willis, 1986a).

C. Data analysis strategies (Schaie, 1973, 1977)
   1. Single-time cross-sectional studies confound age and cohort.
   2. Single-cohort longitudinal studies confound age and period.
   3. Time lag studies measure samples of the same age at different times; they confound cohort and period.
   4. Cohort-sequential studies assess two or more cohorts at two or more ages; period effects are assumed trivial.
   5. Time-sequential studies assess two or more age groups at two or more times; cohort effects are assumed to be trivial.
   6. Cross-sequential studies assess two or more cohorts at two or more times; age effects are assumed trivial.

III. Experimental aging studies (Baltes, Reese, & Nesselroade, 1977; Kausler, 1982)

A. Within-age group studies (Krauss, 1980; Schaie & Willis, 1986b)
   1. Subjects are randomly assigned to treatment and control conditions.
   2. Age ranges should be reasonably small.
   3. Otherwise, because of great variability, pretest data are needed to assure comparability of groups in small-sample studies.

B. Age-comparative studies (Kausler, 1982)
   1. Age groups need to be matched on demographic and relevant base status variables, without matching away aging effects.
      (a) Education or occupation might be matched relative to the cohort average on these variables.
   2. Groups with differential prior experience could be given practice to reach a common base criterion.
   3. Main effects are of little interest, since the treatment will be selected because it is effective and older adults are likely to perform at a lower level on average than do younger subjects.
   4. Age by treatment interaction is the effect of interest to determine whether treatment has differential effects by age.

IV. Threats to internal and external validity

A. Internal validity (Campbell & Stanley, 1967; Schaie, 1977, 1988b).
   1. History, general environmental impact across the time period during which aging is studied.
   2. Maturation, of specific interest to aging researchers, design must avoid controlling for this effect!

3. Reactivity (testing), effect of measurement operation on dependent variables being studied.
4. Instrumentation, variability of protocol or equipment across measurement occasions or comparison groups.
5. Statistical regression, effect of unreliability in spuriously raising means of low-scoring groups and lowering means of high-scoring groups over time (Baltes, Nesselroade, Schaie, & Labouvie, 1972; Furby, 1973; Nesselroade, Stigler, & Baltes, 1980; Rogosa, 1988).
6. Experimental mortality, the effect of selective sample attrition (Cooney, Schaie, & Willis, 1988; Riegel & Riegel, 1972; Siegler & Botwinick, 1980).
7. Selection, tapping particular ranges of behavior on independent variables affecting the dependent variable of interest (Nesselroade, 1988).
8. Interactions, combinations of the above threats.

B. External validity (Cook & Campbell, 1979; Schaie, 1978)
1. Measurement variables, the same instrument may not measure same construct in different samples.
2. Treatment conditions, laboratory paradigms may have differential effects on different samples.
3. Settings, effects obtained in the laboratory may not generalize to field settings.
4. Subgroups of individuals, the same phenomenon may have different temporal patterns in different population subgroups (Schaie, Orshowsky, & Parham, 1982).
5. Occasion of measurement, developmental phenomena may be specific to particular historical periods.

V. Sampling issues in aging studies (Schaie, 1973, 1988c)
A. Representative samples, may represent defined populations or sets of behaviors (Nesselroade, 1988).
B. Stratified samples, select from specific ranges of subject characteristics or phenomena.
1. Identification of demographic characteristics (e.g., age, sex, education, occupation, income).
2. Identification of health status (Elinson, 1988).
C. Oversampling of special populations, the study of abnormal or rare behaviors of interest.
D. Age boundaries, should be selected so as to minimize age/cohort confounds (Schaie, 1988c).

VI. Measurement development (Schaie & Hertzog, 1985)
A. Selection of age-appropriate materials (Schaie, 1978)
1. Participant-observer approach (Keith, 1988)
2. Ethnographic/psychometric approach (Scheidt & Schaie, 1978)

B. Equivalence of measures across age and time (Alwin, 1988; Schaie & Hertzog, 1985; Schaie, Willis, Hertzog, & Schulenberg, 1987)

VII. Major analysis techniques of particular interest to aging researchers
A. Analysis of variance (Schaie & Hertzog, 1985)
1. Repeated measurement designs
2. Factorial designs
3. Analysis of covariance
B. Multivariate analysis of variance
C. Multiple regression analyses
1. Hierarchical versus simultaneous analyses
2. Path analysis
D. Linear structural analysis (Alwin, 1988; Schaie & Hertzog, 1985)
1. Factor analysis (Cunningham, 1978)
(a) Exploratory
(b) Hypothesis testing
2. Causal modeling (Alwin, 1988; Campbell, 1988)
E. Study of individual change (Rogosa, 1988; Schaie, 1988b)
F. Event history analysis (Allison, 1984)

VIII. The ethical reporting of research results (Schaie, 1988a)
A. The problem
1. Effect size, magnitude of individual differences explained (Cohen, 1977).
2. Overlap across adjacent age groups, or between young and old comparison group.
3. Ecological relevance of age effect, does the observed age difference have a practical consequence?
B. Collateral data that should be reported
1. Measures of variability, standard deviations and ranges
2. Measures of effect size, Omega Squared or Cohen's *d* (Cohen, 1977)
3. Measures of population overlap, percent joint distribution.
4. Effect sizes in other relevant demographic categories (e.g., gender, education)

ANNOTATED BIBLIOGRAPHY

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

This monograph applies accelerated failure and proportional hazard models to the analysis of life events for which time-ordered data are

available. This technique has received recent prominence in sociological and epidemiological research. It appears to be a promising method for application to longitudinal behavioral data. Heuristic examples and descriptions of relevant computer programs are included.

Kind	*Recommendation*	*Readability*	*Audience*
HB	** 1/2	3	AG

Alwin, D. F. (1988). Structural equation models in research on human development and aging. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. W. Rawlings (Eds.), *Research methodology in studies of aging* (pp. 71–170). New York: Springer.

This chapter provides a tutorial describing the foundation of linear structural analysis and describing applications for the study of measurement equivalence and the testing of multivariate hypotheses for the study of developmental phenomena with several illustrations from the aging literature. Although highly technical, it is a more complete presentation than most other similar chapters, and should make the topic accessible to graduate students and researchers.

Kind	*Recommendation*	*Readability*	*Audience*
TC	** 1/2	3	AG

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. 61–88). New York: Academic Press.

Provides definitions and a thorough discussion of cohort effects in behavioral research.

Kind	*Recommendation*	*Readability*	*Audience*
TC	* 1/2	2	AG

Baltes, P. B., Nesselroade, J. R., Schaie, K. W., & Labouvie, E. W. (1972). On the dilemma of regression effects in examining ability-level related differentials in ontogenetic patterns of intelligence. *Developmental Psychology, 6*, 78–84.

Discusses the Cronbach–Furby arguments about the unreliability of change scores, and describes a time-reversal technique for assessing the seriousness of regression effects in limiting the interpretation of two-point change data. Provides a fully analyzed empirical example assessing change in cognitive function from the Seattle Longitudinal Study.

Kind	*Recommendation*	*Readability*	*Audience*
E	**	2	AG

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods.* Monterey, CA: Brooks-Cole.

A classic textbook (once again in print) describing methods for the study of developmental phenomena. While many technical issues are raised, the book approaches these problems so as to make them accessible for the beginner. Suitable for advanced undergraduates and service professionals as well as researchers at large. Brought up to date with some more technical supplementation [such as Schaie, Campbell, Meredith, & Rawlings (1988) (see Alwin, 1988)] it could also serve as text for a first-year graduate developmental research course.

Kind	*Recommendation*	*Readability*	*Audience*
TB	***	1	AU,SP

Birren, J. E., & Cunningham, W. (1985). Research on the psychology of aging: Principles, concepts and theory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 3–34). New York: Academic Press.

Discusses the paradigms most useful in behavioral research on aging and provides an outline of questions that must be addressed by the prospective researcher.

Kind	*Recommendation*	*Readability*	*Audience*
RC	**	2	AG

Campbell, D. T., & Stanley, J. C. (1967). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

This classic monograph defines the difference between experiments and descriptive research and specifies appropriate paradigms for the conduct of the latter variety of quasi-experimentation, most common in research on aging.

Kind	*Recommendation*	*Readability*	*Audience*
HB	***	2	AG

Campbell, R. T. (1988). Integrating conceptualization, design, and analyis in panel studies of the life course. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. W. Rawlings (Eds.), *Research methodology in studies of aging* (pp. 43–69). New York: Springer.

A thoughtful overview of the design of longitudinal panel studies, particularly appropriate for survey researchers, but providing many important recommendations for all researchers planning longitudinal studies.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	2	AG

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

An essential reference tool that needs to be known and kept on the shelf of every serious behavioral researcher. Contains the tables required to determine power of significance tests and to ascertain sample sizes required for valid hypothesis testing of relationships or group differences.

Kind	*Recommendation*	*Readability*	*Audience*
HB	*	3	AG

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

An expansion of Campbell and Stanley (1967) with emphasis on issues of external validity. Another important reference work for those employing descriptive research paradigms in field settings.

Kind	*Recommendation*	*Readability*	*Audience*
HB	* ½	2	AG

Cooney, T. M., Schaie, K. W., & Willis, S. L. (1988). The relationship between prior functioning on cognitive and personality variables and subject attrition in longitudinal research. *Journal of Gerontology, 43*, 12–17.

Provides an example of an empirical study of the effects of experimental mortality, segregating attrition effects into those associated with death, illness, or voluntary dropout. Data come from the Seattle Longitudinal Study.

Kind	*Recommendation*	*Readability*	*Audience*
E	**	3	AG

Costa, P. T., Jr., & McCrae, R. R. (1982). An approach to the attribution of aging, period, and cohort effects. *Psychological Bulletin, 92*, 238–250.

An intuitive approach to resolving the age-cohort-period problem. To understand this article, it would help to be familiar with Schaie (1965 or 1977).

Kind	*Recommendation*	*Readability*	*Audience*
E	* ½	3	AG

Cunningham, W. R. (1978). Principles for identifying structural differences: Some methodological issues related to comparative factor analysis. *Journal of Gerontology, 33*, 82–86.

A useful introduction to the design and interpretation of factor analysis studies in aging research.

Kind	*Recommendation*	*Readability*	*Audience*
RC	**	2	AG

Elinson, J. (1988). Defining and measuring health and illness. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. W. Rawlings (Eds.). *Research methodology in studies of aging* (pp. 231–248). New York: Springer.

Definitions are given for alternate ways to measure health status in behavioral and social science studies. Emphasis is on ways to measure quality of life as impacted by disease.

Kind	*Recommendation*	*Readability*	*Audience*
RC	* ½	1	AG

Furby, L. (1973). Interpreting regression toward the mean in developmental research. *Developmental Psychology, 8*, 172–179.

A classic article that led to the premature abandonment of change scores by many researchers. Provides background for material by Nesselroade et al. (1980) and Rogosa (1988).

Kind	*Recommendation*	*Readability*	*Audience*
TC	*	3	AG

Garfein, A. J., Schaie, K. W., & Willis, S. L. (1988). Microcomputer proficiency in later-middle-aged and older adults: Teaching old dogs new tricks. *Social Behaviour, 3*, 131–148.

Example of an empirical within-group aging study involving the teaching of a novel skill to adults and relating effects to prior status on ability measures.

Kind	*Recommendation*	*Readability*	*Audience*
E	* ½	2	AG

Hertzog, C. (1987). Applications of structural equation models in gerontological research. In K. W. Schaie (Ed.), *Annual review of gerontology and geriatrics* (Vol. 7, pp. 265–294).

An up-to-date review of applications of the new linear structural equation methods to topics in aging research.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

Horn, J. L., McArdle, J. J., & Mason, R. (1984). When invariance is not invariant: A practical scientist's look at the ethereal concept of factor invariance. *Southern Psychologist, 1*, 179–188.

Distinguishes between measurement invariance that involves complete equivalence of all psychometric characteristics (metric invariance) from a more relaxed criterion that requires only equivalence of structural patterns (configural invariance). Important for researchers planning comparative factor analyses studies.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	2	AG

Hultsch, D. F., & Hickey, T. (1978). External validity in the study of human development: Methodological considerations. *Human Development, 21*, 76–91.

A full disscussion of issues of generalizability and relevant paradigms as applied to aging research.

Kind	*Recommendation*	*Readability*	*Audience*
RC	* ½	2	AG

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.

This is the basic presentation of schemes for the application of methods of linear structural analysis to longitudinal data. Should probably not be attempted before working through Alwin (1988).

Kind	*Recommendation*	*Readability*	*Audience*
TC	* ½	3	AG

Kausler, D. H. (1982). *Experimental psychology and human aging*. New York: Wiley.

An excellent graduate level presentation of designs and research findings of age-comparative manipulative studies with older persons. Needs to be supplemented by additional material on descriptive research.

Kind	*Recommendation*	*Readability*	*Audience*
TB	**	2	AG

Keith, J. (1988). Participant observation: A modest little method whose presumption may amuse you. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. W. Rawlings (Eds.), *Research methodology in studies of aging* (pp. 211–230). New York: Springer.

A very accessible presentation of the use of ethno-methodology for nonanthropologists. Contains empirical examples of the application of participant observation techniques in a residence for older persons.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	1	AG

Krauss, I. (1980). Between- and within-group comparisons in aging research. In L. W. Poon (Ed.), *Aging in the 1980s* (pp. 542–551). Washington, DC: American Psychological Association.

Discusses under what circumstances studies of aging require young control groups and when within-age-group research is more appropriate.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	2	AG

Labouvie, E. W. (1982). Issues in life-span development. In B. B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 54–62). Englewood Cliffs, NJ: Prentice-Hall.

Another view of design issues in descriptive aging research.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

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

Questions of external validity and the selection of samples in aging research are considered within a general sampling framework that (perhaps for the first time) systematically includes the temporal dimension.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

**Nesselroade, J. R., & Labouvie, E. W. (1985). Experimental design in research on aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook*

***of the psychology of aging* (2nd ed., pp. 35–60). New York: Academic Press.**

Reviews issues of designing descriptive studies in aging research and extends and updates the material first presented in Schaie (1973 and 1977).

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	2	AG

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

This article reviews the Cronbach–Furby objection to the use of change scores in developmental research, and shows under what circumstances this objection may be valid or spurious.

Kind	*Recommendation*	*Readability*	*Audience*
TC	* ½	3	AG

Riegel, K. F., & Riegel, R. M. (1972). Development, drop, and death. *Developmental Psychology, 6*, 306-319.

A historically important treatment of the experimental mortality problem in behavioral research on aging. Data come primarily from the authors' German longitudinal studies.

Kind	*Recommendation*	*Readability*	*Audience*
TC	* ½	3	AG

Rogosa, D. (1988). Myths about longitudinal research. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. W. Rawlings (Eds.), *Research methodology in studies of aging* (pp. 171–209). New York: Springer.

A provocative review of traditional views of research methodology appropriate for longitudinal data. Rogosa supports the use of change scores, and emphasizes the utility of time series analyses.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

Salthouse, T. A. (1987). Age, experience and compensation. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 142–157). Norwood, NJ: Ablex.

A good introduction to the molar equivalence–molecular decomposition method of studying age differences. Describes the author's work on expert typists, a prime example of a data set fitting the decrement-with-compensation model.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

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

This is the original exposition of the age-cohort-period problem in developmental research in psychology. More accessible (less technical) presentations are provided in Schaie (1973, 1977) and Schaie and Willis (1986a).

Kind	*Recommendation*	*Readability*	*Audience*
TC	***	3	AG

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

Provides a less technical discussion of the age-cohort-period problem than the original (1965) exposition. Also examines issues of sampling methodology in aging studies.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	2	AG

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

Applies the work of Campbell and Stanley (1967) to aging research, discusses the age-cohort-period problem, and provides paradigms for the analysis of descriptive data sets, including controls for attrition and practice.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

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

Applies the consideration of external validity (cf. Cook & Campbell, 1979) to descriptive studies of intellectual aging.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	1	AG

Schaie, K. W. (1983). What can we learn from the longitudinal study of adult psychological development? In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 1–19). New York: Guilford Press.

A succinct summary of the rationale for conducting longitudinal studies in aging research. Also provides a convenient schematic for the different methods of collecting and analyzing descriptive developmental data.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	1	AU

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

Provides the conceptual framework for finessing the age-cohort-period problem by redefining one of the components of development on a noncalendar basis.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

Schaie, K. W. (1987). Research methods in gerontology. In G. Maddox & R. Corsini (Eds.), *Encyclopedia of aging* (pp. 570–573). New York: Springer.

A nontechnical review of research issues in gerontology designed for the general reader.

Kind	*Recommendation*	*Readability*	*Audience*
RC	**	1	AU,GP,SP

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: Psychological and social perspectives on time, self and society* (pp. 41–58). New York: Springer.

Reviews paradigm shifts resulting from the introduction of new schemes of data collection and data analysis. Examples cover the sequential methods and linear structural analyses applied to aging data.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	2	AG

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

A complete discussion of internal validity threats, their experimental control and analysis, and completely worked-out examples applying each method to data from the Seattle Longitudinal Study.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

Schaie, K. W. (1988c). Methodological issues in aging research: An introduction. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. W. Rawlings (Eds.), *Research methodology in studies of aging* (pp. 1–11). New York: Springer.

A nontechnical summary of major issues requiring attention by researchers on human aging.

Kind	*Recommendation*	*Readability*	*Audience*
RC	* ½	1	AG

Schaie, K. W. (1988d). Ageism in psychological research. *American Psychologist, 43*, 179–183.

Discusses the manner in which inappropriate inferences from aging research can lead to policy decisions that will unfavorably affect the lives of older persons. Includes recommendations for cautions to be observed in the reporting of research results.

Kind	*Recommendation*	*Readability*	*Audience*
RC	**	1	AU,SP

Schaie, K. W. (1988e). Individual differences in rate of cognitive change in adulthood. In V. L. Bengtson & K. W. Schaie (Eds.), *The course of later life: Research and reflections.* (pp. 65–85). New York: Springer.

An empirical example of methods for the study of individual differences in change over time in intellectual functioning.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

Schaie, K. W., & Hertzog, C. (1982). Longitudinal methods. In B. B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 91–115). Englewood Cliffs, NJ: Prentice-Hall.

A formal presentaion of considerations for the application of longitudinal methods, and an exposition of research designs suitable for the analysis of longitudinal data.

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

Schaie, K. W., & Hertzog, C. (1985). Measurement in the psychology of 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.

Provides a technical discussion of most measurement issues addressed in this module. The discussion of multivariate approaches to the study of aging is somewhat more accessible but not as complete as Alwin (1988).

Kind	*Recommendation*	*Readability*	*Audience*
TC	**	3	AG

Schaie, K. W., Orshowsky, S. J., & Parham, I. A. (1982). Measuring age and socio-cultural change: The case of race and life satisfaction. In R. C. Manuel (Ed.), *Minority aging: Sociological and social psychological issues* (pp. 223–230). Westport, CT: Greenwood Press.

An empirical application of sequential research methodologies to the study of minority aging.

Kind	*Recommendation*	*Readability*	*Audience*
E	**	2	AG

Schaie, K. W., & Willis, S. L. (1986a). *Adult development and aging* (2nd ed.). Boston: Little, Brown.

A comprehensive account of adult development from the stage of family formation to old age and death. The introductory chapter is of particular relevance for this module as it contains an accessible account of sequential-longitudinal methodology.

Kind	*Recommendation*	*Readability*	*Audience*
TB	**	1	AU

Schaie, K. W., & Willis, S. L. (1986b). Can decline in adult cognitive functioning be reversed? *Developmental Psychology, 22*, 223–232.

Empirical example of an intervention study with older persons, conducted within the context of a longitudinal study.

Kind	*Recommendation*	*Readability*	*Audience*
E	**	2	AG

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.

Empirical example of the application of comparative factor analysis to a pretest-posttest design cognitive intervention study.

Kind	*Recommendation*	*Readability*	*Audience*
E	*1/2	2	AG

Scheidt, R. J., & Schaie, K. W. (1978). A taxonomy of situations for the elderly population: Generating situational criteria. *Journal of Gerontology, 33*, 872–883.

Empirical example of the application of ethno-methodology and psychological scaling methods for the development of a technique for the study of perceived competence in the elderly.

Kind	*Recommendation*	*Readability*	*Audience*
E	**	3	AG

Siegler, I. C., & Botwinick, J. (1979). A long-term longitudinal study of intellectual ability of older adults: The matter of selective attrition. *Journal of Gerontology, 34*, 242–245.

Kind	*Recommendation*	*Readability*	*Audience*
E	**	3	AG

Empirical example of the application of methods for experimental mortality analysis to data from the Duke Longitudinal Study.

Willis, S. L. (1987). Cognitive training and everyday competence. In K. W. Schaie (Ed.), *Annual review of gerontology and geriatrics* (Vol. 7, pp. 159–182). New York: Springer.

Reviews the literature on intellectual change and cognitive intervention as related to everyday behavior.

Kind	*Recommendation*	*Readability*	*Audience*
RC	**	2	AG

Willis S. L., & Schaie, K. W. (1988). Gender differences in spatial ability in old age: Longitudinal and intervention findings. *Sex Roles, 18*, 189–203.

Empirical example of a method to disaggregate changes over time in speed and accuracy.

Kind	*Recommendation*	*Readability*	*Audience*
E	**	3	AG