Universität des Saarlandes and West Virginia University

Longitudinal and Cross-Sectional Sequences in the Study of Age and Generation Effects

P.B. Baltes

The following methodological considerations deal with that part of developmental psychology which is usually described as aging research. Aging research consists of observing samples of different age levels in order to obtain age-functional relationships. Since the concept of age is defined in this context by its chronological aspects, aging research may be seen as an interdisciplinary task. The research approaches of various disciplines such as biology, demography, medicine, and psychology differ only with respect to the measurement variables which are used. In the field of psychology, such age-comparative studies represent the bulk of developmental research. In fact, aging research is often thought of as being synonymous with developmental psychology. As Kessen (1960, p. 36) put it: "a characteristic is said to be developmental if it can be related to age in an orderly or lawful way".

To date two types of research plans have been used to study age-functional relationships, (a) the cross-sectional method and (b) the longitudinal method. Schaie (1965) called these two methods conventional designs for the study of aging. Despite the importance of these conventional designs for developmental psychology, it is surprising that nowhere in the psychological literature a satisfactory comprehensive discussion can be found. Only recently some more

---

1 This article is based partially on the writer's dissertation which was completed at the Universität des Saarlandes, Germany, in 1967. The writer, while bearing the sole responsibility for what is presented here, would like to express his gratitude to Margaret M. Baltes, E.E. Boesch, K. Eyferth, E. Labouvie, G. Reinhert, F. Sintpl and K. Wender for their constructive remarks, and to Th. Glass for his valuable assistance in preparing the present English version.
names, i.e. cross-sectional and longitudinal method. It is not clear, however, who first used these concepts in psychology. Gesell (1925, p. 441) employed the term cross-section in the context of developmental stages. The term longitudinal study appeared in the work of Blatz and Bott (1927). Yet in both these instances, the authors failed to designate these terms as specific methods of developmental research. In his classical contribution to the methodology of developmental psychology, Anderson (1931) pointed out that the terms cross-section and longitudinal had been introduced recently as technical terms, but without citing any references as to their origin. The operations which underlie both strategies had been described earlier in connection with other concepts. In the case of the longitudinal method, for example, such terms as consecutive tests, continuous record, follow-up, repeated measurement, repeated re-examination, repeated testing, repeated tests, retests, etc., had been used.

Research Designs

Based on the general terminology and classification of experimental designs (cf. Campbell and Stanley, 1963) both conventional strategies can be characterized as unisfactorial designs, with age as the single factor (Table I).

The critical difference between the cross-sectional and the longitudinal designs lies in the use of independent sampling \((S_1 \rightarrow S_n)\) in the cross-sectional method, and the use of dependent sampling \((S_1)\) in the longitudinal method. In both models, differences between the observations on the different samples \((S_1O_1 \rightarrow S_nO_1\) and \(S_1O_1 \rightarrow S_2O_n)\) are attributed to the various levels of the factor age \((A_1 \rightarrow A_n)\). Both designs are most powerful, if this interpretation is correct, i.e. if the differences between the age groups can be interpreted exclusively as age effects. In general, the significance of a plan with repeated measurement lies in its capacity to take into account intraindividual variation. As a result, a plan with dependent sampling (a) may be more sensitive, and (b) permits the analysis of individual trends. On the other hand, a plan using

---

* The earliest reference to different strategies in aging research known to the author comes from the field of medicine. Camerer (1910) described explicitly a “generalisierende” and an “individualisierende” method of studying age relationships. The operations basic to these methods are identical to the cross-sectional and the longitudinal technique respectively.
repeated observations also has some short-comings which are a consequence of the repeated measurement of the same individuals. The question of the choice between independent and dependent sampling will be dealt with in a subsequent section.

In the light of present standards of research methodology, both research designs appear to be relatively naive. Campbell and Stanley (1963) would not classify either model as a true experimental design. Their designation as pre-experimental designs implies that both models have serious methodological deficiencies which make questionable the interpretation of sample differences as pure age effects. The methodological deficiencies which are inherent to the conventional designs are possible a posteriori explanations for the divergent findings between cross-sectional and longitudinal studies. In this paper, the following problem areas will be discussed: (a) selective sampling, (b) selective survival, (c) selective drop-out, (d) testing effects, and (e) generation effects. The last two problem areas will receive closest attention since these sources of error can be controlled by the sequential models to be presented in a later section.

Most of these sources of error have as their basis the fact that the different age samples (S₁A₁ — SₙAₙ and S₁A₁ — S₁Aₙ) do not differ only with respect to age but also with respect to other factors as well. As a result, the differences between age groups can not be interpreted as pure age effects. In the sense of Campbell and Stanley (1963), these sources of error may impair either the internal or the external validity of both research designs.

Selective sampling. Because of the repeated participation which is required of subjects in a longitudinal study, longitudinal samples almost always fail to meet criteria of representative sampling. Conversely, cross-sectional samples usually can fulfill this requirement without difficulty. Some empirical studies (Baker, Sontag and
NELSON, 1958; KODLIN and THOMPSON, 1958; ROSE, 1965; STREIB, 1966) demonstrate that longitudinal samples already from their onset run the risk of being selectively biased in a positive direction. For example, subjects who volunteer for longitudinal studies tend to be of a higher average intelligence and tend to be of a higher socio-economic status. Such selective sampling biases in longitudinal studies impair (a) the comparability of cross-sectional and longitudinal investigations, and (b) the generalizability of longitudinal findings.

Selective survival. The phenomenon of selective survival applies to longitudinal as well as to cross-sectional samples. Selective survival implies that a given population at birth (cohort) changes in its composition in conjunction with the aging process as a result of death or incapacitation (BIRREN, 1959; CAMPBELL and STANLEY, 1963; DAMON, 1965; JARVIK and FALEK, 1963; RIEGEL, RIEGEL and MEYER, 1967a; 1967b). This population change is selective to the extent that the survival rate is correlated with the measurement variables. BIRREN (1959, p. 30) pointed out that "survivors might be, for example, taller or shorter, brighter or duller, happier or unhappier than their non-surviving cohort". This assumption of selective survival is further substantiated by the well-known fact that members of certain psychopathological groups have a shorter life expectancy. Subsequently, DAMON (1965) has demonstrated that alone on the basis of a negative correlation between height and life expectancy, the average height of older subjects tends to be less than that of younger subjects. RIEGEL, RIEGEL and MEYER (1967b) describe a gerontological study in which the survivors were on the average more intelligent, less dogmatic and less rigid than the non-survivors. In such cases of selective survival, the age samples no longer represent the population of birth cohorts and are therefore biased. As was suggested in the case of selective sampling, the effect of selective survival tends to be in the direction of positive selection.

Selective drop-out. Whereas theoretically the different age groups of a longitudinal study are completely homogeneous, in reality they become heterogeneous due to drop-outs. Such drop-outs, in the sense of CAMPBELL and STANLEY'S (1963) concept experimental mortality, occur during the course of the experiment as some subjects lose interest, change their residence, etc. This drop-out can be called selective if the loss of subjects does not follow a random
pattern, i.e. if there is a correlation between the dependent variable and the characteristics related to drop-out. Selective drop-out, as defined here, is a characteristic of the sample under investigation and not a characteristic of the underlying population from which it was drawn. This is in contrast to the case of selective survival where attrition can be attributed to changes in both the underlying population and the sample.

As a consequence of selective drop-out, the longitudinal sample as measured later in the experiment \((S_tA_nO_n)\) is no longer comparable to the original sample \((S_tA_1O_1)\). Again in the case of selective drop-out, it has been shown that the sample seems to become progressively more and more biased in a positive direction (AMES and WALKER, 1965; ANDERSON and COHEN, 1939).

**Testing effects.** The problem area of testing effects again applies primarily to the longitudinal method. The longitudinal designs, plans with repeated measurement, have as a basic assumption that the repeated observation of the same sample \((S_tO_t - S_tO_n)\) has no effects on the dependent variable. This assumption is highly improbable for many psychological variables when one considers short- and long-term practice and satiation effects. Such effects have been demonstrated with achievement tests (cf. ANASTASI, 1958) as well as with personality tests (cf. WINDLE, 1954). In connection with the longitudinal design, practice effects combined with increased test-sophistication and deliberate coaching have been cited as sources of error (ANASTASI, 1958; ANDERSON, 1954; KODLIN and THOMPSON, 1958; KUHLEN, 1963; MILES, 1934; OWENS, 1953; SONTAG, BAKER and NELSON, 1958; WELFORD, 1961; 1964). The importance of testing effects as confounding variables should not be underestimated since many longitudinal studies employ a high number of retests. For example, in the well-known Berkely Growth Study the majority of subjects were tested no less than 38 times over a period of 18 years (BAYLEY, 1949).

As a matter of fact, it is relatively easy to control testing effects in longitudinal studies by introducing adequate control groups. For the purpose of controlling testing effects, it is necessary only to draw two or more equivalent samples at the beginning of a longitudinal study and to vary them systematically in terms of number of observations made in the sense of a "posttest-only control group design" (BALTES, 1967a; CAMPBELL and STANLEY, 1963; SCHAEI, 1965). In a later section dealing with the simultaneous application
of longitudinal and cross-sectional sequences we will refer more explicitly to such a research plan.

*Generation effects.* The issue of generation effects as a source of error impairs the internal validity of cross-sectional studies and the external validity of longitudinal designs. With respect to the cross-sectional method it has been argued (Birren, 1959; Damon, 1965; Jerome, 1959; Kuhlen, 1963; Rosler, 1966; Schae, 1965; Welford, 1964) that the age samples \((S_1A_1 - S_nA_n)\) differ not only with regard to age, but also simultaneously as to generations in the sense of cohorts. First formulated by Kuhlen (1940), this issue has been stated more precisely by Anastasi (1958, p. 220): "Differences between 20- and 40-year-olds tested simultaneously (in 1940 or 1960) would reflect age changes plus cultural differentials, especially differences in the conditions under which the two age groups were reared". The relationships between age and generation effects in the cross-sectional design is presented in Figure 1.

![Diagram](image)

*Fig. 1.* A hypothetical example for the effects of generation differences on the results of a cross-sectional study.

In Figure 1 the broken lines represent the hypothetical average developmental gradients for the generations 1900–1950 for the age range from 10–60 years. These developmental gradients are based on two assumptions concerning the dependent variable: (a) age development as a linear progression, and (b) an acceleration from generation to generation. In a cross-sectional study conducted in 1960 there is available only one observation per cohort at a specific age level. The cohort of 1900 is observed at the age of 60, the cohort of 1910 at the age of 50, the cohort of 1920 at the age of 40,
etc. The resulting age curve (solid line) corresponds neither to any of the single developmental gradients nor to the average of all of them. Therefore, if generation effects are present, the results of a cross-sectional study can not be interpreted as pure age effects. In the obtained sample differences \( (S_1A_1O_1 - S_nA_nO_1) \), age effects and generation effects are confounded. This confounding might take various forms. The example in Figure 1 was chosen since it explains the well-known textbook age curves of intelligence as a function of the interaction between age and generation effects, whereas usually this curve is interpreted as a function of age alone.

The sample differences of a longitudinal study \( (S_1A_1O_1 - S_nA_nO_n) \) have been obtained from a single generation. Consequently, the age effect of a longitudinal study is generation specific. Therefore, the internal validity of a longitudinal design is not attenuated by the existence of generation differences. The external validity of age effects, however, found in longitudinal studies is seriously restricted. If, for example, the results of the Berkeley Growth Study, begun in 1928, are to be generalized to the birth cohort 1965, the generation gap is almost 40 years.

In summary, one must conclude that the conventional methods are in no way adequate research designs for the assessment of age effects. Due to the various methodological deficiencies it is not legitimate to interpret the sample differences obtained in longitudinal and cross-sectional studies as pure age effects. Besides the age effects a number of other uncontrolled sources of variance are confounded. These uncontrolled factors are alternative explanations for the obtained sample differences. The picture can become very complicated indeed, since the effects need not occur in the same direction and to the same degree in different measurement variables. Because of the hopelessly confounded sources of error it does not seem feasible to make \textit{a posteriori} statements with respect to differential validity of the results obtained by cross-sectional versus longitudinal investigations. If one considers the process of making at least one controlled comparison as the basis of securing scientific evidence (Campbell and Stanley, 1963), both conventional designs have such a total absence of control as to be of almost no scientific value. More differentiated research designs are needed to control these various deficiencies. In the following sections such research plans for the analysis and control of generation and testing effects will be presented.
General Developmental Model of Schaie

The literature discloses a number of proposals to refine the conventional designs in order to eradicate one or another of the methodological shortcomings (Anastasi, 1958; Bell, 1953; Birren, 1959; Kessen, 1960; Kuhlen, 1963; Miles, 1934; Rao, 1966; Schaie, 1959; ThomaE, 1959a; Welford, 1961). In most cases, these proposals are not explicitly worked out. They consist either of the suggestion to perform several cross-sectional or longitudinal studies or to combine both conventional designs with one another.

Schaie (1965) was the first to present a well explicated theoretical conception suitable to clarify the study of age effects. Schaie elaborated a general model for the study of developmental phenomena. This model provides a basis for the evaluation of certain characteristics of the conventional methods and for the formulation of new sequential research designs. Whereas the conventional methods are based on the unifactorial formula \( R = f(A) \), Schaie expanded this formula into a multifactorial one. He builds his general developmental model on three components: age \((A)\), cohort \((C)\), and time of measurement \((T)\). Applied to a concrete example of human development the definitions and the interrelationships between the three components are illustrated in Table II. The example refers to the cohorts of 1880–2000, with an assumed life expectancy of 80 years per cohort. The corresponding time of measurement extends from 1880–2080, in 20 year intervals.

Table II

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1880</td>
<td>0   20   40   60   80</td>
</tr>
<tr>
<td>1900</td>
<td>0   20   40   60   80</td>
</tr>
<tr>
<td>1920</td>
<td>0   20   40   60   80</td>
</tr>
<tr>
<td>1940</td>
<td>0   20   40   60   80</td>
</tr>
<tr>
<td>1960</td>
<td>0   20   40   60   80</td>
</tr>
<tr>
<td>1980</td>
<td>0   20   40   60   80</td>
</tr>
<tr>
<td>2000</td>
<td>0   20   40   60   80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1880</td>
</tr>
</tbody>
</table>
Concern with potential generation effects seemed to be the primary impetus for extending the unifactorial approach into the trifactorial approach \( R = f(A, C, T) \). Since there are a number of alternative cross-sectional and longitudinal studies for the corresponding age, cohort, and time of measurement conditions, single cross-sectional and longitudinal studies represent only special cases of the general developmental model. Each column in Table II (p. 154) represents a cross-sectional study, and each row (in the case of repeated measurement) a longitudinal study. Each diagonal represents a time-lag design, whereby subjects of the same age from different cohorts are compared. Although applied often (cf. Cattell, 1951; Emmett, 1950; Ramsey and Nelson, 1956; Scottish Council for Research in Education, 1949; Tuddenham, 1948; Wheeler, 1942; etc.), this third conventional method was first explicitly named by Schaie (1965). Schaie's developmental model illustrates again the consequences of generation differences. When generation differences exist, then (a) it is necessary to qualify the concept of age by specification of the generation membership, (b) the age effects of a cross-sectional study are confounded with generation effects, and (c) the age effects of a longitudinal study can not be generalized over other cohorts.

Two assumptions are inherent to the conceptualization and the analysis of the general developmental model proposed by Schaie (1965).

(a) The components: The three components age, cohort and time of measurement are conceptualized as separate entities which represent three different sources of developmental change in the sense of different antecedent conditions. In general, age differences represent maturation effects within the organism and time of measurement differences represent environmental effects. Schaie does not define the functional meaning of cohort differences unequivocally. Depending upon the specific case, cohort differences might signify hereditary effects and/or environmental effects.

(b) Separation of the components: Despite the formal interdependencies which exist between the three components age, cohort and time, it is possible to isolate their effects. An analysis of the interdependent (confounded) effects is obtained by algebraic substitutions and by alternate empirical checks of assumptions.

Schaie's proposed analysis of the effects of age, cohort, and time of measurement is based on the conventional cross-sectional
(CS), longitudinal (LO), and time-lag (TL) methods. Two of the three components are, however, always confounded in the sample differences (CSd, LOd, TLd) produced by the three conventional methods. According to Schaie, the following equations hold for the sample differences of the conventional methods:

\[
\begin{align*}
    \text{CSd} &= \text{Ad} + \text{Cd} \\
    \text{LOd} &= \text{Ad} + \text{Td} \\
    \text{TLd} &= \text{Td} + \text{Cd}
\end{align*}
\]

By alternate substitutions and subsequent solutions for the component effects, the effects of the three components age, cohort, and time can be derived in terms of the conventional methods:

\[
\begin{align*}
    \text{Ad} &= \frac{\text{CSd} - \text{TLd} + \text{LOd}}{2} \\
    \text{Cd} &= \frac{\text{TLd} - \text{LOd} + \text{CSd}}{2} \\
    \text{Td} &= \frac{\text{LOd} - \text{CSd} + \text{TLd}}{2}
\end{align*}
\]

The equations (4)–(6) are the basic operations for the formulation of Schaie's sequential strategies. The effects of the three components age, cohort, and time of measurement can be separated when independent estimates for the equations (4)–(6) are available. Such independent estimates of the three components can be obtained when the three conventional methods are applicable simultaneously for a specific subset of six samples of the general developmental model. By analysing the developmental model in this manner, Schaie formulated three sequential strategies: (a) the cohort-sequential method, (b) the time-sequential method, and (c) the cross-sequential method. It is expected that these sequential strategies unconfound the three sources of developmental change, i.e. will analyse age, cohort and time of measurement effects in their functional connotation as indicated above.

Reformulation of Schaie's Developmental Model

Bifactor model. The analysis of the developmental model, as proposed by Schaie, is not adequate for two kinds of reasons (Baltes, 1967a; 1967b): The first objection concerns the formal definitions of the three components, while the second deals with problems of their measurability.
(a) The conditions of Schaae's developmental model can be described adequately by only two components. The introduction of a third component is redundant and leads to an overinterpretation of the developmental model. The three components display the following formal relationships:

\[
\begin{align*}
A &= T - C \\
T &= A + C \\
C &= T - A
\end{align*}
\]  

(7) \hspace{1cm} (8) \hspace{1cm} (9)

These formal relationships result from the fact (Table II) that after two components have been defined the third is unequivocally fixed. The existence of these mutual dependencies signifies that the three components do not satisfy the qualifications of three true experimental variables, namely they can not be defined and varied independently.

With respect to the trifactorial approach of the developmental model the equations (7)–(9) indicate that one of the three components in Schaae's \( R = f(A, C, T) \) formula can always be replaced by substitution. For example, by substituting \( A + C \) from equation (8) for \( T \) in Schaae's basic trifactorial formula, the following equation results:

\[
R = f(A, C, A + C)
\]

(10)

From equation (10) it follows that all developmental phenomena inherent to the frame of reference of Schaae's developmental model can be described by only two components (parameters). In other words, all individuals considered by the developmental model can be classified adequately on the basis of only two characteristics.

[b) Schaae's suggestion to interpret the components content-wise, i.e. age differences are maturation effects, time differences are environment effects, etc., is inadequate. The variation of age, cohort, and time consists only of a classification of individuals into different segments of the time continuum. This classification is derived from two chronological criteria, time of birth and age. No other characteristics, such as hereditary or environmental conditions, are used. A variation of the time continuum, however, cannot isolate maturation from environmental effects. A functional interpretation of the components, as intended by Schaae, would require a direct measurement and subsequent variation of neurophysiological and environmental conditions. Granted the functional conditions, as referred to by Schaae, exist theoretically in
the corresponding sections of the time continuum, they can not be teased out by the arrangement of the developmental model. Additional research designs which extend far beyond the realm of the developmental model would be necessary to interpret the components in the sense of different functional sources of developmental change. The inability of the model to give the components different functional meaning is especially distinct in the case of the relationship between age and time of measurement, as formulated in Schaie's equation (2): \( \text{LOd} = \text{Ad} + \text{Td} \). In this instance, Schaie attempts to separate maturation and environmental effects in the aging process by applying his sequential strategies. It is clear, however, that such a separation is impossible by simply repeating observations on samples which differ only with respect to their placement on a time continuum.

In summary, it must then be concluded that (a) two components are sufficient to describe the developmental processes inherent to the general developmental model, (b) that the functional interpretations of the component effects, as suggested by Schaie, are not adequate, and therefore (c) the analyses of the developmental model, i.e. the sequential strategies sensu Schaie, are inappropriate. It is proposed herewith (Fig. 2) to conceptualize Schaie's developmental model as a bifactorial model with the components of age and cohort.

![Diagram](image)

Fig. 2. An example of the bifactorial developmental model with the components of age and cohort and its relationship to the conventional longitudinal (LO), cross-sectional (CS) and time-lag (TL) methods. Time of measurements are included for illustration only.

The meaning of both components corresponds only in its formal characteristics with the components of age and cohort, as formulated by Schaie. The components refer to the fact that
individuals can be assigned to specific levels of age and cohorts (generations).

*Age and generation effects.* In the case of a parametric analysis, the bifactorial developmental model may be conceptualized as a *bifactorial analysis of variance design*. The components of age and cohort represent the factors in the sense of fixed factors, where number of age and cohort groups used is identical with the number of factorial levels respectively. Consequently, the arrangement in Figure 2 consists of a $7 \times 5$ bifactorial design since seven levels of cohort and five levels of age are employed. A $2 \times 2$ design, in which two cohorts are observed at the same two age levels, is the minimum bifactorial design which can be derived from the developmental model.

With respect to the factor age, the bifactorial analysis of variance might be conceived with *independent* as well as with *repeated measurements*. A repeated measurement design over the factor age is defined as one sample of the cohort y observed repeatedly, as is done in the case of the conventional longitudinal method. In contrast, independent measurements are present when several comparable samples are drawn from the cohort y which are all observed only once at a specific age level.

*Table III*

Bifactorial Developmental Model. Analysis of Variance Designs for Independent and Repeated Observations (Age)

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Independent observations</th>
<th>Repeated observations on factor age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>df</td>
<td>Source of variation</td>
</tr>
<tr>
<td>Age</td>
<td>A-1</td>
<td>Between Ss</td>
</tr>
<tr>
<td>Cohort</td>
<td>C-1</td>
<td>Cohort</td>
</tr>
<tr>
<td>Age × Cohort</td>
<td>(A-1) (C-1)</td>
<td>$S$s within groups</td>
</tr>
<tr>
<td>Error</td>
<td>AC(n-1)</td>
<td>Within $S$s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age × Cohort</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age $\times$ Ss (within)</td>
</tr>
<tr>
<td>Total</td>
<td>nAC-1</td>
<td>Total</td>
</tr>
</tbody>
</table>

Such an analysis of variance design (Table III) permits the separation of the main effects of age and cohort. Simultaneously, in the presence of a significant interaction between age and cohort, it is possible to infer different age effects for different levels of cohorts. In other words, it is possible to examine the extent to
which (a) the observed behavioral characteristic is affected significantly by age, (b) the observed behavioral characteristic is affected significantly by cohort, and (c) the extent to which the effect of age is different when age is combined with specific levels of cohort. By application of suitable \textit{a posteriori} tests it is subsequently possible to localize the single effects more precisely within the levels of age and cohort. Furthermore, it might be tested whether the obtained age- and cohort-functional relationships fit to some developmental trends provided by theoretical considerations or other empirical studies.

The decision between a design with \textit{repeated} and \textit{independent measurements} is dependent upon various criteria. According to the general theory of experimental design, a design with repeated measurement is more adequate if it is feasible (a) to have a more sensitive test of significance, and (b) to analyze individual variation.

The increase in \textit{sensitivity} seems worthwhile only if the effect of the factor under consideration (age) is expected to be relatively small, if the extent of confounding carry-over effects can be assumed to be negligible, and if it is not too difficult to observe the same individuals repeatedly. Therefore, a general decision between a design with repeated and independent observations to increase sensitivity does not seem reasonable. By considering the criteria of the extent of age, carry-over effects as well as of economy, a design with independent measurements, however, appears to be more often adequate for the study of age and generation effects. Age effects are generally considerable and testing effects are not to be underestimated in most psychological variables. In addition, because of selective drop-out, it is not easy to maintain the original samples over longer time intervals.

On the other hand, a design with repeated measurement is necessary for the analysis of \textit{individual trends} and for the control of \textit{selective survival} effects. The analysis of individual developmental trends corresponds largely to the advantage attributed often to the conventional longitudinal approach, namely that the longitudinal study alone is able to analyze the dynamics and individuality of developmental processes (cf. \textit{Anderson}, 1954; \textit{Baldwin}, 1960; \textit{Jones}, 1958; \textit{Kessen}, 1960; \textit{Thomae}, 1959a, 1959b). Besides the main effect of $S$s (Table III) which indicates the existence of over-all individual differences, it is the interaction term $S$s $\times$ Age in the repeated measurement design which provides information
on the need for an idiographical kind of treatment of developmental trends. It is not yet sufficiently clarified to what extent it will be necessary to study developmental processes in an idio-
graphic frame of reference. Should it be necessary, however, to adapt such an idiographic approach, this would have serious con-
sequences for the generalizability of developmental findings.

*Hypothetical examples.* In Figure 3, six possible outcomes of a bifactorial design having three levels of each age (A₁–A₃) and cohort (C₁–C₃) are considered. It is usual (cf. Winer, 1962, p. 235) to illustrate the results of factorial analyses in the form of geometric representations of the sample means on the dependent variable.

![Fig. 3 a-f. Six hypothetical outcomes of a 3x3 factorial design of the bifactorial developmental model (● = means, A₁–A₃ = three levels of age, C₁–C₃ = three levels of cohort).](image)

Figures 3a–3f represent main and interaction effects of the components age and cohort. In Figure 3a there is neither a main nor an interaction effect since all means are on a single horizontal line. Figure 3b shows the results of a study where a main effect of cohort only was obtained, i.e. the cohorts C₁–C₃ gained progressively higher scores on the dependent variable, while age had no effect. Figure 3c represents the reverse situation. In Figure 3d there is a main effect of age and cohort so that with increasing age as well as with increasing cohort higher scores on the dependent variable are observed. In Figure 3e and 3f there are interaction effects in addition to, or instead of, the main effects of age and

---

11 Human Development, Vol. 11, No. 3 (1968)
cohort. In Figure 3e, besides the main effects, the interaction effect is expressed by a steeper age gradient for \( C_9 \) than for \( C_9 \) and \( C_1 \). Finally, Figure 3f represents a main effect of cohort and an interaction effect since the age gradients are different for the three cohorts. There is, however, no main effect of age in Figure 3f since the means for the three cohort groups per age level are all on a horizontal line.

As was discussed above, it is not allowed to attribute to the components different functional developmental conditions, as was suggested by Schae. In the case of age and cohort effects, for instance, it is legitimate only to infer differential developmental conditions without any specification as to their functional composition. In the case of age and cohort, these differential developmental conditions might refer to all internal and external factors whatsoever influencing subpopulations which are defined on the basis of time characteristics, i.e. age and time of birth. To interpret obtained age and cohort effects contentwise, it would be necessary to make additional assumptions or experiments which extend far beyond the general developmental model.

**Testing effects.** It must be recognized that the problem of testing effects is not dealt with in a repeated measurement design over the factor age. As it holds true for a single longitudinal study, a plan with repeated measurements confounds age and testing. When a design with repeated observations is desirable because of its higher sensitivity and its capacity to analyse individual trends, testing effects might be controlled for by the introduction of adequate control groups. In the realm of the bifactorial model the introduction of adequate control groups corresponds to a

\[
\begin{array}{cccc}
C_{11}A_{1}O_{1} & C_{11}A_{2}O_{2} & C_{11}A_{5}O_{3} & C_{11}A_{4}O_{4} & C_{11}A_{6}O_{5} \\
C_{12}A_{5}O_{1} & C_{12}A_{2}O_{1} & C_{12}A_{4}O_{1} & C_{12}A_{4}O_{1} & C_{12}A_{4}O_{1} \\
\end{array}
\]

Table IV

Control of Testing Effects per Cohort by Simultaneous Application of Longitudinal and Cross-Sectional Sequences

\( C = \) Cohort, \( A = \) Age, \( O = \) Observation

\( C_{11} - C_{16} \) are six comparable samples drawn from the population of a single cohort \( C_1 \). \( C_{11} \) represents the sample of the longitudinal design while \( C_{15} - C_{16} \) represent the samples of the different cross-sectional designs.
simultaneous application of the developmental model as design with independent as well as with dependent observations. Such a simultaneous application of dependent and independent measurements leads per cohort (C₁) to the arrangement shown in Table IV.

The arrangement in Table IV permits the analysis of over-all testing effects. This analysis of testing effects is possible since for each age group of the repeated measurement design (C₁₁A₁—C₁₁A₉) a control group is available which has been observed only once at a specific age level (C₁₂A₁—C₁₆A₉). For the control of testing effects the observations, placed one beneath the other in Table IV, are therefore to be compared. For example, C₁₁A₂O₂ is compared with C₁₅A₂O₁. This control of testing effects can be expanded to all cohort and age levels of the bifactorial model under study. A problem is, however, that such a control group design contains simultaneously information with regard to selective drop-out effects. In other words, in the differences between the samples to be compared testing and drop-out effects might be confounded. Therefore, it should be recognized that even more complex research designs are needed to fully clarify these sources of error. The above example is intended only as a first attempt at developing the kind of experimental set-ups which are required for an adequate control of the effects of testing and selective drop-out.

*Longitudinal and cross-sectional sequences.* It remains to decide how to name the strategies of data collection when transforming the bifactorial developmental model into a research design. It seems reasonable to do this in terms of the conventional methods. Therefore, the relationship of the conventional methods to the bifactorial model will be clarified first.

The bifactorial developmental model can be separated into *unifactorial designs.* Such unifactorial designs consist of varying the levels of one factor only, while holding the second factor constant. The formulation of such unifactorial plans is legitimate since they produce effects which can be interpreted unambiguously. The status of the conventional methods has been illustrated already in Figure 2 (p. 158). It can be seen that both the conventional longitudinal and time-lag methods correspond to such unifactorial designs. In the case of the longitudinal method, the factor age is varied while the factor cohort is held constant. The reverse is true for a time-lag study where the factor cohort is varied while age is
held constant. Therefore, both the conventional longitudinal and time-lag methods can be considered as adequate designs. Consequently, the effects of the longitudinal and time-lag methods can be interpreted as pure component effects, i.e. age or cohort effects. Within the bifactorial developmental model the findings of both methods are limited only with regard to their generalizability over the second factor, i.e. cohorts and ages respectively. The conventional cross-sectional method, however, must be considered as an inadequate research design. In a single cross-sectional study age and cohort conditions are varied simultaneously and consequently age and cohort effects are inevitably confounded. The suggestion to interpret the sample differences obtained by the conventional longitudinal and time-lag methods as pure component effects is contrary to the proposal made by Schaie (1965). On the basis of his trifactorial model, Schaie considered the sample differences of all three conventional methods as the confounded effect of two components.

After the status of the conventional methods within the bifactorial developmental model has been described we can proceed to designate the strategy of data collection needed to fill the cells of the bifactorial model. The kind of terminology should serve to distinguish between an independent and a repeated measurement design with regard to the factor age.

![Fig. 4. Examples for longitudinal (LOS) and cross-sectional sequences (CSS) as data collection strategies.](image)

When applying a design with repeated measurement all cells can be set by observing one sample of all cohorts at all age levels. Since, per cohort, this procedure corresponds to a conventional longitudinal study, the total strategy of data collection can be called a *longitudinal sequence* (Fig. 4). A longitudinal sequence,
then, consists of observing two or more cohorts at the same two age levels, each in the sense of a conventional longitudinal method. In the simplest case one is dealing with a $2 \times 2$ design with repeated measurements over the factor age.

In the case of a design with independent observations with regard to the factor age, it is possible to obtain the required observations by performing successive cross-sectional studies. Therefore, this strategy of data collection may be called a cross-sectional sequence (Fig. 4). Again in the simplest case we are dealing with a $2 \times 2$ design with independent observations over both cohort and age. When applying a cross-sectional sequence one has to keep in mind, however, that a single cross-sectional study represents an inadequate design since it confounds age and cohort effects. For this reason, some of the single cross-sections, composing the total of the cross-sectional sequence, need not necessarily cover all age levels. From the first cross-section, for instance, only the results of the first age level are usable. From the second those of the first two age levels, and from the third cross-section the results of the first three age levels, etc., are usable for the bifactorial model. In principle, it is then possible to conceptualize the first cross-sectional studies over fewer age levels. If the first cross-sections, however, are expanded over all age levels under consideration, the obtained information is not necessarily useless. The observations might be applied to unifactorial designs in the sense of single longitudinal and time-lag methods.

When the bifactorial model is applied both with repeated as well as independent observations concerning the factor age some of the observations are applicable for both the cross-sectional and the longitudinal sequences. All samples which are observed for the first time in the course of the longitudinal designs may be used simultaneously for the cross-sectional sequences as well. Conse-

---

4 On the other hand, these restrictions inherent to the concept of cross-sectional sequence suggest that one should perhaps dispense with the conventional terminology. In this case, the author proposes to designate any cohort-specific age study an age-sequential design or an age sequence. The term age sequence, therefore, would include the conventional longitudinal design as a specific case, i.e., age sequences might be conceived with independent as well as with repeated measurements. The second kind of unifactorial design within the realm of the bifactorial model might be called a cohort-sequential design or a cohort sequence which consists, therefore, of an age-specific cohort study in the sense of Schaar's time-lag method. Consequently, instead of cross-sectional and longitudinal sequences one could call the strategies of data collection age sequences and cohort sequences.
quently, when longitudinal and cross-sectional sequences are applied simultaneously, the single cross-sections need not include the first age level of the bifactorial model under consideration.

Discussion

The general developmental model and the sequential strategies based thereon were first presented by Schaeie (1965). Their purpose was (a) to analyse critically the conventional methods in research on aging, and (b) to formulate more adequate research strategies. The present paper proposed a reformulation of Schaeie's developmental model and of the sequential strategies in an attempt to eliminate methodological deficiencies inherent to Schaeie's arguments. Furthermore, an adequate analysis and interpretation of the developmental model was sought. It is hoped that application of the longitudinal and cross-sectional sequences presented in this paper will clarify the puzzling discrepancies between cross-sectional and longitudinal findings and may contribute to the construction of more valid developmental gradients with regard to age as well as to cohort.

Since the proposed sequential strategies require such a great investment of time and energy, the question of utility might be raised. On the one hand, the amount of investment is mainly dependent upon the range and the number of intervals used regarding both factors age and cohort. On the other hand, the amount of investment depends on the choice of a design with independent and/or a design with dependent observations concerning the factor age. There is, however, no general rule to rely upon. In each single case, the investigator must make a new decision which will depend, for example, upon the kind of dependent variable under consideration, the underlying hypotheses with regard to the extent of age and cohort effects, and the intended range of generalizability.

For many variables of a physical or a biological nature, for instance, it might be feasible to restrict the number of levels with regard to the cohort factor, while in the same study the age factor might concentrate on the lower and higher age levels, or even on a specific smaller age range only. In other instances, as in the investigation of consumer motivation, it might be more interesting to select more levels of cohort and to concentrate on the age range
of early adulthood. In the long run, nevertheless, we need in addition more systematic approaches covering wide ranges of age and cohort as well. This necessitates the cooperation of many people, institutions, and even members of different generations.

With regard to the choice between a design with independent and/or a design with dependent measurements over the factor age, this writer considers a design with independent observations as generally more feasible, at least in light of the current state of knowledge. A design with independent observations is in general more economical and as valid as a design with dependent observations. As a consequence, a design with repeated measurements seems worthwhile only when there are strong arguments for the need of a test with higher sensitivity concerning the factor age and/or strong arguments for the need of an idiographical kind of analysis. In this connection, it has to be recognized that in the realm of the bifactorial developmental model it is not the cross-sectional method which is the appropriate alternative to the longitudinal method but a design where different samples from the same cohort are observed, each at a specific age level.

The question of the relation between investment and utility of the bifactorial developmental model and the sequential strategies derived therefrom, therefore, cannot be answered \textit{a priori} on a theoretical level. The answer is rather contingent on the intentions of the investigator, the empirical relevance of the postulated age, cohort and error effects, as on the need for the analysis of individual variation in the case of a repeated measurement design over the factor age. It might well be that the effects of generation differences, for instance, are very small or even absent in a specific case. Nevertheless, as long as this has not been proved empirically, the hypothesis of generation differences seems worthy of investigation with regard to all variables which might be dependent upon hereditary and/or environmental conditions.

Finally, regarding the utility of the sequential strategies, we should keep in mind that the scientific progress in aging research, as in any other discipline, is contingent largely upon the quality of its methodology. In dealing with continuous changes, as is true in the study of such time variables as age and cohort, we will have to become adapted to much more refined research planning. Strictly speaking, one could conclude that it will not be possible at all to investigate these levels of age and cohort which have
not been studied to date, since the variation of age and cohort is of an ordered and irreversible nature. These restrictions with regard to the logical status of age and cohort as experimental variables underline the necessity for a careful methodology in the field of aging.

Summary

The conventional methods of aging research, the cross-sectional and longitudinal designs, are discussed in a historical frame of reference. An analysis of the research designs underlying both methods demonstrates that the two models exhibit serious deficiencies which make questionable the interpretation of sample differences as pure age effects. The following five methodological shortcomings are described: (a) selective sampling, (b) selective survival, (c) selective drop-out, (d) testing effects, and (e) generation effects.

Recently, Schaefer has presented a general trifactorial developmental model consisting of the components age, cohort and time of measurement. This developmental model was the basis for deriving new strategies which Schaefer called sequential designs. Some inconsistencies in the formal definitions and the measurability of the components sensu Schaefer are pointed out. These inconsistencies necessitate a reformulation of the general developmental model and the sequential strategies derived therefrom. In order to obtain an adequate analysis and interpretation it is proposed that the trifactorial developmental model be transformed into a bifactorial model using only the components of age and cohort (generation). This bifactorial developmental model can be conceptualized as a design with independent as well as dependent measurements over the factor age. An analysis of variance design permits the separation of the main effects of age and cohort and their interaction.

The data collection strategies can be called in the case of independent measurement, cross-sectional sequences, and longitudinal sequences in the case of repeated measurement. In the frame of reference of the bifactorial model, the conventional longitudinal and time-lag methods can be considered as adequate unifactorial designs. However, the cross-sectional method must be considered as inadequate since it confounds age and generation effects.
References


GESELL, A.: The mental growth of the preschool child. A psychological outline of normal development from birth to the sixth year, including a system of developmental diagnosis (MacMillan, New York 1925).


HARRIS, C.W.: (Ed.) Problems in measuring change (Univ. of Wisconsin Press, Madison 1963).


Author's address: Dr. Paul B. Baltes, West Virginia University, Department of Psychology, Morgantown, W. Va. (USA).