CHAPTER 19

Plasticity and Enhancement of Intellectual Functioning in Old Age

Penn State's Adult Development and Enrichment Project (ADEPT)

Paul B. Baltes and Sherry L. Willis

College of Human Development
The Pennsylvania State University
University Park, Pennsylvania

Objectives of ADEPT

The Penn State Adult Development and Enrichment Project, with the acronym ADEPT, is a basic research program aimed at examining the modifiability of intellectual functioning in later adulthood and old age. Intelligence as referred to in ADEPT research indicates performance on psychometric tests of intelligence rather than process-oriented indices of cognitive functioning. The ADEPT domain of psychometric intelligence is defined primarily by the theory of fluid-crystallized intelligence. Later adulthood and old age covers the age range from approximately 60 to 80 years of age.

The research program consists of a series of interrelated, short-term, longitudinal studies that extend over varying portions of time, from approximately one

ADEPT is supported by a grant from the National Institute on Aging (No. R01 AG00403) to Paul B. Baltes and Sherry L. Willis. Results reported cover data analyses up to 1980.

Paul B. Baltes is now at Max Planck Institute for Human Development and Education, Lentzeallee 94, 1000 Berlin 33, Federal Republic of Germany.

month to close to two years. Throughout, the emphasis is on the study of the range of intraindividual modifiability (plasticity) in psychometric intelligence that can be effected. Training strategies are primarily educational in nature. They include practice and instructional programs that are geared towards the teaching of problem-solving skills relevant for adequate performance on selected intelligence tests. Abilities chosen as targets for intervention center on fluid intelligence. Assessment of training effectiveness considers two dimensions of generalization: training transfer to a broader spectrum of intelligence than is addressed in training, as well as maintenance of training effects over time.[1]

On Variability and Plasticity in Intellectual Aging

Our general perspectives on the state of the art in life-span and gerontological research on intelligence have been summarized elsewhere (Baltes & Labouvie, 1973; Baltes & Willis, 1979a,b; Willis & Baltes, 1980). For the present purpose, only selected observations are offered. For example, during the recent decade, research on psychometric intelligence in adulthood and old age has moved at an increasing rate from a descriptive cataloguing of the average course of intelligence (Horn, 1970, 1978; Schaie, 1979) to a differential and causal-analytic posture (Baltes & Labouvie, 1973; Baltes & Willis, 1979a; Botwinick, 1977). Descriptive evidence on average trends continues to be important in our search for a better understanding of life-span intelligence. However, in our view, one of the current frontiers is in examining variability in its various meanings. Three such meanings of variability are discussed here: interindividual, interability, and intraindividual. For the latter meaning of variability, we reserve the term *plasticity*.

Variability

A first meaning of variability refers to *differences between individuals* (interindividual differences). Interindividual differences are of multiple sorts. For example, they can denote within-cohort differences, such as static differences at a given point in time or longitudinal differences between ontogenetic trajectories of distinct individuals. Another type of interindividual differences concerns between-cohort variation. For example, descriptive evidence has suggested that there is much interindividual variability in the level and direction of intellectual functioning between cohorts. Schaie's (1979) groundbreaking work using cohort-sequential designs is a symbol of the evidence.

Descriptive research has also shown much evidence for a second type of variability, *interability variation*. This phenomenon appears under the label of *multidimensionality*. Multidimensionality refers to the need to consider intelligence not as a unified construct, but as a system of abilities with varying courses of ontogeny. Three research programs are particularly convincing on this point. One is research with the Wechsler scale demonstrating differential change functions for distinct categories (verbal versus performance) of the WAIS (Botwinick, 1977). The second and third, and perhaps most persuasive, are the ones associated with Thurstone's primary mental abilities and Cattell and Horn's theory of fluid-crystallized intelligence. In this research, rather distinct age-change functions for different psychometric factors of intelligence have been reported (Cattell, 1971; Horn, 1970, 1978; Schaie, 1970, 1979).

In general, it is our opinion that descriptive research on age-, cohort-, and ability-related variability has shown that there is little power, conceptual or empirical, in organizing the development of adult and gerontological intelligence around a single overarching framework such as universal (across individuals and abilities) decline. On the contrary, we interpret the evidence as suggesting much variability and relatively little age-related invariance. Exceptions are the fluid-crystallized (or similar) distinction and the notion that, for advanced old age (beyond age 70) and for samples of older persons suffering from noted loss of health, more general decline can be observed (Baltes & Willis, 1979a; Schaie, 1979). Except for these qualifications, however, the striking feature of descriptive aging research is one of much variability, both between persons and within persons for distinct abilities.

Plasticity

The third major meaning of variability denotes variability at the *individual* level of analysis, again either at one point in developmental time or over the course of life. In this framework, variability becomes intraindividual modifiability or plasticity. In ADEPT, we are addressing this question of intraindividual variability for a specific age group, that of older persons. Rather than comparing individuals *of different ages or cohorts*, our primary effort is to identify *the range and conditions* of individual modifiability within a given age range.

The term plasticity is used, then, to indicate the range of functioning at the individual level, whether ontogenetic (how variable is the individual course of

[1] ADEPT includes one additional set of intervention analyses dealing with transfer of cognitive training to aspects of personality functioning. This research is being conducted for two reasons. A first is ethical and deals with the question of possible negative side effects of sustained cognitive training on personality functioning in the elderly. The second is substantive and deals with the interface between ability and personality. Several aging- and intelligence-related measures have been constructed for that purpose involving such dimensions as locus of control in intellectual and aging contexts (Lachman, Baltes, Nesselroade, & Willis, submitted). Preliminary analyses show that there is no negative side effect on personality as a function of training. The evidence on more specific relationships involving personality and intellectual ability is not yet fully analyzed.

development?) or concurrent (how variable is performance at a given point in ontogenetic time?). Why do we assign much importance to questions of plasticity? In general, it is our position that theories of psychometric intelligence that do not include or consider statements about plasticity are incomplete (Baltes & Baltes, 1980; Baltes & Willis, 1979a). There are three main reasons. One is related to conceptions of intellectual aging, the second to the nature of psychometric intelligence and its relationship to processual models of cognition, the third to questions of social policy and the optimization of human development.

Plasticity and Conceptions of Intellectual Aging. We do not claim that intellectual aging is completely variable and plastic. Such an extreme view would seriously undermine the basics of the field of developmental psychology which requires some acknowledgement of invariance of ontogenetic process (Baltes & Baltes, 1980; Wohlwill, 1973). However, we do postulate that gaining knowledge about conditions and range of plasticity is legitimate and that such knowledge requires an explicit orientation toward such knowledge and a particular methodology.

In general, the theoretical and methodological orientation associated with research on plasticity involves exposing the same individuals to different life conditions including learning programs. The strategy requires a direct or quasi-experimental form of intervention. In the past, such an orientation and such methodology have not been the mainstream of research in psychological gerontology. Using an interventive approach, we learn not only about how intellectual aging looks in a given set of circumstances, but also how it could look if conditions were different (Baltes & Danish, 1980; Baltes & Willis, 1977). Thus, similar to research on testing the limits and learning diagnostics (Brown & French, 1979; Guthke, 1976; Schmidt, 1971), the focus is on examining the range of intellectual aging under conditions not normally existent in either the living ecology of older persons or in the standard assessment situation provided by classical tests of psychometric intelligence. Note here that, for ethical reasons, our research into plasticity is restricted to conditions assumed to be performance-enhancing. From a theoretical point of view, the study of conditions related to performance decrement is equally important.

Plasticity and Process Research. The second major reason why we consider research on plasticity important involves the question of the how and why of psychometric intelligence. Research on modifiability and plasticity of psychometric intelligence facilitates another desirable movement in developmental research on intelligence, that of reaching a stronger interface between psychometric intelligence on the one hand and processual research on cognition on the other.

Work on psychometric intelligence has been predominantly a static-correlational enterprise with a primary focus on product rather than process. Cognitive research, contrariwise, has been oriented primarily toward understanding the mechanisms associated with the processing of information. On a conceptual level, the need for a rapproachment between a psychometric-correlational and a

cognitive-processual approach toward understanding intelligence has been suggested with dogged perseverance for several decades. Recently, there appears to be a growing commitment to moving beyond rhetoric and isolated work to demonstrate the usefulness of such a rapproachment on the empirical and programmatic level as well (e.g., Resnick, 1976; Sternberg & Detterman, 1979). We believe that research on the modifiability or plasticity of psychometric intelligence is one of the steps in that direction. Such research leads quickly to questions about processes and mechanisms involved in intelligence test performance, thereby requiring explicit attention to models and conceptions of cognition and information processing.

Plasticity and Social Policy. The third main reason why research on plasticity is important relates to questions of social policy. Descriptive research with its focus on averages and normative patterns can lead easily to misinterpretations of the potential of older persons and the comparative level of functioning if different age groups are contrasted. For example, descriptive work on average decline in short-term memory or fluid intelligence can be taken as suggesting that older persons are deficient in cognitive effectiveness. At the same time, such an interpretation would not consider the substantial degree of interindividual variability observed nor the fact that whatever decline is observed in aging is relatively small in terms of proportion of variance accounted for (Schaie, 1979, 1982). In fact, Schaie argues that, despite whatever decline is observed in old age, a central impression must be the large number of older persons who either do not experience any decline into the 70s or who in the 70s function at levels above the mean observed for younger age groups.

Research on plasticity augments such a perspective of large interindividual variability in intellectual aging. Research on plasticity leads not only to information about how older persons perform in a given situation but also to what they would be able to do if conditions were different. It examines, even in the presence of decline, whether and how it is possible to use intervention strategies to achieve, for example, age-equivalent levels of functioning. As a consequence, research on plasticity is inherently aimed at providing a knowledge base apt to suggest procedures for optimization (Brandtstädter & Schneewind, 1977) and redistribution of education resources according to a life-span perspective (Baltes & Willis, 1979b; Reinert, 1980; Schaie & Willis, 1978). In this sense, research on plasticity contributes to a foundation of social policy that is inherently preventive, corrective, and equity-oriented rather than discriminatory or defeatist.

Context and Hypotheses of ADEPT Research

In addition to the general background summarized in the preceding section, there are two more specific domains of research and theory that set the context for the ADEPT research program and index our own belief systems about research priorities in the field of intellectual aging. A first domain is given by ecological and performance considerations of aging behavior. The second is provided by our view and application of Horn and Cattell's theory of fluid-crystallized intelligence.

Intellectual Aging: Role of Performance Factors and Ecological Deficits

On Performance versus Potential. Much of our thinking about intellectual aging has been stimulated by the conceptual distinction between performance and potential (capacity) as well as by considerations of the ecology in which elderly persons are living. Related considerations provide the context for the general hypotheses or propositions tested in ADEPT as also evident in Figure 6, shown later in this chapter.

The performance–potential (capacity) distinction appears in a variety of concepts including the differentiation between performance and competence (Bijou, 1971; Botwinick, 1977; Flavell & Wohwill, 1969). While the conceptual distinctions are not always clear, we accept their basic premise. Performance refers to observable behavior, and any given observable behavior is but a sample of the intellectual behavior that a given individual could display. Potential denotes that level (or range) of performance that individuals exhibit if alternative conditions are considered or introduced. The goal is to specify the conditions under which different classes and levels of intellectual behavior occur. Our position may be misinterpreted because the language chosen is similar to that used often in behaviorism. A behavioristic constraint is not intended. The conditions for performance variation can include nonenvironmental and cognitive determinants.

Accepting the significance of a capacity–performance distinction leads to the conclusion that the organization of knowledge about intelligence is facilitated if the data base includes, from the beginning, empirical statements about plasticity. This is so because an observed level of functioning is but one (though concededly important) possible behavior or developmental outcome. In fact, one could take the position that potential or capacity is inherently an unknown, dialectical quantity (see also Figure 6; Riegel, 1976). What is known only is a given outcome resulting from the transaction of capacity with experience; and experiences, of course, can differ markedly, especially if a lifelong view of development is taken.

Recognizing the facts that intellectual behavior is regulated by conditions of performance and that information about variability and plasticity due to performance factors is an important ingredient for a comprehensive theory of intelligence leads to the conclusion that theories of performance (rather than theories about latent abilities and their interrelationships) are a neglected theme in research on psychometric intelligence. It appears to us that psychometric research on intelligence has ignored such an emphasis to such a degree that the resulting body of psychometric theory is out-of-balance. We know relatively little about the questions of how, why, and under which conditions of performance on intelligence tests.

Intellectual Aging and Performance Factors. What about the role of performance factors in the study of intellectual aging? First, there is an *ecological* argument. The need for understanding the conditions for intellectual performance is enhanced in all situations in which individuals do not represent normative or "average" situations of living. Examples are members of some socially disadvantaged groups. In the past, related arguments have been made especially for members of minorities or other intellectually disadvantaged populations. In each of these instances, the position is that the performance conditions for such "disadvantaged" individuals are not at a comparable level of optimization as is true for the mainstream or for socially advantaged segments of a population.

We believe that the elderly often live in "disadvantaged" conditions. On a general level, this view is illustrated in the work of Lawton and his colleagues (Lawton & Nahemow, 1973). As to intellectual functioning, although the evidence is far from adequate, one of us argued several years ago (e.g., Labouvie, Hoyer, Baltes, & Baltes, 1974) that it is likely that many older persons live in environments which are not conducive to high levels of intellectual functioning. Another example of an ecological deficit view of psychological aging is Margret Baltes's and her colleagues' research on the social ecology in nursing homes (Baltes & Baltes, 1980; Barton, Baltes, & Orzech, 1980) and its role in the aging-related increase of dependence.

The second rationale for suggesting a lowered performance context for intellectual functioning in old age is *ontogenetic* in the life-span development sense. To what degree is performance on intelligence tests an aspiration or a desirable developmental goal (task) of older persons? It does not appear to us that performance on intelligence tests is a particularly salient goal in the life of many elderly persons. Rather, note that psychometric tests of intelligence are more appropriate for the developmental tasks characteristic of the first part of the life span (Havighurst, 1948). Intelligence tests have been developed with a substantive focus on education and occupational life, both important characteristics of young adulthood. It appears difficult to argue that a similar demand for intelligence-related performance exists in the environmental system of older persons.

As a consequence of both perspectives (ecological, ontogenetic), it is our general position that the level of intellectual performance observed in older persons, if observed without performance-enhancing conditions, is probably an underestimate of what older persons could do if they attended to and practiced the intellectual tasks involved. The growing body of evidence on enhancement of intelligence- and cognition-related performance of older persons (for additional reviews, see Denney, 1979; Labouvie-Vief, 1976; Sterns & Sanders, 1980) appears to be in agreement with such a position.

Quantitative versus Structural Change in Intellectual Aging. There is another conceptual reason why a performance orientation is salient in the study of intellectual aging. This rationale deals with the relative importance of developmental-structural determinants at various stages of the human life span.

In the area of cognition, developmental structuralism, such as Piaget's, has been one fruitful approach suggesting fairly invariant sequences and mechanisms of cognitive development in childhood and adolescence. However, theoretical and empirical work stemming from childhood-based structuralism has not yet been extended very fully to the second part of the life span. There are several conceptual efforts to push such an enterprise using, for example, dialectical-contextualistic (e.g., Clayton & Birren, 1980; Labouvie-Vief, 1980; Labouvie-Vief & Chandler, 1978; Riegel, 1976; Schaie, 1977–1978), multivariate-structural (Baltes, Nesselroade, & Cornelius, 1978; Reinert, 1970), but also evolutionary (Brent, 1978) perspectives as the main ingredients for formulating a posture of continued structural development in adulthood.

In future work, such "structuralistic" efforts are important and we underscore their importance. At the same time, particularly where existing data bases are concerned, it appears to us that the current center piece of explanation of adult-developmental change in psychometric intelligence is one of accounting for quantitative rather than structural variability. A performance orientation appears to offer a parsimonious strategy for that task, without prejudging the long-range fruitfulness of a structural-qualitative addition or alternative.

Theory of Fluid-Crystallized Intelligence as Framework

Definition. The second specific context for ADEPT is provided by Cattell and Horn's theory of fluid-crystallized intelligence (Cattell, 1971; Horn, 1970, 1978; Horn & Cattell, 1966, 1967). Our intent in this regard is not to use our research program as a vehicle for testing features of that theory. This could be possible, but it is not our primary intent.[2] Rather, we see the theory as a useful, developmental framework for measurement of psychometric intelligence and our work as contributing to knowledge associated with the theory, namely, about the degree and conditions of plasticity associated with tests indexing some of the abilities making up the theory. The work reported here, then, is seen primarily as offering supplemental rather than confirmatory or falsificatory information on Cattell and Horn's theory.

[2] At the same time, we should acknowledge that our decision not to utilize the present work for purposes of testing the Cattell-Horn theory has substantive reasons as well. The theory, in our view, is not sufficiently precise or developed to permit unequivocal predictions when it comes to the role of experience and the direction of causality.

The meanings of fluid and crystallized intelligence are the following. Both represent broad (second-stratum) dimensions of psychometric intelligence. *Fluid intelligence* (Gf) involves problem solving in tasks dealing with complex relations and novel (relatively culture-free) materials. It is "characterized by processes of perceiving relationships, educing correlates, maintaining span of immediate awareness in reasoning, abstracting concept formation, and problem solving" (Horn, 1978, p. 220). *Crystallized intelligence* (Gc), on the other hand, operates in ways and areas in which the intellectual judgments are not novel but involve educational and cultural knowledge. Like fluid, "crystallized intelligence also involves the processes of perceiving relationships, educing correlates, reasoning, etc. . . .but the content of the tasks that best characterize Gc indicates relatively advanced education and acculturation" (Horn, 1978, p. 221–222). Developmentally, fluid intelligence is prior to crystallized intelligence, both from a temporal and causal point of view. It is also seen as reflecting strongly the neurophysiological state of the organism.

Role in ADEPT Training Research. Specifically, the fluid-crystallized (Gf/Gc) theory provides a useful rudimentary theory base for cognitive intervention research in aging for two main reasons. A first strength is its life-span perspective on quantitative developmental change. For the present purpose, this feature sets the theory apart from some of its competitors such as Guilford's structure-of-intellect model. According to the theory, the two broad dimensions of intelligence exhibit differential "average" or normative change patterns across the life span. Fluid intelligence is said to develop early in the life span and to follow a normative pattern of gradual decline beginning in early adulthood, whereas crystallized intelligence is postulated to show a longer trajectory of increment and a fair amount of stability for most of adulthood into old age.

Because fluid intelligence is the dimension assumed to exhibit aging-related decline, our choice was to concentrate on the system of fluid abilities. The emphasis on examining the range of short-term modifiability in fluid performance does not negate the possibility that an average pattern of decline in fluid intelligence functioning could occur in certain circumstances. Nor do we interpret the Cattell–Horn theory to suggest necessarily that such decline was irreversible, although with the heavy emphasis on neurological and other biological antecedents of fluid intelligence (Horn, 1970), some researchers might reach this conclusion. There is, however, another feature of fluid intelligence that suggests that it is subject to experiential modifiability. This feature involves its comparative standing on a dimension of "degree of learning." While crystallized intelligence is associated with overlearning, performance on fluid tasks is assessed in contexts or with materials that are relatively novel. Thus, one could expect fluid performance enhancement as amount of learning or experience with fluid task increases.

The second main reason for choosing the Gf–Gc theory as framework concerns its measurement characteristics and their relevance for the assessment of training effectiveness. A major task of cognitive training research is to work within a theory and measurement framework that permits (a) definition of training substance and (b) assessment and evaluation of training effectiveness in a multivariate and systems context (Glaser & Resnick, 1972). The hierarchical and multivariate theory of fluid-crystallized intelligence offers this opportunity. For the most part, the theory makes available a battery of ability-indexed tests and statements about their interrelationships at different levels of aggregation, such as first- and second-stratum factors. Using such a model of measurement permits an orderly approach to the specification and assessment of intervention, both in terms of expected effect patterns as well as in terms of scope (breadth) of training transfer.

General Design Characteristics

The various short-term longitudinal studies conducted in ADEPT share several design characteristics and objectives. Although these characteristics are not identical for each one of the studies, they are sufficiently general to warrant an introductory summary rather than repetitive presentation.

Subjects

Because the studies were conducted over a period of several years, it was not possible to draw the participants in the various studies from the same parent population and then randomly assign the individuals to the different projects. Each of the studies is based on an independent strategy of sample selection and random subject assignment. However, all participants come from the same general area and for each project from several community sites. Thus, on the average, the study samples represent fairly comparable parent populations.

Specifically, all participants come from rural Central Pennsylvania. They are volunteers, live in community settings, and were recruited through various local organizations such as churches or senior citizens centers. Participants were reimbursed, on the average, approximately $2.00 per hour for their participation, with the funds being paid more often to organizations rather than individuals.

In terms of subject characteristics, aside from being volunteers and affiliated with at least one community organization, the participants on the whole were fairly unselected. In terms of health, the only requisites were physical mobility, to permit attendance at a local testing site, and adequate sensory (vision, hearing) and psychomotor (writing) ability to engage in psychometric testing. The age level of the participants ranged typically from 60 to 80, with an average age of approximately 70. In accordance with demographic characteristics, the samples

usually included a larger number of women than men, at a ratio of better than 2:1. In terms of education, most samples are above the mean level for their cohort-census for Central Pennsylvania. Mean number of years of schooling is close to 11 years, which is about 2–3 years above the median educational level of their age/cohort.

In general, then, participants are positively selected in terms of health and education when compared with census data for the region involved. However, because the participants come from many communities and because of reasonably large within-study variability, it is our conclusion that the samples are fairly heterogeneous and representative of a large segment of the total population of older persons living in Central Pennsylvania.

Measurement Battery

The general framework of the measurement space used in ADEPT is summarized in Table 1. First, in line with the Gf–Gc theory, its two major second-stratum dimensions were included: fluid intelligence and crystallized intelligence. In addition, one of the other broad dimensions of psychometric intelligence, speediness, was incorporated. Then a set of relatively pure primary mental abilities representative of the three general dimensions were identified as marker factors with associated tests. The hypothesized relationships between the broad factors and tests of primary mental abilities summarized in Table 1 are based primarily on Cattell (1971) and Horn (1970, 1975, 1978). As a total package, the measurement battery overrepresents the dimension of fluid intelligence. This strategy was chosen because training is focused on that dimension and a refined assessment within the second-stratum factor of fluid intelligence is desirable.

At the present time, the assessment of training transfer presented in this manuscript is based for the most part on a subset of the 17 tests shown in Table 1. These tests are set in italics. Their selection is guided by the wish to achieve a refined assessment of transfer within the fluid domain and a more gross assessment of transfer to crystallized intelligence and the factor of perceptual speed. One of the training studies reported (attention training), in addition, involves tests of memory span and a battery of tests designed to measure attention and other aspects of information processing. This latter training study is still in process and awaits final data collection and analysis, and only very preliminary findings are reported in this chapter. The tests chosen or developed for the attention/memory domain include letter matching, letter digit, continuous paired associates, number–word, Stroop, concentration, and word recognition. These tests are designed to index four categories of attention: perceptual discrimination, selective attention, attention switching, and concentration–vigilance. In format and rationale, the tests are constructed in analogy to work on information processing (e.g., Hunt, Frost, & Lunneborg, 1973; Posner & Mitchell, 1967).

Table 1. *Schematic Design of Measurement Battery: Hypothesized General Intellectual Dimensions, Hypothesized Primary Mental Abilities, and Marker Tests*[a]

General dimension	Primary ability	Test	Source
Gf	CFR	*Culture Fair Test (Scale 2, Form A)* and *Power Matrices (Scale 3, Form A [1963 ed.] and Form B [1961 ed.])*	Cattell & Cattell, 1957, 1961, 1963
	CFR	*ADEPT Figural Relations Test (Form A)*[b]	Plemons, Willis, & Baltes, 1978
	CFR	*Raven's Advanced Progressive Matrices (Set II)*	Raven, 1962
Gf	I	*ADEPT Induction Test (Form A)*[b,c]	Blieszner, Willis, & Baltes, 1981
	I	*Induction Composite Test*[b]	Ekstrom *et al.*, 1976; Thurstone, 1962
Gf	Ms	Visual Number Span	Ekstrom *et al.*, 1976
	Ms	Auditory Number Span Backwards	After Ekstrom *et al.*, 1976
	Ms	Auditory Number Span Backwards with Delayed Recall	After Ekstrom *et al.*, 1976
Gf/Gc	CMR	Verbal Analogies I	Guilford, 1969a
	CMR	Word Matrix	Guilford, 1969b
Gc	EMS	Social Translations (Form A)	O'Sullivan & Guilford, 1965
	EMS	Social situations (EPO3A)	Horn, 1967
Gd	V	Verbal Meaning (9–12)	Thurstone, 1962
	V	*Vocabulary* (V-2, V-3, V-4)[b]	Ekstrom *et al.*, 1976
Gs	PS	Finding A's	Ekstrom *et al.*, 1976
	PS	*Number Comparison*	Ekstrom *et al.*, 1976
	PS	*Identical Pictures*	Ekstrom *et al.*, 1976

[a]Most training research reported in this manuscript includes all measures set in italics. Research on attention training also included Memory Span (Ms) and newly developed attention measures as reported in text. Based on Baltes *et al.*, 1980.

[b]Tests labeled ADEPT have been developed by the investigators for assessment purposes in cognitive training research.

[c]Induction and Vocabulary are composites of several "subtests" (see text). The Induction Composite subtests include Letter Sets (Ekstrom *et al.*, 1976), Number Series, and Letter Series (Thurstone, 1962).

As we have mentioned already, the intervention analyses reported in this chapter are based on test scores at the primary ability level. We have conducted factor-analytic research, however, that yields information on the structure of the 17 tests shown in Table 1 (Baltes, Cornelius, Spiro, Nesselroade, & Willis, 1980). These structural analyses provided information on the nature of the interrelationships of the tests contained in Table 1 as well as on the relationships at the factor level. Such information is important because past research using the Gf–Gc theory has assumed largely that there is factorial invariance across the adult life span into old age. At the same time, there is evidence from other life-span work on the structure of psychometric intelligence (Baltes, Nesselroade, & Cornelius, 1978; Cunningham, 1980; Reinert, 1970) suggesting structural transformation rather than invariance, namely, a change with aging from a highly differentiated structure in adulthood towards one that exhibits features of integration (or dedifferentiation or neointegration).

With ADEPT data based on $N = 109$ elderly subjects (age range 60–89), we obtained evidence to support such an integration position for old age. Specifically, on the level of test correlations, the 17×17 matrix showed fairly high intercorrelations, with the majority in the range from .4 to .6. Moreover, when comparing different factor models, varying in degree of differentiation versus integration, "integrated" models with fewer factors (including one containing a general factor) provided better fits to the data using COFAMM (Sörbom & Jöreskog, 1976) as a strategy of confirmatory factor analysis. Thus, while the intervention analyses reported in this manuscript are based on test scores (thereby representing psychometric intelligence at the level of primary abilities), it is important to keep in mind these factor analytic outcomes reported in Baltes *et al.* (1980). They suggest a high degree of relatedness (commonality) among the tests and, in addition, some structural transformations that are not in complete agreement with Gf–Gc findings reported for younger adults. Such a high degree of commonality, for example, may suggest that transfer of training from one ability to another is more likely in older than younger adults.

Rationale and Substance of Training

What distinguishes ADEPT training research from earlier training research involving intellectual aging by others and us? As reviewed, for example, by Denney (1979), there has been an increasing number of cognitive intervention studies during the last decade. Many of these earlier studies resulted in positive outcomes. Few, however, were programmatic in the sense that they considered: (a) a theory-based selection of intervention targets, (b) theory-based definition and assessment of scope of intervention transfer, and (c) maintenance of intervention effects beyond immediate posttest. These are the special features of ADEPT training research.

The present chapter provides intermediate results of five independent but related training studies. A first study (Hofland, Willis, & Baltes, 1981) focuses on the role of retesting itself as a condition of training. The remaining studies deal with training of ability-specific cognitive skills. Three first-stratum abilities of the second-order dimension of fluid intelligence are the target of a total of four cognitive training studies: *figural relations* (Willis, Blieszner, & Baltes, 1981; Willis, Blieszner, & Baltes, submitted), *induction* (Blieszner, Willis, & Baltes, 1981), and *attention/memory* (Willis, Cornelius, Blow, & Baltes, 1982). Results from the last training study are highly tentative.

The training studies reported here reflect two levels (strategies) of cognitive intervention.

Test Practice. A first level involves minimal intervention procedures in that only *test familiarity* or *test sophistication* is manipulated. This was implemented by conducting a study using retesting with identical tests, under no-feedback conditions, as the independent variable. The goal of this research was to obtain baseline information on plasticity as a function of testing experience alone.

This strategy of intervention has a long history (Greene, 1941; Vernon, 1954). In recent decades, however, some exceptions such as the controversy about the role of coaching in educational selection tests (e.g., Pike, 1978; Wing, 1980) notwithstanding, the question of practice and its role in performance on intelligence tests has received comparatively little interest. This is particularly true for research on aging.

Factor (Ability)-Specific Problem-Solving Skills. The second and more direct level of training involved the instruction on *ability-specific problem-solving skills* associated with the three primary factors defining fluid intelligence: figural relations, induction, attention–memory. For each of the three primary abilities indexing fluid intelligence, a training program was developed. Each training program consisted of material and instructions for five one-hour training sessions. The content of the training programs was based on a task analysis of marker tests (rules required for solution and task format) for each of the intelligence factors. In the task analysis, rules and problem-solving strategies required for solution of test items were identified.

For example, in the case of figural relations, the intelligence test used for program development was the Culture Fair Test (Scale 2) consisting of four subtests (figure series, figure classify, matrices, topology). From the task analysis, relational rules (e.g., size, shape, position) utilized in solving items were identified. Training problems were then developed based on the most frequently occurring relational rules. Only the relational rules were used in developing training items; none of the training items was identical to those on the test itself.

A similar strategy of defining training substance was employed in the development of training programs for the two remaining factors of fluid intelligence. In the case of the induction factor, the induction tests from the ETS Kit of Factor-Referenced Tests (Letter Sets) and Thurstone's (1962) PMA battery (letter series, number series) were used as criteria for program development. For the attention/memory factor, training content was taken from the seven tests mentioned above defining four categories of attention/information processing: perceptual discrimination, selective attention, attention switching, and concentration–vigilance.

Training was conducted by one instructor in small groups ranging usually from four to eight persons. The primary instructional strategy utilized in training is best characterized as an adaptation of verbal rule learning which has been effective in both prior problem solving (Wittrock, 1963, 1966) and Piagetian training research (Beilin, 1976). Emphasis was also placed on presenting the material and the rules with much concern for individual styles of problem solving rather than focusing on a single set of correct solution strategies. Training sessions for the different ability-specific programs varied in format. However, in general, the five training sessions focused on the relational rules associated with the various test problems. For example, training on figural relations (involving the four types of problems defining the figural-relations subtests) consisted of individual practice with paper and pencil training materials, oral feedback by the instructor, and group discussions of problems. The fifth and last session consisted of an overall review of the types of problems indexing the fluid factor considered in the training study.

Assessment of Training

Two major criteria are used in assessment of training. The first deals with scope or breadth of training, the second with maintenance of training over time.

Transfer of Training (Scope). Scope or breadth of training is examined within a theory-based measurement paradigm provided by the Gf–Gc model of intelligence. With the Gf–Gc model as measurement framework (see Table 1), training is assessed not only with regard to the target ability factor for which a training program was developed, but also with regard to other ability dimensions contained in the Gf–Gc model, including perceptual speed. The tests used in a given training assessment are ordered along a continuum of near-transfer to far-transfer. In line with this ordering, a hierarchical, theory-based pattern of differential training effects (differing in magnitude) is predicted.

The structure of the predicted transfer pattern for each of the studies follows both from the Gf–Gc model (e.g., its factor space) and the degree of similarity of the transfer measures to the content of the training program. It was not expected that the range of training transfer would cover all dimensions of the Gf–Gc model. What is always expected, however, is that training itself would be effective, as assessed at the nearest measure of transfer, most notably with the test developed for that purpose. Near (within-ability) transfer might also involve a training effect on several tests most similar to the target ability trained, usually indexed by tests marking the same ability factors. As to transfer beyond the ability trained, we should also note that, in line with the high degree of integration of psychometric intelligence obtained (Baltes *et al.*, 1980) in elderly persons, some transfer to abilities other than the target ability was expected. In general, however, the breadth of transfer is seen as a question of empirical analysis, except for the prediction that if training extends to additional factors or abilities, it should be ordered hierarchically in magnitude.

Maintenance of Training. The second criterion for assessment of training involved its maintenance with time. All evaluation measures, therefore, are given at several posttraining occasions: one week, one month, and six months following training. If training effects were to be interpreted as representing modification of the level of functioning on the "latent" target fluid ability, the extent of temporal maintenance of such effects is critical.

Following training, no further efforts are made to have participants continue practice with problem-solving skills or engage in related activities. In line with general conceptions of memory, therefore, one could argue that maintenance of training would display the typical pattern of a forgetting curve.

The intervention studies contain an additional condition of performance associated with retesting. The repeated testing involved in assessment of the maintenance of training is expected to represent a performance-related enhancement condition. This is especially true for test-naive populations such as the elderly. Retest effects must be distinguished from ability-specific transfer of training effects. In contrast to ability-specific training effects, we predict that, by and large, retest effects are general (extending across all abilities) rather than hierarchically ordered. Description and analysis of retest effects associated with posttraining assessment, however, are not presented in any detail in this chapter (see, however, Blieszner *et al.*, 1981).

Design and Analysis. All studies are based on a treatment-control group comparison, with random assignment to the respective condition. In all but one instance, the control groups are no-contact control groups, that is, control subjects participate in assessment sessions only. There is no effort made to provide control subjects with a time- and setting-equivalent amount of experience comparable to the five training sessions delivered to the treatment subjects.

Two studies in progress, however, are aimed at specifying further the role of the nature of contact *per se* as a possible treatment condition and the confounded effect of retesting as a function of multiple posttraining assessments. The first study (Willis *et al.*, 1982), reported here in the context of attention/memory training, compares two levels of control: no-contact with social contact. The second study (not yet analyzed) uses a design that varies systematically the frequency of posttraining assessment.

Data analysis of the various studies involves, for the most part, analysis of variance, covariance, and the systematic application of planned comparisons and *a posteriori* tests. In each instance, an effort is made to achieve first an overall assessment and then a test with maximum statistical power. In addition, the statistical procedures selected are such that the structure of the measurement space is not altered. Thus, whenever assessment instruments are selected to represent particular cognitive abilities, univariate analyses rather than multivariate analyses (requiring computation of composites with unknown test validity) are used.

A final note on data standardization: In order to permit a base of comparison, data from the training studies are transformed to a common metric, separately for each of the assessment instruments involved. The metric used is derived from the performance of the control group at the first posttraining occasion. This procedure is the best estimate for level of performance at baseline. Using the control group's performance at the first posttest as criterion, the metric is set at a mean of 50 with a standard deviation of 10. This procedure facilitates comparative assessment and evaluation of the magnitude of the effects shown in the figures summarizing training outcomes.

Results of Intervention Studies

In the following, main results from the five studies are presented. More detailed information is contained in the publications listed in conjunction with each of the projects.

Retesting or No-Feedback Practice

A first study (Hofland, Willis, & Baltes, 1981) deals with the examination of the effect of practice with intelligence tests *per se*. Two timed measures (Culture Fair, induction composite) of fluid intelligence were used representing the fluid abilities of figural relations and induction. Thirty older subjects (mean age: 69 years) participated in eight one-hour practice (retest) sessions distributed over approximately one month. At each session, the same tests were administered. No feedback on individual performance was given.

Based on indirect evidence involving younger adults, we expected performance improvement for anywhere from two to five retest sessions. Figure 1 summarizes the outcome for elderly adults showing mean percentage for correct solutions for each of the two fluid measures. Statistically, there is a continuous incremental trend in performance across the eight retest trials. Total improvement in mean scores on both measures is slightly more than one standard deviation. No apparent asymptote is reached after eight trials.

The Hofland *et al.* study itself does not permit examination of two ques-

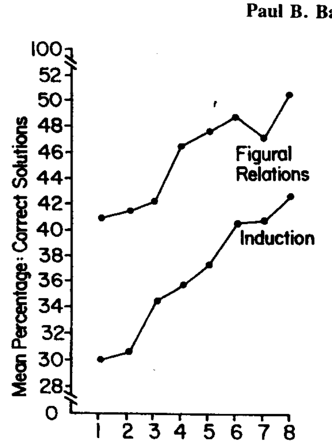

Fig. 1. Mean percentage of correct solutions across retest trials for tests of figural relations and induction. (After Hofland, Willis, & Baltes, 1981.)

tions: (a) whether participants would have reached a similar level of performance if the tests had been given in an extreme power (no time limit) condition at the first retest occasion; and (b) whether retesting itself leads to transfer of training to other ability tests. Relevant studies are currently in progress and initial findings on the role of power versus speeded conditions of assessment are contained in Hofland *et al.* (1981).

However, the data were analyzed for changes in test validity as a function of practice and for evidence on interindividual differences in performance trends. When correlating retest performance with a set of external marker tests (reasoning, memory span, crystallized knowledge, and perceptual speed), there is very little evidence for a testing-related change in the validity of the two retests. Thus, in terms of correlational validity, what is measured at the eighth retest, is similar to what the two tests measured at the beginning of the eight retest trials. Correspondingly, performance increment with retesting does not appear to reflect increasing test specificity associated with the acquisition of test-specific performance skills. Rather, the observed retest increments, in line with one of Greene's (1941) early interpretations, are likely the result of increments in solution strategies (level and/or speed) associated with the two ability factors, figural relations and induction.

As to interindividual differences in performance gains with retesting, intertrial stability coefficients were computed. Throughout, these stability coefficients are very large and approximate the level of estimated reliabilities. We take this information as evidencing that performance improvement as a function of retest practice is similar for all levels of initial functioning. This conclusion is

further supported by the finding that trial-related interindividual variability does not show substantial changes, although there is a slight increase in variability toward the last trials of the retest schedule.

Training of Figural Relations Ability

In earlier ADEPT pilot research (Plemons, Willis, & Baltes, 1978), it was already possible to show that performance of elderly persons on figural relations could be enhanced by means of an eight-session training program. That study also showed transfer of training as predicted by the theory, with no transfer to two training-far tests, induction and verbal comprehension.

This pilot evidence has now been enriched by two additional ADEPT studies dealing with figural relations. A first (Willis, Blieszner, & Baltes, 1981) is a replication of the Plemons *et al.* work using an improved training program and an extended battery of tests permitting a more refined assessment of transfer of training. A second study (Willis, Blieszner, & Baltes, submitted) extends the same training program into a two-year longitudinal design involving multistage training. This procedure involved a second application of the training program to the same persons after the three posttests for the first stage of training were completed.

Replication and Extension. The results for the new training study (Willis *et al.*, 1981) targeting on figural relations are based on 58 older adults with a mean age of approximately 70 years. Seven tests served as criterion measures at the three posttraining occasions. Figure 2 summarizes the outcome.

Based on statistical analysis, both the predicted training effect on training-near measures and the predicted pattern of differential transfer were judged to be significant. The pattern of training transfer is maintained across all three post-

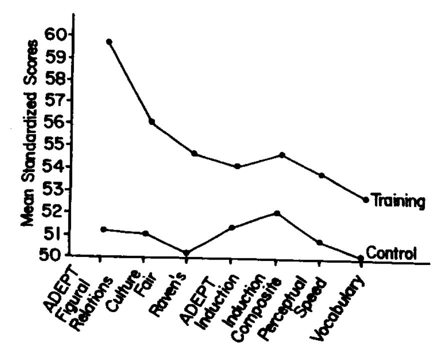

Fig. 2. Standardized mean scores on seven transfer measures for training and control groups averaged across three posttraining occasions. (After Willis, Blieszner, & Baltes, 1981.)

tests, spanning a six-month period. On the nearest transfer measure, the magnitude of training corresponds to approximately one standard deviation if quantified in terms of the baseline performance of the control group at the first posttest.

If the statistical analysis is conducted separately for each of the seven assessment instruments, significant differences (in favor of the figural relations training group) are obtained for the three nearest transfer measures: ADEPT figural relations, Culture Fair, and Raven. However, training is also to some degree effective for the four far-transfer measures. When increasing the statistical power by using a repeated measures analysis of covariance conducted on just the four far-transfer measures, a significant treatment main effect for the four far-transfer tests is obtained as well.

Two-Stage Training. Sixty-three percent (N = 36) of the subjects completing the Willis *et al.* (1981) replication and extension study agreed to participate in and completed a second stage of figural relations training. This resulted in a two-stage training project extending over a total period of approximately 18 months.

Approximately four months after the six-month posttest for stage 1, participants received five additional hours of training of figural relations involving the same trainer and training program as in stage 1. Subsequent to stage 2 training, subjects were given again three posttests (one week, one month, six months) using the same battery of transfer measures as was true for posttesting of stage 1. Various control analyses for dropout effects were also conducted using stage 2 performance as the dependent variable. There was no strong evidence that stage 2 participants differed in intellectual performance during stage 1 from those subjects who did not continue from stage 1 into stage 2 or did not complete stage 2.

Figure 3 shows the main outcomes related to effect of training and training transfer. Statistical analyses supported the conclusion that averaged over both stages (a) training was effective and (b) training resulted in differential transfer. On the individual level of assessment instruments, significant training effects were found for ADEPT figural relations, Culture Fair, Raven, and perceptual speed. Planned comparisons also indicated that the overall pattern of test performances did not differ for stage 1 and stage 2.

Did stage 2 add anything significant to stage 1 beyond ensuring long-term maintenance of training effectiveness? It appears to us that there is no good evidence that stage 2 training contributed to the level of performance reached by the training subjects following stage 1 training. Stage 2 performance on near-transfer measures is higher than that following stage 1. However, differences between stage 1 and stage 2 are accountable in terms of retesting rather than ability-specific training.

This two-stage training study of figural relations, then, can be seen as a further replication and longitudinal extension of the results obtained with one stage of training. What is particularly noteworthy is not only that training was effective but that the pattern of differential transfer was maintained across two stages of training. On the nearest test of transfer (ADEPT figural relations), the magnitude of training (in addition to retest benefits) is about 1.2 standard deviations as measured by the baseline performance of the control group.

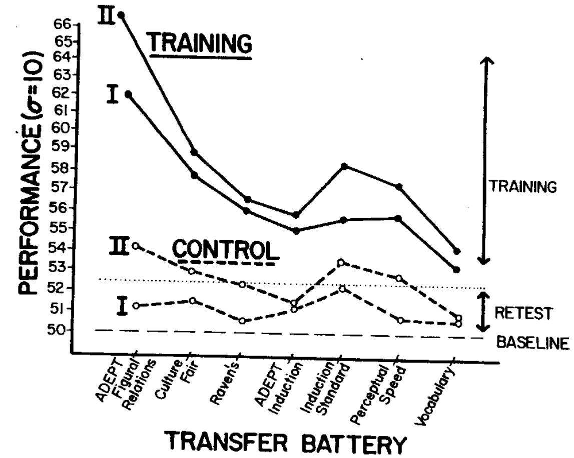

Fig. 3. Standardized mean scores on seven transfer groups following two stages of training (data are averaged for each stage across three posttraining occasions. (After Willis, Blieszner, & Baltes, submitted.)

Training of Induction

The next intervention study (Blieszner, Willis, & Baltes, 1981) concentrated on induction, the second of the three markers of fluid intelligence. Within the fluid-crystallized model, induction has been defined as the eduction of relationships in reasoning tasks that do not have semantic content.

Three published tests indexing induction were identified as substantive domain: letter sets, number series, and letter series. These tests served as the content for which task analyses were performed to determine relational rules (strategies) necessary for solving induction-type problems. The most frequently occurring relational rules were then used to develop items for the nearest test of training transfer (ADEPT induction test) and for problems used in the training sessions. Training consisted, as was true for figural relations, of five one-hour training sessions.

A total of 52 older persons (average age: 70) participated. Figure 4 displays the main outcomes. Training was effective, though to a lesser degree than was

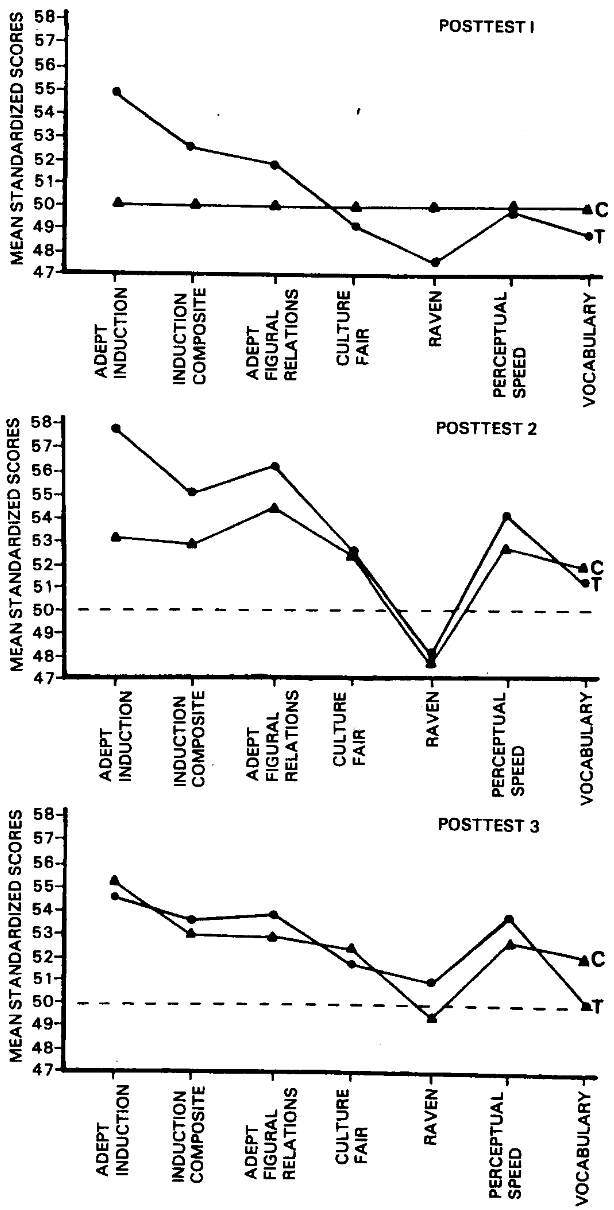

Fig 4. Standardized mean scores on seven transfer measures for training and control groups at three posttraining occasions. (After Blieszner, Willis, & Baltes, 1981.)

true for figural relations. There is statistical evidence for (a) a significant training effect on the nearest test of transfer and (b) a differential pattern of transfer at the first two posttests. At the third posttest, however, the control group (both as a function of retest increments and perhaps because of dissipation of training effect on the part of the training group) has reached the same level of performance as the training group. The fact that the second nearest test of transfer (induction composite), while in the appropriate direction, did not reach statistical significance indicates that induction training was rather specific if compared to the outcome involving training of the ability of figural relations.

Is the plasticity of induction less than that of figural relations? We are not sure yet. Informal assessment of the effectiveness of the induction program during training by the trainer and the second author suggested (already prior to data analysis) that the program may not have been developed optimally. This is an alternative explanation for the fact that induction training was less effective and narrower than training on figural relations. That alternative interpretation is also supported by the fact that induction and figural relations exhibited comparable increments due to retesting alone (Figure 1). Research is in progress to examine this question in a replication study using a modified training program. That study is also designed in such a manner that the effect of training can be identified independent of the effects of retesting. This is accomplished by adding subgroups that are posttested for the first time at occasion 2 and occasion 3, respectively.

Training of Attention/Memory

Development of a training program for the third major ability indexing the second-stratum factor of fluid intelligence focused on several cognitive processes considered to be associated with attention-memory (Willis, Cornelius, Blow, & Baltes, 1982). This research is not yet complete. Thus, only preliminary data, restricted to near transfer (within attention/memory) and the first posttraining assessment, can be reported at this time.

There is some question whether attention/memory should be seen as an intrinsic part of the Gf/Gc realm or as primarily a nonintellectual performance factor more similar to broad speededness than to fluid intelligence. The theoretical and empirical relationship between attention/memory and fluid intelligence is not precise, in part because Horn and his colleagues (1975, 1978; Horn, Donaldson, & Engstrom, 1981) are in the process of refining the meaning and multivariate location of various indices of memory and attention. Our research focus derives from earlier writing, where Cattell and Horn (Horn & Cattell, 1967; Horn, 1970) had identified at the primary ability level a memory component of fluid intelligence. In addition, we consider work on information processing aimed at relating attentional to memory processes. Relevant work on attention and memory processes includes that of Hunt, Frost, and Lunneborg (1973) and, in the aging literature, that of Botwinick and Storandt (1974), Rabbitt (1977), and Craik (1977). Historically, the memory component of fluid intelligence is indexed by measures of secondary memory with backwards memory span serving as a marker test. In our own conception, and very much in line with Horn's current writings, we relate memory functioning of the fluid type to work in the information-processing literature.

Whether this attempt to combine memory as a component of fluid intelligence with memory–attention processes of the information-processing kind will be successful is an open question at present. In addition to the attention/memory training study reported here, we have research in progress (Cornelius, Willis, Nesselroade, & Baltes, submitted) that is aimed at identifying the correlational and factorial relation between the measurement system summarized in Table 1 and the attention/memory tests used in the present intervention study.

Table 2 summarizes the attention/memory measures used. The tests were modified for administration in group settings and for use with older adults. Scoring of the tests involved number of correct answers in a time-limited situation. The given tests deal with four dimensions of attention/memory: *perceptual discrimination* (speed and accuracy in discriminating similarities and differences in perceptual tasks), *selective attention* (ability to focus on a task or stimulus while ignoring irrelevant aspects), *attention switching* or reorientation (ability to shift attention from one task to another and back), and *concentration–vigilance* (sustained attention to or persistence in simple tasks over long periods of time).

The attention/memory training program was delivered in five one-hour sessions with groups of three to eight subjects. The first four sessions each dealt with one of the four dimensions of attention. Practice problems were derived from the marker tests; none of the training problems was identical in content with the test items. For example, in the training session involving selective attention, subjects were shown rows of colored stars and dots and asked to count as fast as possible the items of a given color (ignoring shape), or to count the items of a given shape (ignoring color). Instructional materials included a trainer's manual and individual practice booklets for each of the sessions, and subjects were encouraged to identify performance strategies of their own. The fifth session was devoted to a review of all four attention/memory dimensions.

Table 2. Attention/Memory Training Study: Assessment Battery

General dimension	Test	Source
Perceptual discrimination	Letter Matching	Hunt, 1978 (after Posner)
Selective attention	Underwood Letter-Digit Task	After Underwood, 1975
	Stroop Color-Word Interference	After Stroop, 1935
Attention switching	Continuous Paired Associates	After Atkinson & Shiffrin, 1968
	Number–Word Test	After Wickens, 1970
Concentration–vigilance	Concentration	After Düker & Lienert, 1959
	Word Recognition	ADEPT, 1979

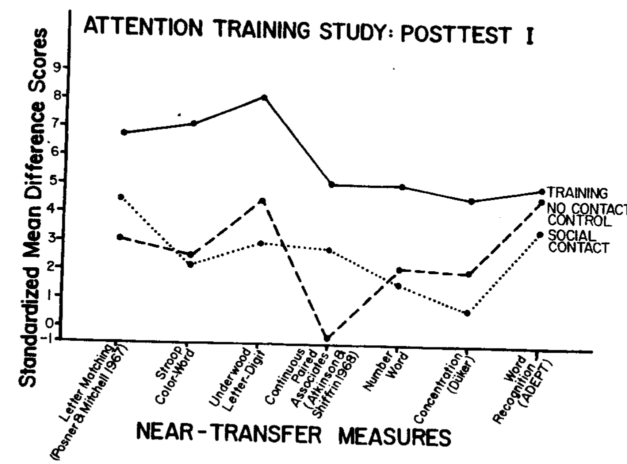

Fig. 5. Standardized mean scores on seven attention/memory measures for training and two control groups at first posttraining occasion. (After Willis, Cornelius, Blow, & Baltes, 1982.)

Figure 5 summarizes available results of the attention/memory training study. Data are based on three randomly assigned groups and a total of $N = 77$ participants with an average age of approximately 70 years. Two no-training control groups were used—one corresponding to past no-contact control groups. The other control group participated in five sessions of social contact with an ADEPT leader. Substance of these social-contact control group sessions involved discussion surrounding questions of life review and friendship.

The results presented in Figure 5 are restricted to assessment of near-transfer *within* the attention/memory domain. Moreover, the data are based on the first posttraining occasion only. Assessment of additional transfer to the measurement battery used in previous training research is not yet complete. Statistical analysis of the data presented indicates a highly significant overall effect of training using a multivariate analysis of variance as strategy of significance testing. Univariate follow-ups specify further that there is significant improvement on three (letter matching, Stroop, continuous paired associates) of the seven tests. Strong suggestive evidence for improvement on two additional tests of the seven (letter–digit and number–word) is present also. Note further in Figure 5 that the direction of the effects is consistent.

The available analyses, then, suggest that there is improvement on attention/memory tests on the one-week posttest following five hours of training. The pattern of outcomes is a rather consistent one. Further analyses are aimed at examining whether this near-transfer is supplemented by far-transfer to other dimensions of the domain of psychometric intelligence, and whether training effectiveness is maintained over longer periods of time.

Conclusions

The entire research program of ADEPT is not yet completed, nor is it fully analyzed. In particular, we await final analysis of the transfer and delayed posttests regarding the attention-training study and outcomes from the replication and extension of the induction training project. The latter is especially important because these results will permit us to be more precise in interpretation of the differences in breadth of training effectiveness observed when comparing figural relations with induction. In addition, this study will help us in specifying more clearly the role of retest effects and their interaction with training effects when assessing training maintenance over several posttraining occasions.

The following observations, then, are preliminary. Yet, we believe that the outcomes are sufficiently consistent to warrant a few tentative conclusions. In general, we conclude that findings from both the retest and ability-specific training studies suggest considerable plasticity in fluid intellectual performance in old age.

Retest Effects

Manipulation of test familiarity associated with retest effects provides the first data base for evidence on plasticity. Substantial retest effects occurred under minimal practice conditions in which subjects received no experimenter-guided feedback on individual performance.

The strongest evidence for the impact of testing *per se* is contained in the Hofland *et al.* (Figure 1) study. This study, because of its direct focus on retesting as a treatment condition with few tests, involved massed practice. Therefore, it represents the best learning situation concerning retesting effects in ADEPT research. In this study, a continuous performance increment (without apparent asymptote) for the two fluid measures was observed. In our view, this is a remarkable outcome. The finding gains additional significance because the retest measures do not appear to change their factorial validity as retesting progresses.

Note further that retest effects, though of a lesser amount, are also found in the control groups of the training studies. These data are not presented extensively in this chapter but are contained in the original papers. Control groups who participated in from three to six posttraining assessments showed consistent improvements in their performance on practically all tests as a function of retesting. What is also important to recognize is that, as predicted, retest effects in the control groups are fairly general. Contrary to the ability-specific training effects, they occur more equally for all tests. In the past (e.g., Hoyer, Labouvie, & Baltes, 1973; Plemons *et al.*, 1978), we interpreted this generalized retest pattern as indexing primarily the operation of general (ability-nonspecific) performance factors. Following Greene's (1941; see also Wing, 1980) early suggestion, however, the acquisition of ability-specific skills as a function of retest experience is a viable alternative interpretation. One example of such an ability-specific interpretation of retest effects is to view them as an increase in the speed it takes to solve test items rather than as an increase in the level of difficulty of cognitive tasks involved. Because most of the tests used in this study are timed, such an interpretation is plausible and deserves more careful examination.

In the absence of age-comparative data, it is difficult to make statements about age differences or age changes in magnitude and scope of retest effects. In other words, are the retest effects observed in elderly persons larger than those reported or expected for younger age groups? We cannot answer this question. However, we are persuaded that the retest effects seen are substantial and that, if retest effects are taken as one index of plasticity, older persons continue to show a marked ability to benefit from testing experience. It is also noteworthy that this ability to benefit from retesting alone seems to apply to elderly of all levels of functioning who have participated in ADEPT research.

Ability-Specific Training and Transfer

Results of the three ability-specific training studies (figural relations, induction, attention/memory) are also consistently in favor of a position of marked plasticity. In principle, enhancement of fluid performance was demonstrated for all three dimensions of fluid intelligence. The extent of breadth and maintenance of training, however, varied.

At present, the evidence for plasticity is strongest for the ability dimension of figural relations (Figures 2 and 3). Marked, sustained, and broad-based transfer resulting from training of figural relations ability has been found in three studies. In the case of induction ability (Figure 4), our results are less impressive. Improvement due to training, while significant, is restricted to near transfer. Whether this outcome pattern is specific to induction (i.e., whether induction ability is less trainable) or a result of inadequate design (deficiencies in training program as well as lack of retest control) is unanswered. Research in progress is aimed at clarifying these possibilities. Results on the third fluid dimension, attention/memory, are incomplete. However, the available evidence (Figure 5) again indicates a very consistent pattern of performance enhancement on all near-transfer measures. Such an outcome appears to be a necessary condition for finding more broad-based transfer in subsequent analyses.

How significant, empirically and theoretically, are the findings on intervention effectiveness? A first observation on the significance of the observed training effects concerns *breadth* and *order* of transfer. Especially for figural relations, breadth of transfer is large, covering all measures contained in the transfer battery, although at the individual test level some of the transfer effects are nonsignificant. Such breadth of transfer is a remarkable outcome. Although one must be careful not to conclude that the findings point unequivocally to modification of ''latent'' abilities, the results do support the notion that training of figural relations results in enhancement of performance on a broad battery of ability tests ranging from Gf to Gc and perceptual speed. It is also noteworthy that, by and large, magnitude of transfer was ordered and in line with predictions derived from the theoretical measurement model.

A second observation regarding the significance of the observed training effects deals with their *magnitude*. On near-transfer measures (including the effects of retesting itself), the magnitude is approximately one standard deviation when assessed in terms of baseline conditions. As it turns out, this magnitude is at least as much or more than what Schaie (1979, 1982) reports as the best estimate for longitudinal age decrements in the primary mental abilities from age 60 to 81. Schaie (1982) finds values between one-half and one standard deviation when cumulating successions of seven-year longitudinal decrements to yield 21-year estimations of decline from age 60 to 81. This age range is similar to the cross-sectional age range covered in ADEPT research.

Comparing Schaie's descriptive findings on longitudinal age decline with ADEPT intervention work on the near transfer measures, then, shows that performance improvement in ADEPT is equivalent to 20-year longitudinal age decline observed in the same age range. Note that this level of improvement is reached after a relatively brief intervention: five one-hour training sessions. Thus, it appears likely that, in terms of ability indicators used in ADEPT, intellectual decline during the sixties and seventies (if it occurs) can be slowed down, halted or even reversed if proper education interventions are designed and implemented.

We realize, of course, that this conclusion is restricted in several respects. For example, it is restricted to the age cohorts studied and to fairly healthy elderly community residents able and willing to participate in training research of the type presented. Moreover, the results do not imply that older persons benefit more from cognitive training experience than younger adults. This is possible, but at present, an untested proposition. We must emphasize the role that birth cohorts may play in restricting the generalizability of the present findings. There are historical changes in health, education, and social ecology that modify not

only the level of cognitive functioning of elderly cohorts and their intellectual plasticity but also the life history of other age groups living at a particular point in time (Baltes & Schaie, 1976). Age-comparative assessment of plasticity, therefore, will need to consider the use of cohort-sequential designs if the interpretation is aimed at isolating the role of chronological age from the effect of historical change. What is shown, then, in the present data is that elderly persons of a particular birth cohort, if they are exposed and attend to performance-enhancing conditions, do benefit substantially from cognitive training; furthermore, that plasticity at the individual level in the elderly is at least as large as average aging-related decrements observed in longitudinal research based on comparable birth cohorts.

Implications: Toward an Integrative View

We began this chapter by summarizing our observations on the role of plasticity and variability in intellectual assessment. The present findings support the theoretical significance of these perspectives from an empirical point of view. We added observations on the role of ecological and performance deficits in the area of psychometric intelligence, noting that elderly persons are likely to suffer from such deficits. Again, the present findings are in concert with such a position, although they do not represent a direct test. Thus, it is likely that when reduced levels of intellectual functioning on psychometric tests are observed in older persons, these are in part due to ecological and performance deficits. This conclusion is particularly important because it applies to that dimension of psychometric intelligence, fluid abilities, that is postulated to show regular and marked aging decline.

Before we proceed to discuss further theoretical implications, it is desirable to digress somewhat to state what the findings and observations do not imply. It has become a dominant theme to view our research as bearing directly on the question of whether there is intellectual decline with aging (Baltes & Schaie, 1976; Horn & Donaldson, 1976, 1977; Schaie & Baltes, 1977). The present studies are not designed to contribute a direct answer to that question. The research is not age-comparative but deals with in-depth examination of one age group, the 1975–1980 elderly from about 60 to 80 years of age. However, in conjunction with earlier descriptive research on interindividual variability including cohort effects, the findings do speak to the need for developmental theories of psychometric intelligence that incorporate not only features of normative age change (such as is true for the bulk of Horn's writings), but also features dealing with the conditions for and range of variability and plasticity (Baltes & Baltes, 1980; Baltes & Willis, 1979a).

Performance versus Potential (Latent Reserve). This digression aside, what are the other implications? In the introduction to this chapter, we also used the distinction between abilities as indicators of performance and abilities as indicators of potential or capacity. Figure 6 summarizes our current view on this matter, preliminary as it is. The figure emphasizes first that observed *performance* is but one assessment of what individuals can do. Although we are not sure, we acknowledge that life-long intellectual performance functions in current cohorts are likely to follow an incremental-plateau-decremental trajectory. The trajectories, however, vary considerably by ability and individual (Baltes & Willis, 1979b).

How can we combine a view of plasticity of gerontological intelligence with the notion that there is aging-related decline in performance? Performance is but one level of achievement. Figure 6, therefore, also emphasizes that there is room for performance enhancement at all ages. This is illustrated in Figure 6 by graphing a second life-span function of intelligence indexing maximum *potential*. In our view, *the level of maximum potential is inherently an unkown or dialectical quantity*. As we learn about conditions for plasticity in performance, the curve for maximum potential is continuously redrawn. However, in line with general notions of biological decline (Strehler, 1977), we acknowledge that the "maximum" level of performance may be lower in old age than in adulthood. Approximation to maximum level would be measured by considering "optimal" conditions of performance such as would be true when graphing world records by age in marathon running (Fries, 1980). When applying such a testing-the-limits

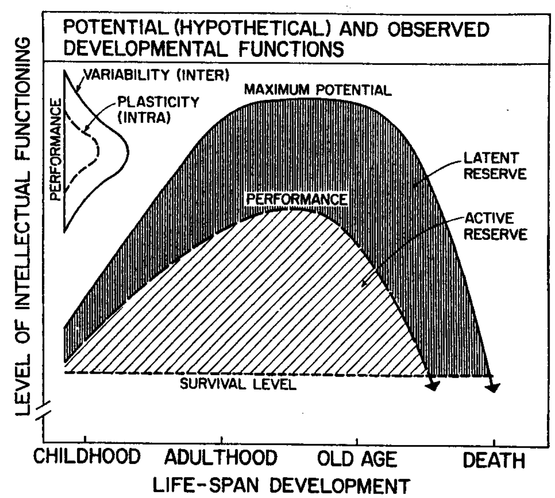

Fig. 6. Life-span development of intellectual performance: Conceptual relationships between potential, performance, and two types of reserve (active and latent).

approach to intellectual aging, we believe (as Fries found for marathon running) that the best older persons would be in the highest percentiles of young adults.

In biology and medicine, the concept of a reserve has been introduced to communicate a similar view. Organisms have reserves that can be called upon if special demands are made (Fries, 1980; Strehler & Mildvan, 1963). In adulthood, for example, the reserves for heart and lung functioning are a multiple of what is necessary during resting stages. In Figure 6, we define as *active reserve* (evidenced in performance) that portion of intellectual potential that has been put into operation. *Latent reserve,* on the other hand, includes that portion of intellectual potential that could be activated if additional energy and time were invested. For most behaviors and certainly for intellectual aging, then, we suggest that there is much latent reserve. Rarely, such as in highly trained athletes or "mental" experts, would the level of performance approximate the maximum potential. Most individuals, including the elderly, therefore, if they choose to do so or are exposed to supportive conditions, can improve their performance. Thus, even in the face of aging-related decline of maximum potential, many older persons could maintain high levels of intellectual functioning. What will be necessary for the elderly is to invest more effort in a given class of behavior than would be true for younger adults.

Figure 6 contains an additional piece of information which completes the gist of our current thinking about intellectual aging. It postulates that there are life-stage differences in the proportional relationship between latent and active reserve. In psychometric intelligence and current cohorts, we postulate a smaller difference between maximum potential and performance for the first part of life. We believe this is so because, for childhood and young adulthood, cultural-ecological conditions are aimed at optimization of intellectual performance. Beginning in adulthood and especially in old age, on the other hand, we assert that the ecology is less supportive of high performance on intelligence. This performance deficit in old age is perhaps also enhanced by intraorganismic conditions of late life that assign little priority to high-level functioning in psychometric intelligence.

Aging as Selective Optimization. In another context, Baltes and Baltes (1980; see also Brent, 1978) have proposed to use the term *selective optimization* to characterize a process of individuation with aging. They argue that, both for biological and environmental reasons, aging involves an increase in biological vulnerability associated with a decrease in environmental adaptability. However, that process (a) varies widely by individuals and (b) results in individual selective and compensatory effects.

This perspective of selective optimization fits the present data. In light of the need to invest, with aging, more and more effort into maintaining high levels of functioning, older persons may choose a strategy of specialization (Brent, 1978) or selective optimization (Baltes & Baltes, 1980) focusing on those classes of behavior judged to be most adequate to the individual's life situation. Yet, if older individuals are prompted, attend, and practice in other areas, such as psychometric intelligence, their performance can be enhanced markedly. In line with a selective optimzation hypothesis, older individuals *can* maintain or even increase their levels of intellectual functioning.

Thus, we believe that the research reported permits the establishment of another linkage, namely that between intelligence and personality. The findings of plasticity can be coordinated with a position that views aging as a process of selective optimization. When older individuals exhibit lower intellectual performance, this does not necessarily mean that they cannot achieve higher levels in principle. Certain older individuals might have opted out, either because investment is too high or because other domains of behavior are more attractive.

We acknowledge that these concluding observations are tentative and preliminary. Not all of them are supported by empirical evidence, nor is it likely that such a simple view will be a panacea. Yet, until they are falsified, we consider taking the views expressed as a conceptual guide in our future explorations into the conditions for variability and plasticity of psychometric intelligence in old age.

Summary

Penn State's Adult Development and Enrichment Project (ADEPT) consists of a series of short-term longitudinal studies aimed at examining the extent of and conditions for modifiability (plasticity) of intellectual performance in old age. Study designs include consideration of theory-guided transfer to other abilities (tests) and maintenance of training at several post-training occasions (1 week, 1 month, 6 months). The measurement focus of ADEPT is on Cattell and Horn's model of fluid-crystallized intelligence.

Targets of intervention are three abilities, all indexing fluid intelligence: figural relations, induction, and attention. Conditions for studying modifiability in the sense of enhancement include manipulation of test familiarity (e.g., retest experience) and of ability-specific problem-solving skills by means of educational training. A progress report summarizing a total of five relevant studies is presented.

In general, modifiability of intellectual performance in the target abilities selected for enhancement is substantial. Training also results in a fair amount of transfer to other tests (perhaps abilities) and the skills are maintained over at least one month following training. The overall results suggest (a) the existence of much latent potential (reserve) for intellectual performance in the elderly, (b) the need for articulation of performance theories, and (c) the general notion of emphasizing knowledge about interindividual variability and intraindividual plasticity (including lack thereof) as a critical feature of any theory dealing with intellectual aging. The findings are seen also as convergent with a theory of aging as selective optimization.

ACKNOWLEDGMENTS

Many thanks are due to several colleagues, Paul A. Games and John R. Nesselroade; research assistants Rosemary Blieszner, Frederic Blow, Steven W. Cornelius, Brian F. Hofland, Marjorie E. Lachman, and Avron Spiro, III; and support staff, Rosalie K. Ammerman, Carolyn S. Nesselroade, and Myrtle A. Williams, who collaborated in many phases of the research program. We acknowledge also the helpful comments by Margret M. Baltes, Steven W. Cornelius, and Elizabeth Loftus on an earlier draft of this manuscript.

References

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press, 1968.

Baltes, P. B., & Labouvie, G. V. Adult development of intellectual performance: Description, explanation, and modification. In C. Eisdorfer & M.P. Lawton (Eds.), *The psychology of adult development and aging.* Washington, D. C.: American Psychological Association, 1973.

Baltes, P. B., & Schaie, K. W. On the plasticity of intelligence in adulthood and old age: Where Horn and Donaldson fail. *American Psychologist,* 1976, *31,* 720–725.

Baltes, P. B., & Willis, S. L. Toward psychological theories of aging and development. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging.* New York: Van Nostrand Reinhold, 1977.

Baltes, P. B., & Willis, S. L. The critical importance of appropriate methodology in the study of aging: The sample case of psychometric intelligence. In F. Hoffmeister & C. Müller (Eds.), *Brain function in old age.* Heidelberg: Springer, 1979. (a)

Baltes, P. B., & Willis, S. L. Life-span developmental psychology, cognition, and social policy. In M. W. Riley (Ed.), *Aging from birth to death.* Boulder: Westview, 1979. (b)

Baltes, P. B., & Danish, S. J. Intervention in life-span development and aging: Issues and concepts. In R. R. Turner & H. W. Reese (Eds.), *Life-span developmental psychology: Intervention.* New York: Academic Press, 1980.

Baltes, P. B., & Baltes, M. M. Plasticity and variability in psychological aging: Methodological and theoretical issues. In G. Gurski (Ed.), *Determining the effects of aging on the central nervous system.* Berlin: Schering, 1980.

Baltes, P. B., Nesselroade, J. R., & Cornelius, S. W. Multivariate antecedents of structural change in development: A simulation of cumulative environmental patterns. *Multivariate Behavioral Research,* 1978, *13,* 127–152.

Baltes, P. B., Cornelius, S. W., Spiro, A., III, Nesselroade, J. R., & Willis, S. L. Integration vs. differentiation of fluid-crystallized intelligence in old age. *Developmental Psychology,* 1980, *16,* 625–635.

Barton, E. M., Baltes, M. M., & Orzech, M. J. Etiology of dependence in older nursing home residents during morning care: The role of staff behavior. *Journal of Personality and Social Psychology*, 1980, *38*, 423–431.

Beilin, H. Constructing cognitive operations linguistically. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 11). New York: Academic Press, 1976.

Bijou, S. W. Environment and intelligence: A behavioral analysis. In R. Cancro (Ed.), *Contributions to intelligence*. New York: Grune & Stratton, 1971.

Blieszner, R., Willis, S. L., & Baltes, P. B. Training research on induction ability in aging: A short-term longitudinal study. *Journal of Applied Developmental Psychology*, 1981, *2*, 247–265.

Botwinick, J. Aging and intelligence. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*. New York: Van Nostrand Reinhold, 1977.

Botwinick, J., & Storandt, M. *Memory, related functions and age*. Springfield: Charles C Thomas, 1974.

Brandstädter, J., & Schneewind, K. A. Optimal human development: Some implications for psychology. *Human Development*, 1977, *20*, 48–64.

Brent, S. B. Individual specialization, collective adaptation and rate of environmental change. *Human Development*, 1978, *21*, 21–33.

Brown, A., & French, L. The zone of potential development: Implications for intelligence testing in the year 2000. *Intelligence*, 1979, *3*, 255–277.

Cattell, R. B. *Abilities: Their structure, growth and action*. Boston: Houghton, 1971.

Cattell, R. B., & Cattell, A. K. S. *Test of "g": Culture Fair (Scale 2, Form A)*, Champaign, Ill.: Institute for Personality and Ability Testing, 1957.

Cattell, R. B., & Cattell, A. K. S. *Measuring intelligence with the Culture Fair tests: Manual for Scales 2 and 3*. Champaign, Ill.: Institute for Personality and Ability Testing, 1961.

Cattell, R. B., & Cattell, A. K. S. *Test of "g": Culture Fair (Scale 3, Form A, 1963 Edition; Form B, 1961 Edition, Second)*. Champaign, Ill: Institute for Personality and Ability Testing, 1963.

Clayton, V. P., & Birren, J. E. The development of wisdom across the life span: A reexamination of an ancient topic. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 3). New York: Academic Press, 1980.

Cornelius, S. W., Willis, S. L., Nesselroade, J. R., & Baltes, P. B. *Convergence between attention variables and factors of psychometric intelligence in older adults*. Manuscript submitted, College of Human Development, The Pennsylvania State University, 1982.

Craik, F. I. M. Age differences in human memory. In J. E. Birren & K. W. Schaie (Ed.), *Handbook of the psychology of aging*. New York: Van Nostrand, 1977.

Cunningham W. R. Age comparative factor analysis of ability variables in adulthoood and old age. *Intelligence*, 1980, in press.

Denney, N. W. Problem solving in later adulthood: Intervention research. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 2). New York: Academic Press, 1979.

Düker, H., & Lienert, G. A. *Konzentrations-Leistungstest*. Göttingen: Hogrefe, 1959.

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. *Kit of factor-referenced cognitive tests* (1976 Rev.). Princeton, N.J.: Educational Testing Service, 1976.

Flavell, J. H., & Wohlwill, J. F. Formal and functional aspects of cognitive development. In D. Elkind & J. H. Flavell (Eds.), *Studies in cognitive development: Essays in honor of Jean Piaget*. New York: Oxford University Press, 1969.

Fries, J. F. *The plasticity of aging*. Unpublished manuscript, Department of Medicine, Stanford University, Stanford, California, 1980.

Glaser, R., & Resnick, L. Instructional psychology. In P. H. Mussen & M. Rosenweig (Eds.), *Annual review of psychology*. Palo Alto, Calif.: Annual Reviews, 1972.

Greene, E. B. *Measurement of human behavior*. New York: Odyssey, 1941.

Guilford, J. P. *Verbal analogies test, I*. Beverly Hills, Calif.: Sheridan Psychological Services, 1969. (a)

Guilford, J. P. *Word matrix test*. Beverly Hills, Calif.: Sheridan Psychological Services, 1969. (b)

Guthke, V. J. Entwicklungsstand und Probleme der Lernfähigkeitsdiagnostik. *Zeitschrift für Psychologie*, 1976, *184*, 215–239.

Havighurst, R. J. *Developmental tasks and education*. New York: McKay, 1948.

Hofland, B. F., Willis, S. L., & Baltes, P. B. Fluid intelligence performance in the elderly: Retesting and conditions of assessment. *Journal of Educational Psychology*, 1981, *73*, 573–586.

Horn, J. L. *Social situations—EP03A* (Unpublished test). Denver: University of Denver, Department of Psychology, 1967.

Horn, J. L. Organization of data on life-span development of human abilities. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory*. New York: Academic Press, 1970.

Horn, J. L. Psychometric studies of aging and intelligence. In S. Gershon & A. Raskin (Eds.), *Geriatric psychopharmacology: The scene today*. New York: Raven, 1975.

Horn, J. L. Human ability systems. In P. B. Baltes (Ed.), *Life-span development and behavior* (Vol. 1). New York: Academic Press, 1978.

Horn, J. L., & Cattell, R. B. Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 1966, *57*, 253–270.

Horn, J. L., & Cattell, R. B. Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 1967, *26*, 107–129.

Horn, J. L., & Donaldson, G. On the myth of intellectual decline in adulthood. *American Psychologist*, 1976, *31*, 701–719.

Horn, J. L., & Donaldson, G. Faith is not enough: A response to the Baltes–Schaie claim that intelligence does not wane. *American Psychologist*, 1977, *32*, 369–373.

Horn, J. L., Donaldson, G., & Engstrom, R. Apprehension, memory, and fluid intelligence decline in adulthood. *Research on Aging*, 1981, *3*, 33–84.

Hoyer, W. J., Labouvie, G. V., & Baltes, P. B. Modification of response speed deficits and intellectual performance in the elderly. *Human Development*, 1973, *16*, 233–242.

Hunt, E. B. Mechanics of verbal ability. *Psychological Review*, 1978, *85*, 109–130.

Hunt, E. B., Frost, N., & Lunneborg, C. L. Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), *Advances in learning and motivation* (Vol. 7). New York: Academic Press, 1973.

Labouvie-Vief, G. Toward optimizing cognitive competence. *Educational Gerontology*, 1976, *1*, 75–92.

Labouvie-Vief, G. Beyond formal operations: Uses and limits of pure logic in life-span development. *Human Development*, 1980, *23*, 141–161.

Labouvie-Vief, G., & Chandler, M. Cognitive development and life-span developmental theories: Idealistic versus contextual perspectives. In P. B. Baltes (Ed.), *Life-span development and behavior* (Vol. 1). New York: Academic Press, 1978.

Labouvie, G. V., Hoyer, W. J., Baltes, P. B., & Baltes, M. M. Operant analysis of intellectual behavior in old age. *Human Development*, 1974, *17*, 259–272.

Lachman, M. E., Baltes, P. B., Nesselroade, J. R., & Willis, S. L. *Cognitive training research with the elderly: Ethical issues and transfer to personality*. Manuscript submitted, College of Human Development, The Pennsylvania State University, 1982.

Lawton, M. P., & Nehemow, L. Ecology and the aging process. In C. Eisdorfer & M. P. Lawton (Eds.), *The psychology of adult development and aging*. Washington, D.C.: American Psychological Association, 1973.

O'Sullivan, M., & Guilford, J. P. *Social translations, Form A*. Beverly Hills, Calif.: Sheridan Psychological Services, 1965.

Pike, L. W. *Short-term instruction, testwiseness and the scholastic aptitude test: A literature review with research recommendations*. Princeton, N.J.: Educational Testing Service, 1978.

Plemons, J. K., Willis, S. L., & Baltes, P. B. Modifiability of fluid intelligence in aging: A short-term longitudinal training approach. *Journal of Gerontology*, 1978, *33*, 224–231.

Posner, M. I., & Mitchell, R. A chronometric analysis of classification. *Psychological Review*, 1967, *74*, 392–409.

Rabbitt, P. Changes in problem solving in old age. In J. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*. New York: Van Nostrand Reinhold, 1977.

Raven, J. C. *Advanced progressive matrices, Set II* (1962 revision). London: Lewis, 1962.

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

Reinert, G. Educational psychology in the context of the human life span. In P. B. Baltes & O. G. Brim (Eds.), *Life-span development and behavior* (Vol. 3). New York: Academic Press, 1980.

Resnick, L. B. (Ed.). *The nature of intelligence*. Hillsdale, N.J.: Lawrence Erlbaum, 1976.

Riegel, K. F. The dialectics of human development. *American Psychologist*, 1976, *31*, 689–700.

Schaie, K. W. A reinterpretation of age-related changes in cognitive structure and functioning. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory*. New York: Academic Press, 1970.

Schaie, K. W. Toward a stage theory of adult cognitive development. *Journal of Aging and Human Development*, 1977–78, *8*, 129–138.

Schaie, K. W. The primary mental abilities in adulthood: An exploration in the development of psychometric intelligence. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 2). New York: Academic Press, 1979.

Schaie, K. W. The Seattle Longitudinal Study: A twenty-one year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development*. New York: Guilford, 1982.

Schaie, K. W., & Baltes, P. B. Some faith helps to see the forest: A final comment on the Horn and Donaldson myth of the Baltes–Schaie position on adult intelligence. *American Psychologist*, 1977, *32*, 1118–1120.

Schaie, K. W., & Willis, S. L. Life-span development: Implications for education. *Review of Educational Research*, 1978, *6*, 120–156.

Schmidt, L. R. Testing the limits im Leistungsverhalten: Möglichkeiten und Grenzen. In E. Duhm (Ed.), *Praxis der klinischen Psychologie* (Vol. 2). Göttingen: Hogrefe, 1971.

Sörbom, D., & Jöreskog, K. G. *COFAMM: Confirmatory factor analysis with model modification*. Chicago: National Educational Resources, 1976.

Sternberg, R. J., & Detterman, D. K. (Eds.). *Human intelligence: Perspectives on its theory and measurement*. Norwood, N.J.: Ablex, 1979.

Sterns, H. L., & Sanders, R. E. Training and education of the elderly. In R. R. Turner & H. W. Reese (Eds.), *Life-span developmental psychology: Intervention*. New York: Academic Press, 1980.

Strehler, B. L. *Time, cells, and aging*. New York: Academic Press, 1977.

Strehler, B. L., & Mildvan, A. S. General theory of mortality and aging. *Science*, 1960, *132*, 14–21.

Stroop, J. R. Studies in interference in serial verbal interactions. *Journal of Experimental Psychology*, 1935, *18*, 643–661.

Thurstone, T. G. *Primary mental abilities, grades 9–12, 1962 Revision*. Chicago: Science Research Associate, 1962.

Underwood, G. *Attention and memory* (1st ed.). New York: Pergamon Press, 1975.

Vernon, P. E. Practice and coaching effects in intelligence tests. *Educational Forum*, 1954, *18*, 209–280.

Wickens, D. Encoding categories of words: An empirical approach to meaning. *Psychological Review*, 1970, *77*, 1–15.

Willis, S. L., & Baltes, P. B. Intelligence and cognitive ability. In L. Poon (Ed.), *Aging in the 1980's: Psychological issues*. Washington, D.C.: American Psychological Association, 1980.

Willis, S. L., Blieszner, R., & Baltes, P. B. *Multi-stage training on the fluid ability of figural relations in the aged*. Manuscript submitted, College of Human Development, The Pennsylvania State University, 1980.

Willis, S. L., Blieszner, R., & Baltes, P. B. Training research in aging: Modification of performance on the fluid ability of figural relations. *Journal of Educational Psychology*, 1981, *73*, 41–50.

Willis, S. L., Cornelius, S. W., Blow, F., & Baltes, P. B. Training research in aging: Attentional processes. *Journal of Educational Psychology*, 1982.

Wing, H. Practice effects with traditional mental test items. *Applied Psychological Measurement*, 1980, *4*, 141–155.

Wittrock, M. C. Verbal stimuli in concept formation: Learning by discovery. *Journal of Educational Psychology*, 1963, *54*, 183–190.

Wittrock, M. C. The learning by discovery hypothesis. In L. S. Schulman & E. R. Keisler (Eds.), *Learning by discovering: A critical appraisal*. Chicago: Rand McNally, 1966.

Wohlwill, J. *The study of behavioral development*. New York: Academic Press, 1973.