In K. W. Schaie (Ed.), Annual Review of Gerontology and Geriatrics, Vol. 7. New York: Springer, 1987

CHAPTER 7

Cognitive Training and Everyday Competence

Sherry L. Willis
Department of Individual and Family Studies
The Pennsylvania State University
University Park, Pennsylvania

This chapter will be divided into three major sections, with each section focusing on an issue associated with cognitive training research with the elderly. In the first section, we examine what are some major questions or goals addressed by cognitive training research and consider the appropriateness of age-comparative research designs in addressing these questions. In the second section, we discuss the implications of cognitive training for enhancing competence in real-life tasks. In the third section, the findings of cognitive training are related to current models of intellectual aging.

MAJOR QUESTIONS ADDRESSED IN COGNITIVE TRAINING RESEARCH

The implicit goal of most cognitive training research has been to increase our understanding of intellectual aging (Denney, 1982; Sterns & Sanders, 1980; Willis, in press). Much of the training literature has focused on three major questions regarding intellectual aging:

1. To what extent can the elderly's cognitive performance be modified (improved), and what is the range of individual differences in the magnitude of improvement?
2. What types of experimental manipulations are effective in modifying the elderly's cognitive performance?
3. What aspects of cognitive performance are acutally modified via cognitive intervention?

Preparation of this chapter was supported by research grants AG03544 and AG05304 from the National Institute on Aging.

Before discussing each of these questions in more detail, one particular limitation of the current training literature should be noted. Virtually all training research has focused on cognitive domains (e.g., memory, problem solving, fluid intelligence) that show the largest age differences or greatest age-related decline. Although targeting training efforts on deficiencies may at first appear reasonable, it remains, nevertheless, that little is known regarding the potential for further enhancement of cognitive domains (e.g., verbal ability) that exhibit little normative decline and thus are considered well-functioning in many elderly (Baltes, Dittman-Kohli, & Dixon, 1984).

This dearth of cognitive training research on normatively stable abilities is problematic both in regard to the development of theories of intellectual aging and in regard to applied concerns. From a theoretical perspective, behavioral plasticity is an issue that needs to be examined with regard to stable as well as declining areas of cognitive functioning in order to understand across cognitive domains the range of variability that can be produced through experimental manipulation (Kliegl & Baltes, in press). Examination of plasticity in well-functioning cognitive domains may be particularly important to understanding areas of growth and development unique to old age.

From an applied perspective, a singular focus on remediation or compensation for deficits may result in a limited and unbalanced approach to behavioral intervention. Some approaches to rehabilitation argue for emphasizing the individual's "strengths," rather than the "weaknesses." For example, verbal ability has been found in several studies to be a significant predictor of working memory in old age; in particular, verbal ability has been related to use of mnemonic strategies (Poon, 1985). However, enhancement of verbal ability as an intervention procedure has received little attention in memory training research (Poon, Fozard, Cermak, Arenberg, & Thompson, 1980).

Examining the Range of Plasticity/Variability in Training Improvement

The most frequently addressed question in training research focuses on the extent to which intervention procedures are effective in improving the performance of subjects receiving training. What was the magnitude of improvement associated with training? This issue has sometimes been discussed in terms of developmental plasticity (Baltes & Baltes, 1980; Baltes & Willis, 1982). What is the range of plasticity in the individual's cognitive performance that can be elicited via experimental manipulation? Plasticity has typically been measured by comparing the individual's normal (unexercised) performance level with his/her optimal performance level, as assessed following some type of training or practice (Denney, 1982).

It is important to note that plasticity is concerned primarily with *intra-*

individual change. The emphasis is on the range or magnitude of change that can be observed within the *same* individual when assessed at two or more time points (e.g., pretest vs. posttest) or under multiple experimental conditions. If longitudinal data are available, plasticity may be assessed in terms of the individual's performance at some earlier point in development (Schaie & Willis, 1986). For example, the researcher may examine whether the magnitude of training effects was such that the individual's level of performance was remediated to the level demonstrated at an earlier developmental age (e.g., middle adulthood); note that the focus is still on intraindividual comparisons.

The researcher is typically also interested in *individual differences* in intraindividual change associated with training. The range of variability may be examined within the same group before and after treatment. For example, does training alter the range of individual differences within the treatment group? Alternatively, the focus may be on *between-group* comparisons of variability. Does the range of variability differ for various groups exposed to the same or different treatments?

In cognitive training research, plasticity and variability have been measured in three ways: mean level of performance, change scores, and overlap in distributions. We will discuss some of the conditions and limitations associated with each procedure.

Mean Level. The most commonly used procedure for evaluating training effects has been mean level of performance. Reporting effects in terms of mean level is most appropriate when the focus is on intraindividual change; the same persons' mean performance scores prior to and after training are compared.

In many training studies, however, the major comparison focuses on individual differences in intraindividual change. For example, the mean scores of the treatment group and a control/comparison group are compared at posttest. If the level of performance after training is to be compared across groups, then the various groups must have a comparable performance baseline at some point prior to training, or group differences must be statistically adjusted (Campbell & Stanley, 1963). The pretest score can serve as this common baseline. The pretest score is used as the common baseline most frequently when comparing groups of the same age cohort. Subjects are assigned to treatment conditions via random assignment or matching procedures so that the mean pretest scores for the groups are comparable. In age comparative research, the assumption is made that the age cohorts performed at comparable levels at a previous developmental period; thus, a common performance baseline is assumed to have occurred in young adulthood. In a later section of this chapter we will discuss why we consider such assumptions problematic.

Change Scores. Computation of change scores is another approach to assessing magnitude of training gain. The difference between the posttest and pretest scores becomes the dependent variable. Change scores have one advantage over raw total scores in that they do not reflect the possible pretest differences among treatment/comparison groups in level of performance. The question becomes, Does the amount of pre–posttest training improvement differ between treatment conditions (or comparison groups)?

There has been considerable controversy over the use of change scores (Cronbach & Furby, 1970; Harris, 1963). However, several developmental methodologists (Campbell & Stanley, 1963; Nesselroade, Stigler, & Baltes, 1980; Nunnally, 1982) have strongly supported the use of change scores in developmental and experimental studies of change as the most direct procedure for assessing change. Gain scores have often been criticized for their unreliability. However, the reliability of change scores is a function of the reliability of the measures from which the change scores were derived. If the reliability of the pre–posttest measures is good, error of measurement problems would be expected to affect interpretation of results only minimally.

Overlap in Distributions. A third approach to assessing training improvement is to examine the amount of overlap in the distribution of scores (Kerlinger, 1973). Statistical tests used to assess training effects when mean scores or change scores are the dependent variables are based on sample size and variability. To understand the actual magnitude of intraindividual change or interindividual differences associated with training, it is useful to examine the degree to which score distributions converge or diverge across occasions or groups.

If variability in intraindividual change is the major question, then the proportion of overlap between distribution of scores at pretest versus posttest for the same group can be examined. If interindividual differences are the major concern, then the proportion of overlap in distribution of posttest scores for various groups can be examined.

Age-Comparative Studies of the Magnitude of Training Effects

The focus of a number of training studies has been on comparing the magnitude of training effects for old versus young adults (e.g., Bellucci & Hoyer, 1975; Le Breck & Baron, 1987; Taub & Long, 1972). It has been argued that a younger comparison group is required in gerontological training studies in order to address the issue of remediation of age-related decline (Denney, 1982). Interpretation of age-comparative training studies has been based on a number of assumptions regarding the course of adult intellectual development. Age-comparative training studies have typically employed mean scores or change scores as dependent variables. We find problematic both the assumptions underlying age-comparative training research and the procedures used to compare training effects.

Several researchers have argued that training effects should be greater for older age cohorts than for younger age cohorts. This hypothesis of greater

training effects for the elderly is based on several assumptions that longitudinal research findings indicate are not tenable (Schaie, 1983). First, it is assumed that the elderly were at a comparable performance level (baseline) as the young when at the same chronological age. Second, it is assumed that virtually all elderly in the sample have suffered age-related decline on the target variable. Third, it is assumed that the younger age group is functioning nearer to its peak level of functioning; thus, the magnitude of training improvement should be less for the young. Given these assumptions, it is argued that training gain should be greater for the older group.

Cohort-sequential longitudinal research findings, however, indicate that for virtually all of the cognitive abilities studied, older cohorts performed at a significantly lower level when at the same age than the younger comparison group (see Cunningham, this volume; Schaie, 1983). The assumption of equivalent performance baselines for different cohorts when at the same chronological age is not tenable.

Figure 7-1 illustrates baseline performance (1970, 1977) and training effects (1984-Pre, 1984-Pt) for the 1910 and 1917 birth cohorts, trained on Spatial Orientation ability (Schaie & Willis, 1985). Note that there are significant differences in the mean performance level of the two cohorts when compared at the baseline ages of 60 and 67, prior to training (i.e., 1970 and 1977 for 1910 cohort; 1977 and 1984-Pre for 1917 cohort). The mean score for the 1910 cohort is significantly below that for the 1917 cohort at both ages. Significant training improvement (1984-Pt) is demonstrated for both age cohorts. That is, intraindividual training gain is demonstrated by the finding that the 1917 cohort's posttraining mean score is above its score 14 years previously, and the posttraining performance of the 1910 cohort is at the mean level of their performance in 1970. Training was effective in remediating, on average, the age-related decline for the 1910 and 1917 cohorts; however, the initial cohort difference in level of performance remains. In order for mean posttest scores to be comparable for the two age cohorts, as argued in age-comparative research, training would have to modify not only age-related decline but also initial cohort differences.

In cross-sectional age-comparative training studies, the researcher has available only the pre–posttest scores from Figure 7-1. The researcher focuses on the age differences in posttest mean scores, and, confounding age and cohort in his/her interpretation, incorrectly concludes that training does not eliminate the "age deficit."

The second faulty assumption often made in age-comparative training research further complicates interpretation of findings. It is assumed that most elderly in the sample have experienced reliable age-related decline on the target variable. However, longitudinal research indicates that there are wide individual differences in the timing and rate of age-related decline. For example, older participants (age, 64–95; $N = 229$) from the Seattle Longitudinal Study were classified as having remained stable or having reliably declined over the previous

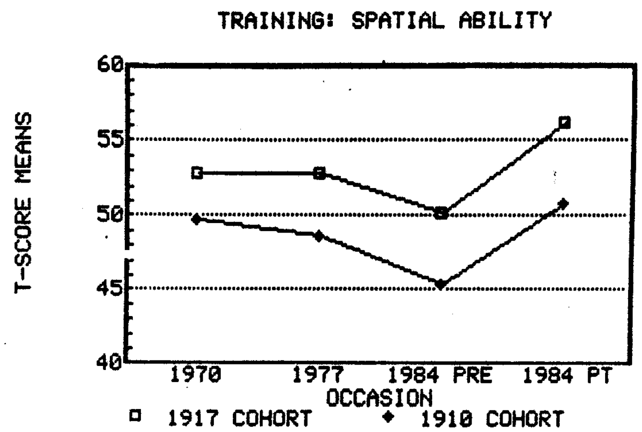

Figure 7-1. PMA Spatial Orientation T-score means for the 1910 and 1917 birth cohorts at four occasions: 1970, 1977, 1984 Pretest, 1984 Posttest.

14-year period on two abilities (Inductive Reasoning, Spatial Orientation) that show early normative patterns of decline (Schaie & Willis, 1986; Willis & Schaie, 1986a). Only 21.8% of the subjects had declined on both abilities; 15% had declined only on Inductive Reasoning; 16% had declined on Spatial Orientation; 46.7% of the sample had not declined on either ability.

These data should not be interpreted as evidence against the reality of age-related decline, but are only shown to demonstrate that there are wide individual differences in the timing and rate of decline. If training researchers selected subjects 80 years of age and older, then they might accurately assume that most of their subjects had experienced at least some age-related decline. However, most older persons included in training studies thus far have been the young-old, that age period in which there are the widest individual differences in patterns of decline (Willis, 1985). Without prior longitudinal data the researcher cannot accurately assume that all of the young-old subjects participating in psychological research have experienced substantial age-related decline.

Figure 7-2 illustrates the complexities of interpreting training effects in older adults (Willis & Schaie, in press). The top figure presents data for a total group of 118 older subjects trained on Spatial Orientation ability. On the left-hand side of the figure, the magnitude of age-related decline over a 14-year period (1970–1984) prior to training is shown separately for men and women. On the right side of the figure, the magnitude of pre–posttest training improvement on spatial ability is shown. It appears that, on average, the magnitude of training gain for men is comparable to the magnitude of age-related decline, and training gain for women exceeds the magnitude of age-related decline. However, an inspection of

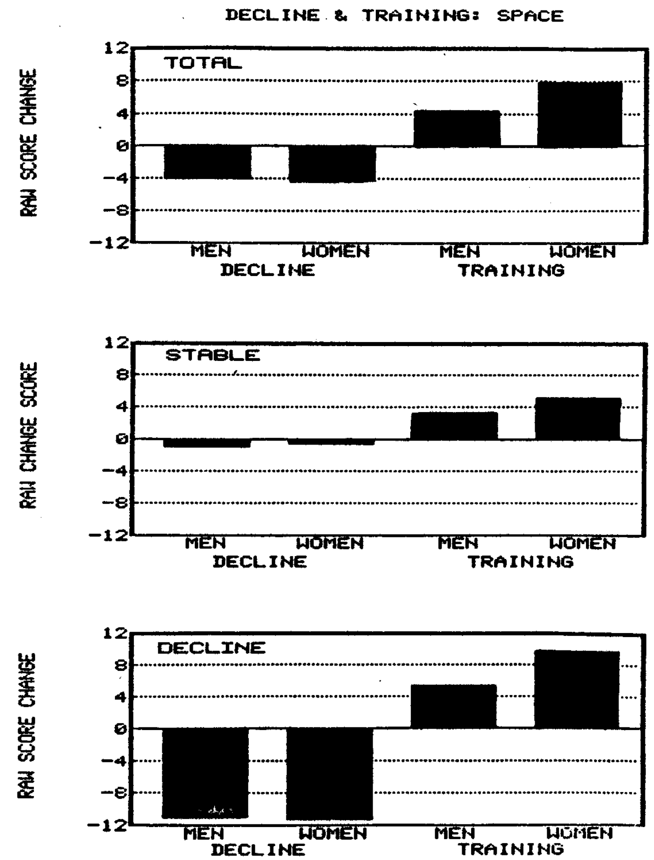

Figure 7-2. PMA Spatial Orientation raw change score means for total group (top figure), stable subjects (middle figure), and decline subjects (lower figure). In each figure the 1970–1984 mean raw change scores are shown on the left side. The pre–posttest mean raw change scores are on the right side.

the middle and lower figures, showing magnitude of age-related decline and training gain separately for stable (N = 51) and decline (N = 67) subjects, indicates that the interpretation given above is too simplistic. Significant training improvement occurred for both stable and decline subjects; however, there are differences in the magnitude and qualitative nature of training improvement for the two groups.

Separate examination of training gains for the stable and decline groups clarifies interpretation of effects for the total group (top figure). It becomes evident that the sizable training gain for the total group of women represents a confound of remediation of age-related loss for decliners and training improvement above prior baseline performance for stables. The investigator involved in age-comparative training research cannot accurately assume that all older adults in the sample have experienced prior age-related decline, and thus that training improvement represents solely the remediation of age-related loss. The magnitude of training gain in older samples will depend partially on the proportion of the sample represented in the stable versus decliner groups; hence, predictions of the relative magnitude of training gain for young versus old are highly problematic in the absence of longitudinal data.

Contrary to the hypotheses of some age-comparative researchers, findings from several studies indicated that the age × treatment interaction was not significant; that is, the magnitude of training gain for old and young did not differ significantly (Beres & Baron, 1981; Coyne, 1981; Erber, 1976; Grant, Storandt, & Botwinick, 1978; Le Breck & Baron, 1987; Taub & Long, 1972). While the decline group exhibited somewhat greater training gain than the stable group, the mean scores of the decliners at posttest will be lower than the mean scores of the stable group, since training for decliners represented partial remediation of prior age-related loss, while training gain for stables represents significant improvement above their prior baseline (1970) level of performance.

Age comparisons of training improvement in terms of level of performance at posttest (e.g., mean scores) is particularly deceiving, given cohort differences in performance level at comparable ages. Matching performance levels of older and younger age cohorts immediately prior to training in order to achieve equivalent baselines of performance is problematic in that such matching procedures may result in the less advantaged subset of the younger age cohort being compared with the more advantaged subset of the older age cohort.

One often neglected but useful approach to age comparisons of training effects involves examining the proportion of overlap in the distribution of scores for the two age groups (Kerlinger, 1973). For example, although there was a significant difference in the mean posttest scores of the 1910 and 1917 cohorts (Figure 7-1), there is considerable overlap in the distribution of posttest scores for the two groups. Figure 7-3 illustrates the overlap in posttest scores for the 1910 and 1917 cohorts trained on Spatial Orientation (Schaie & Willis, 1986). (Scores are reported in T-score points, scaled to the total group's pretest score.) In comparing overlap in distributions, one must consider the proportion of scores in the lower-scoring group (i.e., 1910 cohort) that fall below the distribution of the higher-scoring group (i.e., 1917 cohort), and also the proportion of scores in the higher-scoring group that lie above the distribution of the lower-scoring group. Approximately 6.4% ($N = 2$) of the scores for the 1910 cohort fell below the distribution of the 1917 cohort. There was complete overlap in the distribu-

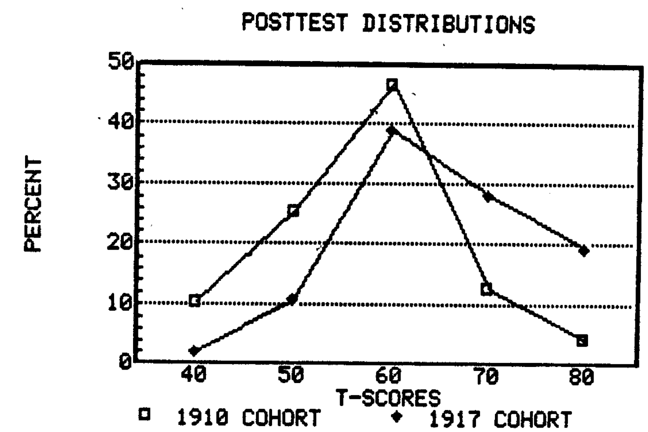

Figure 7-3. PMA Spatial Orientation score distribution (in T-scores) at posttest for 1910 and 1917 birth cohorts.

tions at the upper end (1910 Score Range: 33–78; 1917 Score Range: 39–78). More significantly, there is a 93.6% overlap in score distributions for the two cohorts. Thus, the majority of subjects in the 1910 cohort are performing within the same range as the 1917 cohort at posttest. In this example, we have compared two relatively close birth cohorts. When more disparate cohorts are compared, cohort differences will increase, and the overlap in distributions at pre- and posttest will, therefore, be smaller.

In summary, what are the implications for both the assessment and interpretation of age-comparative training research? First, there are serious limitations in the extent to which age-comparative training research can inform us about the process of age-related change in cognitive functioning. Assumptions that a younger age cohort can be used as a proxy for the older age cohort at an earlier developmental period appear untenable. Cohort differences in level of performance have been reported for many cognitive abilities when cohorts are compared at the same chronological age. Assumptions regarding comparable levels of performance for the two cohorts when in young adulthood are consequently violated. The use of mean scores to assess age differences in training effects is not appropriate when the focus is on the modifiability of age-related change. Cohort differences in mean scores may persist although substantial intraindividual training gain has occured.

Change scores are also not useful in understanding age-related change, given that a comparable performance baseline cannot be assumed to exist across cohorts. However, change scores can be employed to answer questions regarding age differences in the range of plasticity associated with training. For example, the researcher may ask whether there are age differences in the magnitude of training gain, as measured by change scores, or the treatment × age interaction. If age differences in change scores are nonsignificant, then it may be argued that the range of plasticity associated with training does not differ by age cohort (Beres & Baron, 1981; Coyne, 1981; Grant, Storandt & Botwinick, 1978). Older adults are capable of improving the same amount, on average, as younger age cohorts, even though age differences in level of performance may remain. Change scores may also be used to assess age differences in the amounts and types of training intervention that may be required to achieve comparable magnitudes of training gains. For example, simple practice experience (without feedback regarding response correctness) may be sufficient to achieve a certain mean gain score for younger age cohorts, while for older age cohorts an intervention involving strategy training might be required to achieve a comparable mean gain score.

Examination of the overlap by age in score distributions may also not be very useful in addressing the age-related change issue in training research, given that researchers will rarely have longitudinal data that could be used to determine which of the elderly subjects have experienced age-related decline. However, examination of overlap will inform the researcher about the proportion of old and young subjects that are functioning within the same performance range following training (Beres & Baron, 1981). Although cohort differences and unremediated age-related decline may result in mean score differences between young and old, there will still be considerable overlap in the distributions. This type of analysis is also useful if a valid external criterion exists. For example, if a job requires a certain level of word processing proficiency, then examination of overlap in age distributions of word processing scores following training would inform the researcher of the proportion of old and young reaching the desired criterion.

Experimental Conditions Associated with Training Effects

A second major question addressed in training research deals with the types of experimental manipulations (training procedures) that are effective in improving cognitive functioning in the elderly. Just as the cognitive abilities targeted for training research have been those showing early age-related declines, so the training procedures employed have involved the cognitive processes and behaviors believed to be associated with age deficits. For example, Scogin, Storandt, and Lott (1985) gave training on four mnemonic strategies (method of loci, novel interacting images, categorization, chunking) often found deficient in the memory performance of older adults. Mergler and Hoyer (1981) examined the effectiveness of social praise in training older adults on classification skills, given that many older adults experience loss of self-confidence and negative

stereotypes of intellectual aging. Although there has been research combining or comparing pharmacological and psychological intervention procedures, the majority of training studies conducted in the social sciences have focused solely on behavioral, educative procedures (Yesavage, Westphal, & Rush, 1981; Rosenberg, Greenwald, & Davis, 1983). The underlying assumption has been that at least part of age differences and/or age-related decline in cognitive functioning is associated with experiential deficits in the elderly's environment.

Types of experimental manipulations have varied along at least two dimensions. First, researchers have varied the degree of *intensity* of the experimental treatment. By intensity, we mean the amount of experience (e.g., number of training sessions, number of problems attempted) the subject received during training and whether the intervention involved procedures such as feedback and modeling/instruction on specific cognitive strategies.

The second dimension focuses on *cognitively oriented* intervention procedures versus *noncognitively oriented* procedures. That is, intervention may focus on cognitive strategies and behaviors specific to the ability to be trained (e.g., use of mnemonic strategies in short-term memory). In contrast, intervention may focus on noncognitive behaviors and variables (e.g., social praise, anxiety reduction, extrinsic reinforcement) that are not specific to a particular mental ability or skill but that are known to affect performance on measures of the target ability; these have been called performance variables.

Intensity of Training Procedures. The elderly's performance on a number of cognitive abilities and skills has been shown to improve with even very limited exposure to intervention conditions. Thus, the major question is the relative effectiveness of various training procedures. A low-intensity "practice only" condition, in which the subject practices on a set of problems related to the target ability but with no feedback regarding correctness of response, has been compared with a "practice with feedback" condition. In turn, practice conditions have been compared with training conditions involving modeling or instruction on specific cognitive strategies, as well as practice and feedback. With some exceptions, more intensive conditions involving modeling/instruction of cognitive strategies and feedback have resulted in greater training improvement than have less intensive practice-only conditions. It should be noted that the relative effectiveness of various training procedures is often difficult to evaluate because intervention procedures have varied with the specific cognitive ability/skill being studied, and there have been few replication studies. We will briefly review some of the training studies that have compared intervention procedures varying in intensity.

Hornblum and Overton (1976) and Schultz and Hoyer (1976) compared the relative effectiveness of practice-with-feedback versus practice-with-no-feedback conditions on training in area conservation and spatial egocentrism, respectively. Practice with feedback was found to be more effective in both studies. In contrast, Hoyer, Labouvie, and Baltes (1973) found no difference in posttest performance between a reinforced practice-with-feedback and a nonreinforced practice-with-no-feedback condition when perceptual speed was the target of training.

In a concept attainment study, Sanders, Sterns, Smith, and Sanders (1975) found that reinforced (token) strategy training and nonreinforced training conditions were more effective than a practice-with-feedback condition; both training procedures involved provision of strategies relevant to problem solution. Denney and Denney (1974) compared the effectiveness of two modeling conditions in training in problem solving with the Twenty Questions task. No difference in training effectiveness was found between the exemplary modeling condition in which the experimenter simply modeled asking constraint-seeking questions, and the cognitive-strategy-modeling condition in which the experimenter both verbalized the strategy for formulating and using constraint-seeking questions and modeled the use of constraint-seeking questions.

The fact that subjects improve significantly even in a practice-with-no-feedback condition suggests that for many elderly the cognitive processes required for problem solving are intact and can be activated when the elderly find themselves in practice situations requiring the repeated use of these skills (Overton & Newman, 1982). However, training procedures that provide immediate feedback regarding the correctness of responses appear to be very helpful, particularly in difficult problem-solving situations (e.g., area conservation, spatial egocentrism problems). Feedback allows the elderly to monitor and evaluate whether they are using problem-solving strategies. In contrast, in very simple problem-solving situations, such as the Hoyer et al. (1973) study involving perceptual speed tasks, the elderly can readily judge the accuracy of their responses and are less in need of external feedback.

The largest training effects occur in training conditions in which there is modeling/instruction on appropriate cognitive strategies, as well as feedback and practice (Baltes & Willis, 1982; Denney, 1979). Descriptive research on a number of cognitive processes and abilities has demonstrated that many elderly do not spontaneously use cognitive strategies relevant to a particular problem-solving situation but that they apparently have these strategies within their repertoire and can utilize them effectively when encouraged or motivated to do so (Poon, 1985; Yesavage, 1983). Thus, the effectiveness of including cognitive strategies in intervention procedures appears to be in activating the elderly's use of these procedures rather than in the training of these strategies de novo (Overton & Newman, 1982). The instructor's modeling of the use of specific cognitive strategies appears to be particularly useful because it provides the elderly with a concrete, observable example of how to employ the strategy effectively, in contrast to the instructor's giving an abstract, verbal rule regarding the strategy without illustrating how to use it (Denney & Denney, 1974).

In most studies that included a strategy training condition, there has been an adult instructor who modeled use of the cognitive strategies, prompted subjects

to utilize the strategies, and gave feedback regarding the effective use of the strategies. A few recent studies have suggested that while inclusion of cognitive strategy training is important, there may be alternatives to the traditional instructor. Blackburn and Papalia (1986) compared the traditional instructor-directed cognitive strategies condition with a self-instructed interactive condition. In the later condition, older adults discussed among themselves how they solved practice problems and explained strategies to one another; feedback on correct answers to problems was also available. There were no significant differences in magnitude of training improvement between the two conditions. Scogin, Storandt, and Lott (1985) examined the effectiveness of a self-taught program of memory skills training. Subjects completed 16 study sessions in their homes, focusing on four mnemonic techniques (method of loci, novel interacting images, categorization, and chunking). Significant training effects were found and maintained at a 1-month follow-up posttest. Since strategy training appears to be activating already existing competencies in the elderly, peer tutoring or self-guided instruction may be viable alternatives to the traditional instructor role and may deserve further exploration. However, the presence of an instructor may be particularly useful when the subjects are learning truly new skills and competencies or when subjects are functioning at a very low level of competence.

Noncognitive Training Procedures. It has been hypothesized that variables such as low motivation, test anxiety, and negative stereotypes of intellectual aging interfere with the elderly's performance on cognitive measures. Several training studies have included intervention procedures targeted at these performance variables.

Denney (1980) examined the effectiveness of three performance variables (monetary reinforcement, manipulation of self-confidence, additional time to plan a game strategy) and found that none of the conditions resulted in improved performance on the Twenty Questions task. Mergler and Hoyer (1981) reported that noncontingent social praise had no effect on classification performance. Hoyer, Labouvie, and Baltes (1973) found no difference in performance on perceptual speed tasks for a practice-only versus a monetarily reinforced practice condition. Sanders, Sterns, Smith, and Sanders (1975) reported no difference in effectiveness of a reinforced-strategy training condition and a strategy-training-only condition; the reinforced condition involved the earning of tokens. Willis, Cornelius, Blow, and Baltes (1983) found that a social contact condition was not effective in improvement of attentional processes.

Review of the above literature suggests that noncognitive variables have not been effective when they were the sole focus of the treatment condition. However, there is some evidence for effectiveness when the emphasis is on selected noncognitive variables that are particularly relevant to the target cognitive ability or strategy and when these noncognitive variables are included in combination with strategy training. For example, Yesavage (Yesavage, Sheik, & Lapp, in press) has conducted training research on face–name memory recognition using an imagery-driven mnemonic strategy. Prior descriptive research indicates that many elderly have difficulty in forming and using images (Poon et al., 1980). Yesavage (1983) has found pretraining on imagery per se, followed by the mnemonic training, to be particularly effective. In addition, pretraining on anxiety reduction has been found to be effective in utilizing imagery strategies (Yesavage & Rose, 1984).

The Nature of Training Improvement

Although most training studies have reported significant improvement in performance as a function of some intervention procedure, there is considerable debate on what it is that is actually modified as a function of training and how changes in performance resulting from training should be interpreted (Donaldson, 1981; Willis & Baltes, 1981). We will discuss the nature of training effects in terms of three issues: training at the level of the construct, the ability-specific nature of training effects, and the particular behaviors and processes modified as a function of training.

Training at the Level of the Construct. In discussing the nature of training effects, it is important to begin with the acknowledgment that the target of training research is a cognitive construct or latent variable (e.g., memory span, inductive reasoning). The primary concern is not improvement on a particular test per se (e.g., WAIS Digit Symbol Substitution task, Raven's Progressive Matrices), but rather to obtain change on the latent construct represented by the particular measure (Schaie, Willis, Hertzog, & Schulenberg, in press; Willis & Schaie, 1986a). Latent constructs cannot be directly observed or measured but rather must be estimated by performance on observable tests that have been shown empirically to represent that construct. Performance, as measured by observable test scores, represents a combination of that proportion of the variance common to the latent construct and that proportion of the variance that is unique to a particular measure. It is that proportion of variance in performance associated with the construct that is of interest in assessing training effects. When training studies rely on only one test to assess training gain, it is impossible to estimate the proportion of variance in the score that is associated with the cognitive construct of interest. Subjects' performances on multiple measures representing the latent construct are needed to assess whether training improvements can be attributed to a particular cognitive construct.

Factor-analytic procedures provide one mechanism for studying training effects at the construct level. Ability factor scores represent that proportion of variance in performance that is associated with a particular latent construct. Although factor-analytic procedures have been most commonly employed in research guided by psychometric models of intelligence, these methods could also be usefully employed in research on memory processes or problem-solving

abilities and would permit researchers to examine their findings more precisely at the latent construct level.

In our recent studies (Willis & Schaie, 1986a), we have found significant training effects at the factorial level for the abilities of inductive reasoning and spatial orientation. These findings are evidence of training improvement at the latent construct level.

Structural Invariance and Training Effects. Although findings of training effects at the level of ability factor scores indicate that training improvement represents more than "teaching the test," our interpretation is based on assumptions regarding the structural stability of the measurement framework from pre- to posttest (Donaldson, 1981; Schaie et al., in press). Structural invariance addresses questions such as the following:

1. Are the same latent constructs represented in the assessment battery at pre- and posttest?
2. Are the observable measures representative of the same latent constructs before and after intervention?
3. Have the relationships among these constructs remained constant from pre- to posttest?

If assumptions regarding structural stability are not met, then interpretation of exactly what was modified as a function of training becomes ambiguous. For example, if a test representing the target ability is shown to "load" primarily on Factor A at pretest but on Factor B at posttest, then it is unclear what cognitive ability/skill that measure represents (and hence, the nature of the training effect). The existing descriptive literature offers strong support for assumptions of structural invariance (Baltes, Cornelius, Nesselroade, Spiro, & Willis, 1980; Reinert, 1970). However, the issue merits further empirical study within the training literature.

In recent training research we have examined the pre–posttest structural stability of a measurement battery representing the five primary mental abilities of verbal meaning, inductive reasoning, spatial orientation, number, and perceptual speed (Schaie, Willis, Hertzog, & Schulenberg, in press). The question examined was whether training on inductive reasoning or spatial abilities significantly altered the relationships among these abilities. We found virtually complete structural stability for the abilities (verbal, number, speed) that were not the target of training. There were some slight pre–posttest shifts in the regression weights for measures representing the trained abilities. However, all measures were still representative of the constructs they were selected to represent at pretest, and the shifts in the regression weights did not alter the interpretation of the ability-specific nature of the training effects.

These findings of structural stability were to be expected, given that the psychometric measures employed are highly reliable and have evidenced strong saturations on the particular ability constructs. However, there is need for further examination of structural stability in the training context, particularly with regard to less reliable (i.e., more statelike) constructs. For example, there is suggestive evidence that changes occur in the relationship between memory span and fluid-type abilities at various phases in the learning curve (Hofland, Willis, & Baltes, 1981; Labouvie, Frohring, Baltes, & Goulet, 1973).

Ability-Specific Nature of Training Effects. Findings from the intervention literature indicate that training effects are specific to the ability/skill that was the focus of training (Baltes & Willis, 1982; Willis, Blieszner, & Baltes, 1981; Willis & Schaie, 1986). Demonstration of ability-specific effects requires two conditions:

1. Training effects must be shown for multiple measures of the ability that was the focus of training; using factor analytic procedures, training effects can be demonstrated at the construct level.
2. No training effects should occur for measures of empirically distinct abilities that were not the focus of training.

Both conditions must be met for an ability-specific interpretation of training effects.

Some have argued that training effects should extend to abilities that are empirically distinct from the ability/cognitive process that was the target of training (Donaldson, 1981). Why this should be so is unclear to us. First, training procedures have typically focused on those cognitive strategies and behaviors that are directly related to performance on the target ability, and thus it would be expected that training effects would be limited to the target ability. Second, the current state of the field of cognitive psychology is not such that it is possible to specify cognitive strategies that are common (i.e., generalizable) across empirically distinct abilities. In mainstream cognitive psychology, it has proved exceedingly difficult to specify and operationalize metacognitive components or executive processes that are truly generalizable across multiple cognitive domains. Componential analyses have typically been limited to one particular ability construct (e.g., analogical reasoning), and attempts to define metacomponents have remained at the level of theoretical discussion (Detterman, 1980; Sternberg, 1982).

Focusing training efforts at the level of a specific ability makes sense from the perspective of longitudinal research findings (Cunningham, this volume; Schaie, 1983). There are wide individual differences not only in the timing of the onset of cognitive decline but also in terms of which particular abilities or skills exhibit early decline. In young–old age, cognitive decline tends to be highly specific and individual. That is, many individuals show decline on one ability but not on another, even when only abilities exhibiting normative patterns of early decline are considered. Moreover, the particular ability showing decline varies by

individual. Thus, a prescriptive, individualized approach to training that would result in ability-specific effects would seem most useful.

Specific Behaviors Modified as a Function of Training. Training improvement, if measured by the total number of items answered correctly, may occur in several ways. As a result of training, the older subjects may respond more quickly, therefore attempting more problems. The total number of correct (and incorrect) responses will increase, as a function of more items being attempted. However, the proportion of correct items to the total number of items attempted may not change; thus, there will be little or no change in accuracy. Alternatively, there may be little increase in the total number of items attempted as a function of training, but the subject may become more accurate, as evidenced by a decrease in the proportion of commission errors and an increase in the proportion of correct responses.

Error analyses of training data are particularly useful in understanding the qualitative nature of behavioral change associated with training. In their study of spatial egocentrism, Schultz and Hoyer (1976) found that training improvement occurred as a result of a significant reduction in perceptual judgment errors (wrong responses) but not in egocentric responses per se. In fact, this sample of elderly made relatively few egocentric judgments, even prior to training. Berg, Hertzog, and Hunt (1982) examined the effects of practice on the spatial ability of mental rotation for older men and women. Practice was found to be particularly beneficial for older women; improved performance for older women was attributed primarily to decreased reaction times, not to decreases in commission errors.

Error analyses indicate that various treatment conditions may differentially affect accuracy of performance. In a training study on fluid abilities, Baltes, Dittman-Kohli, and Kliegl (1986) found a significant increase in the number of correct responses for both the cognitive training and a no-treatment control group. However, cognitive training was also effective in decreasing the proportion of commission errors, thus raising the accuracy level. In contrast, the no-treatment control group exhibited an increase in commission errors. These findings are supported by results from two studies involving practice with no feedback conditions (Hofland, 1981; Hofland et al., 1981). Subjects in the practice conditions exhibited significant increases in the total number of items attempted and in the number of correct responses; however, the proportion of commission errors did not significantly decrease. Thus, while an increase in number of correct responses can be elicited via a variety of treatment conditions, a cognitive strategy training condition appears to be particularly effective in improving accuracy.

Our recent training research also suggests that the qualitative nature of training improvement may vary with the particular ability that is the target of training improvement and with whether the training subject has experienced previous decline on the target ability (Willis, 1986; Willis & Schaie, 1986a). For example, the elderly's performance on fluid abilities, such as figural relations and inductive reasoning, is characterized by a large number of commission errors. Therefore, an important outcome of training on these fluid abilities is a reduction in the proportion of commission errors (i.e., an increase in accuracy). Error analyses of our training research on inductive reasoning support this conclusion (Willis, in press.)

Descriptive research on spatial orientation ability (mental rotation) indicates that the major source of individual differences (including age differences) is in the speed of mental rotation (Berg et al., 1982; Cerella, Poon, & Fozard, 1981); the rate of commission errors is relatively low. Thus, one would predict that training on mental rotation would result in an increase in the number of items attempted, reflecting increased speed of mental rotation, but that there would be relatively little change in error rate. Error analyses of our mental rotation training study only partially support these hypotheses. For subjects experiencing no prior decline in the ability, training improvement does reflect primarily an increase in response speed. However, for subjects experiencing prior decline, training improvement reflects improved accuracy as well as some increase in speed of mental rotation (Willis & Schaie, in press).

RELATION OF COGNITIVE TRAINING TO COMPETENCE IN REAL-LIFE ACTIVITIES

The types of mental abilities and processes that have been the target of cognitive training research have been derived from psychometric, Piagetian, and information processing approaches to the study of intellectual aging. The training research literature cited above indicates that there is considerable plasticity in the elderly's functioning in these domains. These findings argue against a position of irreversible cognitive decrement for all individuals and abilities.

There remains, however, the important question of the implications of cognitive training research for improving the competence of older adults in activities of daily living. What are the relationships between the types of cognitive domains examined in cognitive training research and competence in activities of daily living? Can performance on real-life tasks be improved? What types of training might be useful in improving performance on real-life tasks? Research related to such questions is very limited. In this final section of the paper we will briefly outline one possible procedure for examining these issues.

Practical Intelligence: Definition and Measurement

We must begin with definitions. The term *practical intelligence* has been used to differentiate between academic intelligence (e.g, psychometric, Piagetian, information-processing approaches) and other forms of intelligence or competence

(Commons, Richards, & Armon, 1982; McClelland, 1973; Sternberg & Wagner, 1986). There is no commonly agreed-on definition of practical intelligence. The term has been used to describe areas such as "practical know-how," professional competence in one's work sphere, social judgment, and the layperson's conception of an intelligent person. In gerontology, unique aspects of intellectual aging, such as wisdom, have also been considered (Dixon & Baltes, 1986; Labouvie-Vief, 1982; Schaie, 1978).

Our approach to the study of practical intelligence in later adulthood proceeds from the assumption that there are classes of everyday activities that are critical for adaptive functioning in common life situations (Willis & Schaie, 1986b). Because a major concern in old age is maintenance of independent living, this approach focuses on tasks associated with independent, effective functioning. For example, inability to perform tasks, such as comprehending a medicine bottle label or utilizing information in the yellow pages of a phone directory, may lead to the curtailment of independent living for many elderly. There is no exhaustive taxonomy of real-life tasks. Our research has focused on a subset of tasks that represents categories of common problems experienced by many elderly.

It is difficult, if not impossible, to observe directly the elderly's performance in many real-life situations; thus, prototypical task types have been developed. These task items represent a simulation of how the elderly might perform a task in daily life (Willis & Schaie, 1986b). Although the specific stimuli in the task item may not have been encountered previously, it is assumed that the subject applies the same relevant cognitive skills and information to the task item that he or she would apply in a similar real-life problem. At least two characteristics distinguish the types of tasks we have studied from other forms of practical intelligence. First, these tasks involve the interpretation/comprehension of printed material. Second, the tasks, for the most part, involve logical reasoning, and therefore there are logically "correct" responses.

The Relationship between Traditional and Practical Intelligence: A Hierarchical Approach

Are the types of cognitive domains studied in training research related to practical intelligence as described above? Some have suggested that the domains of traditional and practical intelligence are independent and often unrelated spheres (Sternberg & Wagner, 1986). There has been little empirical examination of the possible relationships between the two domains. For the types of practical intelligence tasks we are interested in, it has been useful to examine a hierarchical relationship between the domains (Schaie, 1978; Willis & Schaie, 1986c).

In this hierarchical scheme, the most elemental components are ability factors, such as the primary mental abilities represented within a psychometric approach to traditional intelligence (Thurstone, 1938). Examples of ability factors include inductive reasoning, verbal comprehension, and spatial relations. Although our examples of ability factors have been derived from a psychometric approach, such factors could also be derived from a Piagetian or information-processing approach.

Because behavior in real-world contexts is of necessity complex, we assume that no single ability factor can adequately predict performance in a specific real-life situation. Rather, some combination of ability factors is required to best predict performance on real-life tasks. Thus, practical intelligence can be described as involving the expression of that combination of ability factors that, given minimally acceptable levels of motivation, will permit adaptive behavior within a specific situation or class of situations. We do not wish to suggest that all of the variability in real-life task performance can be accounted for by a combination of ability factors, but we believe such factors represent an important first strata in a hierarchical approach to the study of practical intelligence (Schaie, 1978; Willis & Schaie, 1986b).

A series of steps has been undertaken to examine the hierarchical relationship between ability factors and a subset of practical intelligence tasks (Willis & Schaie, 1986b). These steps include: (1) Identification of clusters of practical intelligence tasks that are intercorrelated. The elderly's scores on prototypical tasks representing practical intelligence were entered into an item factor analysis to determine task clusters sharing common variance. Seven task clusters were identified. (2) Examination of the relationship between task clusters and ability factors. The proportion of variance that each task cluster shared with the ability factors in an established structural model of psychometric intelligence was examined via a series of extension analyses. The structural model of psychometric intelligence represented the Thurstonian (1938) primary mental ability factors of verbal meaning, inductive reasoning, spatial orientation, number, and perceptual speed. In an extension analysis for each practical task cluster, a loading of the cluster on each of the five ability factors was obtained.

For a cluster involving tasks such as comprehending labels/directions on bottles (e.g., medications, household substances), the following loadings were obtained: Verbal Meaning, 0.40; Inductive Reasoning, 0.31; Spatial Orientation, 0.14; Perceptual Speed, −0.04; and Number, 0.02. The factor loadings were statistically significant for Verbal Meaning, Inductive Reasoning, and Spatial Orientation. For a cluster involving tasks such as comprehending technical documents (e.g., guarantees) and newspaper editorials, the following loadings were obtained: Inductive Reasoning, 0.55; Verbal Meaning, 0.13; Perceptual Speed, 0.11; Spatial Orientation, 0.02; and Number, 0.02. Only the factor loadings for Inductive Reasoning and Verbal Meaning were statistically significant.

These preliminary data suggest several points regarding the relationship between traditional intelligence factors and tasks of practical intelligence. First, there are significant and reliable relationships between traditional abilities and many tasks of daily living encountered by the elderly; traditional and practical intelligence are not totally unrelated spheres of competence. Second, the relationships between practical and traditional intelligence domains can be empirically determined through procedures such as those described above. Third, tasks of practical intelligence are complex and involve the utilization of several mental abilities. The particular combination of abilities will vary across different types of practical tasks. Moreover, the relative importance of a particular mental ability will vary from task to task. For example, verbal, reasoning, and spatial abilities were all found to be important in comprehending bottle labels and directions. Although verbal ability shared the most common variance with task performance, reasoning ability was also important in interpreting the label directions and determining what dosage of a medication would be appropriate for an individual with certain characteristics (age, weight) as specified in the directions. In interpreting technical documents, the role of reasoning ability was even more salient.

Cognitive Training and Practical Intelligence

If substantial relationships exist between ability factors and some important tasks of daily living, does cognitive training enhance performance on these practical tasks? There is suggestive evidence that training is effective in improving performance on certain real-life tasks. For example, in our own training research on inductive reasoning, we found significant training effects on the cluster of tasks involving comprehension of technical documents and newspaper editorials. Recall that in the extension analyses above this particular task cluster showed a substantial loading on the inductive reasoning factor.

The critical issue is that for cognitive training to be effective, the practical tasks of interest must share substantial common variance with the cognitive abilities trained. Empirical analyses would need to be undertaken to examine the relative proportion of common variance shared between various cognitive abilities and the subsets of practical tasks in which the elderly exhibit deficiencies. If analyses indicate that more than one ability shares substantial common variance with the practical task of interest, then cognitive training may need to focus on multiple abilities. In this approach, therefore, the particular cognitive abilities targeted for training would be determined by the types of practical tasks that are of interest or in need of enhancement.

Why train on cognitive abilities, rather than on the practical tasks themselves? This is a question that deserves empirical investigation. However, the number of specific tasks of daily living on which training might be useful is very large. Training on cognitive abilities that underlie performance on these tasks may therefore be more efficient and result in broader transfer of training than training that is restricted to a specific practical task. This conclusion is subject to least two conditions: (1) Clusters of empirically related practical tasks must be identi- fied. and (2) a small subset of ability factors must be identified that share significant common variance with the practical task cluster.

TRAINING RESEARCH AND MODELS OF INTELLECTUAL AGING

Recent conceptions of intellectual aging have focused on the multidimensional and multidirectional nature of the aging process (Willis, 1985; Willis & Baltes, 1981). Intelligence in old age needs to be viewed as a multidimensional phenom- enon rather than as a homogeneous or global entity. Various dimensions of intellectual functioning exhibit different patterns of developmental change (Schaie, 1983). Within the classical pattern of aging, abilities and skills requiring abstract reasoning, novel material, and/or highly speeded responding have been shown to have an earlier onset of normative decline, in contrast to overlearned, well-practiced abilities and skills that show relative stability until late old age (Botwinick, 1977).

Findings from cognitive training research suggest that this multidimensional, multidirectional perspective of intellectual aging is further complicated in that issues of plasticity need to be considered, as well as normative patterns of developmental change. In comprehensive theories of developmental aging, it is necessary to consider not only how the elderly perform on average but also to explore the range of potential performance under a variety of contexts (Baltes & Baltes, 1980; Kliegl & Baltes, in press).

As our perspective of intellectual aging expands, the meaning of the well-worn and overextended terms *stability* and *decline* need to be reexamined and refined. Although both terms have come to hold multiple meanings, the term *decline* is of particular concern to intervention researchers. Just as we have come to un- derstand that intellectual aging must be conceptualized and studied, not as a global phenomenon but in terms of differential patterns of developmental change, we believe that those studying intellectual aging must now begin to distinguish between various types of decline.

Reversible Decline

In earlier research on intellectual aging there was the implicit, if not explicit, assumption that the term *decline* meant an irreversible change in performance. However, it is now quite evident that not all decline in cognitive performance is irreversible. Whether or not a decline is irreversible cannot be determined solely

from longitudinal research on age change but must be examined empirically through medical and/or behavior intervention research (Smith, 1980). Language is now needed to differentiate between various types of decline in terms of remediability and/or compensation (Bäckman, in press). The term *remediation* may be used to refer to that type of decline for which there is reversal (or partial reversal) of age-related decline in cognitive performance via behavioral interven- tion procedures.

There are two types of remediation that should be differentiated. First, there is *remediation-in-kind*. That is, the intervention results in remediation such that the behavior after intervention is qualitatively the same as the behavior prior to decline. This type of remediation appears most likely to occur for simple cognitive skills and processes. For example, several practice studies have demonstrated significant improvement in performance on simple perceptual speed tasks (Hoyer et al., 1973). Assuming that improvement with practice does reflect a reversal of age-related decline for some individuals, then remediation- in-kind would have occurred if the posttraining behavior was qualitatively (e.g., number of attempted items, accuracy rate) like the subjects' performance prior to decline.

For more complex abilities and skills, reversal of decline may be more likely to involve *remediation-with-compensation*. Although intervention may result in significant performance gains, there may be qualitative differences between behavior prior to decline and behavior after intervention. For example, with regard to spatial orientation ability, age-related decline may result in slowing in the rate of mental rotation that is only partially reversible with intervention. However, the effect of intervention may be such that the elder compensates by becoming more accurate. Loss in one aspect of the outcome behavior is com- pensated for by increases in another aspect.

Irreversible Decrement

In very old age or in the case of pathologies, irreversible decline in cognitive functioning becomes increasingly likely. In some cases, however, there is the potential for *decrement-with-external-compensation*. Just as prosthetic devices such as eyeglasses and hearing aids have aided the elderly in compensating for some irreversible sensory change, so prosthetic environments can be constructed to enable the impaired elderly to compensate, at least partially, for irreversible cognitive loss (Zarit & Anthony, 1985). Interactive memory devices can aid memory-impaired elderly in self-medication and self-care tasks (Wilson & Mof- fat, 1984). Wandering devices are being employed to monitor the movements of dementia patients. The technology is currently available for a significant expan- sion in the types of external compensation systems to assist the impaired elderly.

Stability

In our training research approximately 45% of the older adults studied had not declined on either inductive reasoning or spatial orientation ability (Willis & Schaie, 1986a). Training improvement for these subjects represented new performance levels beyond those previously exhibited. There is need, then, for further differentiation in the meaning of the term *stability*.

Findings from longitudinal research indicate that for abilities that are regularly exercised and overlearned (e.g., recognition vocabulary), there is considerable stability in performance until the 70s are reached. Indeed, some modest improvement in verbal ability occurs through the early 50s, on average. Although there has been little empirical research, it is likely that functioning on these abilities probably represents *maintenance-in-kind*. That is, there is little qualitative change in the nature of verbal performance in late middle age or early old age.

In contrast, for other overlearned, regularly exercised skills, stable performance with age on an important outcome measure may best be described as *maintenance-with-compensation*. The work of Salthouse (this volume) is particularly relevant to this issue. Salthouse first matched older and younger experienced typists on the outcome measure of typing proficiency. Although older typists were found to be slower on reaction-time measures related to typing speed, the older typists compensated for increased response time by processing larger chunks of information.

Finally, there is an increasing body of research that indicates that older adults can acquire totally new skills and bodies of information. For example, several studies have reported the acquisition of various types of computer skills by the elderly (Clark, Lanphear, & Riddick, 1987; Garfein & Schaie, 1986).

SUMMARY AND CONCLUDING REMARKS

In this chapter, we began by discussing what were the major questions addressed by cognitive training research. While many researchers and interpreters of the cognitive training literature have assumed that remediation of age-related decline is the central question, we have argued that most cognitive training studies cannot address this issue directly, since prior longitudinal data are not available on the subjects. Given that many subjects who participate in cognitive training studies are in good health, comparatively well educated, and are the young-old, it cannot be assumed that all or most of these subjects have experienced substantial age-related decline. This is not to argue that these subjects will not experience age-related decline (most assuredly will), but that it is the young-old who show the most variability in the timing and rate of decline, even for cognitive variables exhibiting relatively early normative decline. To interpret training effects in terms of age-related decline (without longitudinal data) is

problematic. Given that significant cohort effects have been demonstrated for the very abilities of interest to training researchers, young adult comparison groups cannot serve as a proxy for the older age cohorts at younger ages (see Cunningham, this volume). Thus, the central question in cognitive training research can, more accurately, be defined as what is the range of plasticity or variability in cognitive performance in old age, as a function of behavioral intervention. The major concern in examining plasticity is the range or magnitude of change in performance, not whether the change reflects remediation of age-related decline or improvement of performance in subjects showing no prior decline.

Two other questions of concern in training research focus on the types of training procedures shown to be most effective in improving cognitive functioning, and the specific aspects of behavior that are changed as a function of training. More intensive training procedures involving the factors of feedback, and modeling/instruction on specific strategies related to problem solution, have been found to result in the largest training effects. Moreover, recent research suggests that various training delivery systems (e.g., instructor-guided, self-directed, peer tutoring) may be effective, and further research on cost–benefit analyses of these various delivery systems is needed.

While findings from training research have indicated considerable plasticity in the elderly's cognitive functioning, some have questioned the practical significance of such training. Are there implications from training research for maintenance of effective functioning in tasks of daily living? To begin to address this issue, it is first necessary to examine the relationship between traditional models of intelligence and tasks of daily living (i.e., practical intelligence). In the second section of this chapter, we briefly outline one possible approach to studying this issue. A hierarchical relationship between basic mental ability factors and practical tasks is proposed; it is hypothesized that effective functioning on many tasks of daily living requires a combination of several abilities. There is need for much further research examining the relationship between traditional abilities and tasks of daily living. However, recent analyses conducted in our laboratory suggest that when there is a substantial relationship between particular cognitive abilities and specific everyday tasks, transfer of training effects may occur such that improvement of performance on tasks of daily living occurs as a function of training.

In the final section of this chapter, we suggest that our models of intellectual aging are now in need of further differentiation of the concepts of intellectual decline and stability. Findings from recent training studies indicate that intellectual aging researchers need to differentiate between remediable and irreversible intellectual decline. Some may question the need for differentiating the various types of decline and stability in cognitive functioning discussed above. However, at the current stage in the development of models of cognitive aging, it appears particularly important. First, these distinctions serve as a constant reminder that irreversibility is not to be assumed in discussions of decline, and that plasticity as well as normative patterns of intellectual aging must be considered. More importantly, if findings from training research are to contribute to our understanding of intellectual aging, it is necessary to examine for what abilities and under what intervention procedures the behavioral processes of decline and of remediation are qualitatively similar and when they are different. Equally important is our understanding of the mechanisms that the aging individual uses to compensate for and to cope with age-related decrements in functioning, and how the external environment may limit or support the individual's attempts to maintain competence.

REFERENCES

Bäckman, L. (in press). Varieties of memory compensation of older adults in episodic remembering. In L. Poon, D. Rubin, & B. Wilson (Eds.), *Everyday cognition in adult and late life*.

Baltes, P. B., & Baltes, M. M. (1980). 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.

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

Baltes, P. B., Dittman-Kohli, F., & Dixon, R. A. (1984). New perspective on the development of intelligence in adulthood: Toward a dual-process conception and a model of selective optimization with compensation. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior*, Vol. 6 (pp. 33–76). New York: Academic Press.

Baltes, P. B., Dittman-Kohli, F., & Kliegl, R. (1986). Reserve capacity of the elderly in aging-sensitive tests of fluid intelligence: Replication and extension. *Psychology and Aging, 1*, 172–177.

Baltes, P. B., & Willis, S. L. (1982). Enhancement (plasticity) of intellectual functioning in old age: Penn State's Adult Development and Enrichment Project (ADEPT). In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes* (pp. 353–389). New York: Plenum Press.

Bellucci, G., & Hoyer, W. (1975). Feedback effects on the performance and self-reinforcing behavior of elderly and young adult women. *Journal of Gerontology, 30*, 456–460.

Beres, C., & Baron, A. (1981). Improved digit substitution by older women as a result of extended practice. *Journal of Gerontology, 36*, 591–597.

Berg, C., Hertzog, C., & Hunt, E. (1982). Age differences in the speed of mental rotation. *Developmental Psychology, 18*, 95–107.

Blackburn, J., & Papalia, D. (1986, November). *Modifiability of fluid intelligence: A comparison of two training approaches*. Paper presented at the Meeting of the Gerontological Society, Chicago, IL.

Blieszner, R., Willis, S. L., & Baltes, P. B. (1981). Training research in aging on the fluid ability of inductive reasoning. *Journal of Applied Developmental Psychology, 2*, 247–265.

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

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Cerella, J., Poon, L., & Fozard, J. (1981). Mental rotation and age reconsidered. *Journal of Gerontology, 35*, 620–624.

Clark, J., Lanphear, A., & Riddick, C. (1987). The effects of videogame playing on the response selection processing of elderly adults. *Journal of Gerontology, 42*, 82–85.

Commons, M., Richards, F., & Armon, G. (1982). *Beyond formal operations: Late adolescence and adult cognitive development*. New York: Praeger.

Coyne, A. C. (1981). Age differences and practice in forward visual masking. *Journal of Gerontology, 36*, 730–732.

Cronbach, L., & Furby, L. (1970). How should we measure "change"—or should we? *Psychological Bulletin, 74*, 68–80.

Denney, N. W. (1979). 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.

Denney, N. W. (1980). The effect of manipulation of peripheral, noncognitive variables on the problem-solving performance of the elderly. *Human Development, 23*, 268–277.

Denney, N. W. (1982). Aging and cognitive changes. In B. B. Wolman (Ed.), *Handbook of developmental psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Denney, N. W., & Denney, D. (1974). Modeling effects on the questioning strategies of the elderly. *Developmental Psychology, 10*, 400–404.

Detterman, D. K. (1980). Understand cognitive components before postulating metacomponents. *Behavioral and Brain Sciences, 3*, 589.

Dixon, R., & Baltes, P. B. (1986). Toward life-span research on the functions and pragmatics of intelligence. In R. Sternberg & R. Wagner (Eds.), *Practical intelligence* (pp. 203–235). Cambridge: Cambridge University Press.

Donaldson, G. (1981). Letter to the editor. *Journal of Gerontology, 36*, 634–636.

Englehardt, K., & Leifer, L. (1984). The VA RR&D center: Where computers and robots serve elders. *Generations, Summer*, 38–39.

Erber, T. J. (1976). Age differences in learning and memory on a digit-symbol substitution task. *Experimental Aging Research, 2*, 45–53.

Garfein, A., & Schaie, K. W. (1986, November). *Microcomputer proficiency in later middle-aged adults: Teaching old dogs new tricks*. Paper presented at the meeting of the Gerontological Society, Chicago, IL.

Grant, E., Storandt, M., & Botwinick, J. (1978). Incentive and practice in the psychomotor performance of the elderly. *Journal of Gerontology, 33*, 413–415.

Harris, C. (1963). *Problems in measuring change*. Madison, WI: University of Wisconsin Press.

Hofland, B. F. (1981). *Practice effects in the intellectual performance of the elderly: Retesting as an intervention strategy*. Unpublished doctoral dissertation, The Pennsylvania State University, University Park, PA.

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

Hornblum, J. N., & Overton, W. F. (1976). Area and volume conservation among the elderly: Assessment and training. *Developmental Psychology, 12,* 68–74.

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

Kerlinger, F. (1973). *Foundations of behavioral research.* New York: Holt, Rinehart, & Winston.

Kliegl, R., & Baltes, P. B. (1987). Theory-guided analysis of mechanisms of development and aging through testing-the-limits and research on expertise. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course.* Norwood, NJ: Ablex.

Labouvie-Vief, G. (1982). Dynamic development and mature autonomy: A theoretical prologue. *Human Development, 25,* 161–191.

Labouvie, G., Frohring, W., Baltes, P., & Goulet, L. (1973). Changing relationship between recall performance and abilities as a function of stage of learning and timing of recall. *Journal of Educational Psychology, 64,* 191–198.

Le Breck, D., & Baron, A. (1987). Age and practice effects in continuous recognition memory. *Journal of Gerontology, 42,* 89–91.

McClelland, D. (1973). Testing for competence rather than for "intelligence." *American Psychologist, 28,* 1–14.

Mergler, N., & Hoyer, W. (1981). Effects of training on dimensional classification abilities: Adult age comparisons. *Educational Gerontology, 6,* 135–145.

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

Nunnally, J. (1982). The study of human change: Measurement, research strategies, and methods of analysis. In B. B. Wolman (Ed.), *Handbook of Developmental Psychology.* Englewood Cliffs, NJ: Prentice-Hall.

Overton, W. F., & Newman, J. (1982). Cognitive development: A competence-activation/utilization approach. In T. Field, A. Houston, H. Quay, L. Troll, & G. Finley (Eds.), *Review of human development* (pp. 217–241). New York: Wiley.

Poon, L. (1985). Differences in human memory with aging: Nature, causes, and clinical implications. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 427–462). New York: Van Nostrand Reinhold.

Poon, L., Fozard, J., Cermak, L., Arenberg, D., & Thompson, L. (Eds.). (1980). *New directions in memory and aging.* Hillsdale, NJ: Erlbaum.

Reinert, G. (1970). Comparative factor analytic studies of intelligence throughout the human life span. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory* (pp. 468–487). New York: Academic Press.

Rosenberg, G., Greenwald, B., & Davis, K. (1983). Pharmacologic treatment in Alzheimer's disease. In B. Reisberg (Ed.), *Alzheimer's disease: The standard reference* (pp. 329–340). New York: Free Press.

Sanders, J. C., Sterns, H. L., Smith, M., & Sanders, R. E. (1975). Modification of concept identification performance in older adults. *Developmental Psychology, 11,* 824–829.

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

Schaie, K. W. (1983). The Seattle Longitudinal Study: A twenty-one year investigation of psychometric intelligence. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 64–155). New York: Guilford Press.

Schaie, K. W., & Willis, S. L. (1982). Life span development. In H. E. Mitzel (Ed.), *Encyclopedia of Educational Research* (5th ed.). New York: Macmillan.

Schaie, K. W., & Willis, S. L. (1985, August). *Differential ability decline and its remediation in late adulthood.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles.

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

Schaie, K. W., Willis, S. L., Hertzog, C., & Schulenberg, J. (in press). Effects of cognitive training upon primary mental ability structure. *Psychology and Aging, 2.*

Schultz, N. R., & Hoyer, W. J. (1976). Feedback effects on spatial egocentrism in old age. *Journal of Gerontology, 31,* 72–75.

Scogin, F., Storandt, M., & Lott, L. (1985). Memory-skills training, memory complaints, and depression in older adults. *Journal of Gerontology, 40,* 562–568.

Smith, A. (1980). Cognitive issues: Advances in the cognitive psychology of aging. In L. Poon (Ed.), *Aging in the 1980s* (pp. 223–226). Washington, DC: American Psychological Association.

Sternberg, R. (Ed.). (1982). *Advances in the psychology of human intelligence,* Vol. 1. New York: Erlbaum.

Sternberg, R., & Wagner, R. (Eds.). (1986). *Practical intelligence.* New York: Cambridge University Press.

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

Taub, H., & Long, M. (1972). The effects of practice on short-term memory of young and old subjects. *Journal of Gerontology, 27,* 494–499.

Thurstone, L. L. (1938). *Primary mental abilities.* Chicago: University of Chicago Press.

Willis, S. L. (1985). Towards an educational psychology of the adult learner. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed.). New York: Van Nostrand Reinhold.

Willis, S. L. (1986, August). *Inductive reasoning: Longitudinal change and cognitive training.* Paper presented at the meeting of the American Psychological Association, Washington, DC.

Willis, S. L. (in press). Improvement with cognitive training: Which old dogs learn what tricks? In L. Poon, D. Rubin, and B. Wilson (Eds.), *Everyday cognition in adult and late life.* New York: Cambridge University Press.

Willis, S. L., & Baltes, P. B. (1980). Intelligence in adulthood and aging: Contemporary issues. In L. W. Poon (Ed.), *Aging in the 1980s.* Washington, DC: American Psychological Association.

Willis, S. L. & Baltes, P. B. (1981). Letters to the editor. *Journal of Gerontology, 36,* 634–638.

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

Willis, S. L., Cornelius, S. W., Blow, F., & Baltes, P. B. (1983). Training research in aging: Attentional processes. *Journal of Educational Psychology, 75,* 257–270.

Willis, S. L., & Schaie, K. W. (1981). Maintenance and decline of adult mental abilities: II. Susceptibility to experimental manipulation. In F. Grote & R. Feringer (Eds.), *Adult learning and development*. Bellingham, WA: Western Washington University.

Willis, S. L., & Schaie, K. W. (1986a). Training the elderly on the ability factors of spatial orientation and inductive reasoning. *Psychology and Aging, 1*, 239–247.

Willis, S. L., & Schaie, K. W. (1986b). Practical intelligence in later adulthood. In R. Sternberg & R. Wagner (Eds.), *Practical intelligence* (pp. 236–270). New York: Cambridge University Press.

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

Wilson, B., & Moffat, N. (Eds.). (1984). *Clinical management of memory problems*. Rockville, MD: Aspen.

Yesavage, J. (1983). Imagery pretraining and memory training in the elderly. *Journal of Gerontology, 29*, 217–275.

Yesavage, J. (in press). Mnemonics as modified for use by the elderly. In L. Poon, D. Rubin, & B. Wilson (Eds.), *Everyday cognition in adult and late life*. New York: Cambridge University Press.

Yesavage, J., & Rose, T. (1984). The effects of a face–name mnemonic in young, middle aged, and elderly adults. *Experimental Aging Research, 10*, 55–57.

Yesavage, J., Sheik, J., & Lapp, D. (in press). Mnemonics as modified for use by the elderly. In L. Poon, D. Rubin, & B. Wilson (Eds.), *Everyday cognition in adult & late life*. New York: Cambridge University Press.

Yesavage, J., Westphal, J., & Rush, L. (1981). Senile dementia: Combined pharmacologic and psychologic treatment. *Journal of the American Geriatrics Society, 29*, 164–171.

Zarit, S., & Anthony, C. (1985). Interventions with dementia patients and their families. In M. Gilhooly, S. Zarit, & J. Birren (Eds.), *The dementias: Policy and management*. Englewood Cliffs, NJ: Prentice-Hall.