Four

Methodological Issues in Behavioral Intervention Research with the Elderly

Sherry L. Willis

I. Introduction

There has been the assumption that the primary goal of behavioral intervention research should be to assess the efficacy of a given intervention—to examine whether there is positive gain in functioning as a result of training. In large-scale intervention research, such as clinical trials research, the primary focus is on the efficacy of an intervention to produce a specific outcome. There is relatively less concern with such issues as the role of individual differences or the specific processes and mechanisms underlying the desired outcome.

Although determining the efficacy of a given intervention is an important objective, programmatic intervention research should be aimed at the broader goal of answering a series of theoretically important empirical questions (Baltes & Willis, 1977; Hazlett-Stevens & Borkovec, 1999; Schulz & Martire, 1999). When couched within a particular theory or conceptual framework, programmatic intervention research seeks to address questions such as: What is the nature of the problem or deficit? What specific mechanisms, processes, or components of the intervention are responsible for the desired change? What individual-difference variables are associated with responsivity to change? How can the change be maintained?

Kastenbaum (1968) has suggested that until the 1960s, gerontological researchers were largely satisfied with "counting and classifying the wrinkles of aged behavior" (p. 282). Since the 1960s there has been increasing focus on the "why" or explanations of aging. The step from explanation to experimentation and intervention was a relatively small and expected one (Baltes & Danish, 1979). In the 1970s a new movement of behavioral intervention research in aging began to appear. Cognitive training research was one manifestation of this trend. These intervention studies were typically small scale, conducted by a single investigator in a single laboratory (Camp, 1999; Kliegl, Smith, & Baltes, 1989; Neely & Backman, 1995; Willis, 1990; Yesavage et al., 1990). In the 1980s and particularly in the 1990s, there has been a shift toward larger scale intervention studies, involving multiple sites and investigators (Appel et al., 1995; Jobe et al., in press; LaRosa et al., 1994; Ory et al., 1993). These large-scale

Handbook of the Psychology of Aging

interventions were often an outgrowth of the findings of the earlier smaller scale intervention work. Although there have been recent reviews of substantive findings from these training studies (Schulz, Maddox, & Lawton, 1999), the evolution of intervention research with the elderly has also resulted in methodological issues specific to behavioral interventions, particularly when conducted from a psychosocial or life span developmental perspective (Baltes & Danish, 1979; Smyer & Gatz, 1986).

This chapter serves as a vehicle for a discussion of what we consider to be some of the most salient current methodological issues in behavioral intervention research with the elderly. Many of the methodological issues discussed are relevant to different types of behavioral interventions with the elderly, such as those involving exercise and nutrition. Given the author's research interests and experiences, however, many of the exemplars are drawn from cognitive training research. The chapter begins by considering several definitions of intervention and by enumerating characteristics that are common to most behavioral interventions. The second section consists of a review of the primary experimental designs employed in behavioral training studies, with discussion of the strengths and limitations of each design. In the third section of the chapter, there is a discussion of the mechanisms or processes that are assumed to underlie the intervention and how these mechanisms mediate intervention outcomes. A final section focuses on several factors related to treatment outcomes. Levels of treatment outcomes are considered, and the critical but difficult phenomenon of training transfer is discussed.

There are numerous other important methodological issues that are not considered in this chapter, such as sampling, attrition, training of the interventionist, quality control, follow-up assessment, and measurement characteristics of the outcome measures (see Schulz et al., 1999). The topics focused on in this chapter, such as mechanisms accounting for treatment outcomes and transfer effects, reflect methodological issues that have recently been identified in the literature as relatively understudied but of particular salience (Salomon & Perkins, 1989; Schulz & Martire, 1999).

II. Defining Behavioral Intervention Research

The concept of behavioral intervention does not have a singular meaning. Schulz and Martire (1999) state that "an intervention study involves actions that alter, or are intended to alter, relationships between observable phenomena" (p. 2). The goal of intervention is that some agent under human control can be manipulated to bring about desired change; thus, the design of choice in most cases is a randomized trial. The action of interest typically has a definable onset and in many cases a clear termination. In a chapter on interventions in life span development and aging, Baltes and Danish (1979) consider a gerontological intervention as programmatic attempts aimed at modification of the course of psychological aging. In a prior edition of the *Handbook of the Psychology of Aging*, a psychological intervention was defined as planned processes of behavioral change that employ a deliberate application of psychological principles and theory (Smyer, Zarit, & Qualls, 1990).

In this chapter, behavioral interventions will be considered to have in general the following characteristics. The intervention is a planned effort with the goal of manipulating or altering behavior. The key independent variable is behavioral in nature. In many instances, the dependent variable(s) are also behaviors (outcomes). The target of the intervention is an

individual, typically an older adult. Given that it is a planned effort, the intervention typically has a defined onset and often a defined duration and termination. The design should involve a comparison group, and individuals are randomly assigned to groups. Ideally, the intervention is grounded in one or more theoretical or conceptual framework(s) and is based on prior descriptive research. The hypothesized causal links between the intervention and the outcome should be stated, and the mechanisms or processes that mediate the intervention should be explicated. Many of the above characteristics are representative of the broader domain of experimental research of which behavioral intervention research is a class. In contrast to many experiments, intervention studies are often of a longer duration involving multiple training sessions, and the temporal durability of outcome(s) is expected to be longer. Moreover, intervention studies often involve multiple levels of outcomes and a broader assessment battery.

III. Design Issues in Behavioral Interventions

The particular experimental design chosen for an intervention study serves to address some of the above issues. That is, the specific control or comparison groups chosen in between-group experimental designs contribute knowledge regarding the nature of the problem or deficit and the processes or mechanisms hypothesized to be responsible for the change. In this section of the chapter, some of the more common between-group designs used in behavioral intervention research are reviewed (Borkovec, 1994; Campbell & Stanley, 1963; Kazdin, 1992). Exemplar cognitive training studies representing a particular design are cited.

A. No-Treatment Control Group

The most common comparison employed in intervention research is the no-treatment group. In clinical research, this group is often referred to as "treatment as usual." This design examines the magnitude of change associated with the intervention compared with the amount of change that would have occurred without the intervention. The no-treatment group receives the same pre- and postintervention assessment battery but otherwise no intervention or contact. The no-treatment control permits assessment of the amount of change that could have occurred simply due to the passage of time, including the pre–post assessment. Change in the no-treatment control group may occur due to factors such as maturation, sociocultural and history-related events, spontaneous remission, effects of repeated testing, changes in measurement procedures, statistical regression, attrition, or the interaction of these factors with each other or participant variables (Campbell & Stanley, 1963). The effect of maturation has traditionally been cited as of particular concern with younger age groups in which ongoing developmental process may result in acquisition of the target cognitive skill, without any specific intervention. In older groups, maturation may represent the normative age-related decline in cognitive functioning that occurs with no intervention; or in pathological conditions, the normative progression of the disease, such as Alzheimer's disease without a treatment.

A no-treatment control group is considered to be essential in early phases of studying a new intervention procedure to provide initial evidence of treatment efficacy. It is also less expensive than other comparison groups, because it involves no time or expenditures for alternative interventions or contact hours. In addition, the largest effect is often

obtained in comparisons of the intervention with the no-treatment group, hence allowing a reduction in the number of subjects needed for adequate statistical power.

There are some particular limitations of a no-treatment control that are related to the recruitment and maintenance of a sample, ethical considerations, and the information gained from the comparison (Hazlett-Stevens & Borkovec, 1999). First, if the problem being addressed is of particular concern to a certain segment of participants or the intervention is believed to be of considerable merit, then participants with these concerns may be reluctant to enroll in the trial, given the possibility of being randomized to the no-treatment group. Hence, the parent population from which the sample is recruited may be biased. For example, persons who believe themselves to have memory problems or who have recently been diagnosed with Alzheimer's disease may be less willing to risk being assigned to a no-treatment group and to make the commitment not to seek other treatment during the intervention trial. Assessment of the long-term maintenance of the intervention, furthermore, may be compromised due to no-treatment controls dropping out of the study as newer, promising treatments become available. Moreover, there is the ethical issue of not providing any treatment to participants with a known problem or deficit.

It is often stated that the no-treatment control group provides the least information regarding the processes or mechanisms associated with the intervention in relation to any other comparison group. There is, however, particularly in cognitive training research, a critical type of information that is best provided by a no-treatment control group. Specifically, the pre–post change in performance of the no-treatment control on cognitive outcome measures provides the best estimate of the magnitude of change that can accrue from repeated testing or being exposed to the assessment battery (see also Schaie, 1988, 1996, for discussion of practice effects in repeated testing in a longitudinal design). The magnitude of pre–post change shown by any comparison group other than the no-treatment control represents a confound of practice effects and the possible enhancement associated with the activity in which the comparison group was involved.

The magnitude of pre–post practice effects is not inconsequential when participants are low-educated elderly and have had little experience with the types of tests or assessment involved in psychological research. Simply becoming familiar with the assessment routine and practicing the tests can lead to significant improvement. In some of our studies of training on fluid abilities, the magnitude of retest effects has approached one-quarter of a standard deviation (Willis, Blieszner, & Baltes, 1981). The magnitude of practice effects or reactivity may vary for different mental abilities or assessment procedures. The largest practice effects are often shown for measures of abstract reasoning or speeded performance representing the fluid abilities that decline earlier and thus have been the target of many cognitive training studies. For example, verbal memory tests or inductive reasoning measures may exhibit larger retest effects than a vocabulary test, representing crystallized intelligence. Differences in reactivity among measures have also been reported in other behavioral interventions (Hazlett-Stevens & Borkovec, 1999).

Given the relative brevity of many cognitive training interventions, the time involved in the pre–post assessment battery may equal or exceed the time spent in the intervention itself. The likelihood of this occurring increases if multiple measures of an outcome are included in order to assess the target outcome at the construct level, as is currently supported from a

methodological perspective (Bentler, 1980; Schaie, 1996). Hence, participating in the pre–post battery becomes an intervention in itself! The challenge for the target training program, then, is to produce an effect that exceeds the performance improvement associated with practice and with increasing familiarity with the assessment routine. The impact of practice effects becomes of further concern with follow-up assessments conducted in order to examine the maintenance of the intervention. The follow-up assessment may serve as a mini-booster intervention, making it difficult to distinguish retest effects versus maintenance effects ascribed solely to the intervention. Additional control groups varying in number of assessment occasions would need to be added at each follow-up occasion to disentangle completely the retest effects versus maintenance.

Although practice effects are more obvious and can be directly measured for objective cognitive ability measures, reactivity effects also occur for other forms of assessment included in pre–post batteries (Hazlett-Stevens & Borkovec, 1999). For example, responding to questionnaires or interviews regarding their self-efficacy beliefs or metacognitive knowledge may increase older adults' awareness of such issues and lead to their further consideration or monitoring of these factors. Many researchers have had the experience, for example, of having an older participant return to a follow-up session and make statements, such as "I've been thinking about those questions you asked me regarding how often I cannot remember someone's name..."

B. Nonspecific Comparison Group

Since at least the 1950s (Rosenthal & Frank, 1956), placebo or nonspecific comparison groups have been given special attention in psychopharmacological and clinical trial interventions. Some have argued that placebo groups provide the only adequate control group and are the sine qua non of adequate experimental design (Klein, 1996; Rosenthal & Frank, 1956). Recently, there has been considerable discussion of the limitations of the traditional placebo control in psychologically oriented interventions (Brown, 1994; Kazdin, 1992; O'Leary & Borkovec, 1978; Strayhorn, 1987)

A brief discussion of the evolution of a placebo control is useful to understand the current debate (O'Leary & Borkovec, 1978; Rosenthal & Frank, 1956). From a historical perspective, there is evidence that many traditional medicines and cures were inert and probably were effective due to psychological placebo-type effects (Shapiro, 1971). Hence the necessity of a placebo control—known as the "sugar pill"—became an important design feature in current pharmacological research. The classical criteria for a placebo condition was that it involved a theoretically *inert* procedure that did not include factors considered critical to the intervention (O'Leary & Borkovec, 1978). A critical aspect of the criteria is that the placebo procedure is assumed to be *inert*. There was no theoretical rationale for why the placebo should result in the hypothesized positive outcome predicted for the treatment condition. A second critical aspect of the placebo effect was the assumption that the same type and level of *expectancies* regarding outcomes would be generated by participants in the treatment and in the placebo groups. The placebo and treatment conditions were presented to participants in such a manner that both groups should form the same expectancies regarding outcomes.

With the increase in psychological-based interventions and in interventions with both pharmaceutical and behavioral intervention arms, a placebo-type control was argued to be critical to the design of these types of interventions also (Parloff,

1986). The placebo-type control is sometimes known as the *nonspecific control* condition. The assumption was that a placebo-type control would account or control for nonspecific factors that were not considered critical to the treatment in behavioral or psychological interventions. The nonspecific-type factors assumed to be represented in a placebo-type control has grown to include variables such as social contact, personal contact with the trainer, expectancy effects, and opportunities for the participant to verbalize their concerns and to receive attention (Brown, 1994; Hazlett-Stevens & Borkovec, 1999; O'Leary & Borkovec, 1978).

There is a growing debate regarding the appropriateness of nonspecific, placebo-type controls for psychological and behavioral interventions (Hazlett-Stevens & Borkovec, 1999; Strayhorn, 1987; Wilkins, 1983). Conceptual, methodological, and ethical questions have been raised. First, it has been questioned whether a truly theoretically *inert* placebo procedure can be developed in psychologically oriented interventions. Nonspecific elements such as social contact, personal contact with the interventionist, and participation in multiple sessions are not psychologically inert. These nonspecific elements can have psychological effects that are not directly analogous to a pharmaceutical inert placebo, which can be assumed to have no beneficial physiological effect.

These nonspecific factors, moreover, are often considered to be necessary for the delivery of the critical theory-specific elements of the intervention (Parloff, 1986). For example, both in cognitive training and in therapeutic interventions, formation of some form of therapeutic alliance between interventionist and participant may enhance or be necessary in the delivery of the critical elements of the intervention (Hazlett-Stevens & Borkovec, 1999). Nonspecific factors such as personal contact with the trainer, positive feedback and support, development of trust and respect are often seen as salient aspects in the development of the therapeutic alliance. It might be argued that although factors such as a therapeutic alliance are considered salient in the traditional manner of treatment delivery, successful interventions have been administered via computers or the internet with less emphasis on a therapeutic alliance (Ball & Owsley, 2000; Finkel & Yesavage, 1989). However, research has also shown that treatment is often more effective when administered by an in-person trainer when recipients are cognitively low functioning individuals or have more severe pathologies (Rebok, Rasmusson, & Brandt, 1997).

A second concern regarding placebo-type controls in behavioral interventions is whether participants can be truly blinded to the assigned conditions and whether equivalent levels of expectancies can be created and maintained across conditions (Kazdin, 1992; Morin, Colecchi, Brink, & Astruc, 1995). The "sugar pill" in pharmaceutical interventions had much more face validity than can be created in many psychological interventions; moreover, recent advances in pharmaceuticals have resulted in reduced side effects such that the patient is even less able to judge the condition to which he or she was assigned (Brown, 1994). In contrast, in psychological interventions it is much more difficult to create a placebo condition that is equally credible compared to the treatment condition. Such credibility is based not only on the type of activities involved in the placebo condition, but also the ability of the intervener to generate and manifest enthusiasm and expectations that are equivalent in treatment and placebo. Many recent studies have involved placebo conditions that bore little similarity to the treatment of interest and thus further limited the likelihood of equivalent expectancies across groups (Clark et al., 1997; Tinetti et al., 1993).

Given these difficulties in developing comparable treatment and placebo–control conditions, some studies have reported a higher drop-out rate for participants assigned to placebo–control versus treatment (Nicholas, Wilson, & Goyen, 1991).

Maintaining the credibility of a placebo condition and of equivalent expectancies across groups becomes more difficult when the treatment is lengthy, which is more common when the intervention is dealing with serious, real-life problems. Expectancies are likely to dissipate across time in the placebo group if there are no effects or there are less effects over time (Brown, 1994). Likewise, maintaining equivalent motivation and enthusiasm in the intervener across groups becomes more challenging with lengthy interventions.

Relatively few cognitive training studies have included a social contact or placebo-type control; studies have more typically included a no-treatment or an alternative treatment condition. The few studies including such a control have generally found the mean performance of the treatment group to exceed that of the social contact control group (Clark et al., 1997). In our lab we compared cognitive training on attention and flexibility to a social contact group that met for an equivalent number of hours focusing on social support and friendship (Willis, Cornelius, Blow, & Baltes, 1983). The social contact group's performance on the target ability measures was significantly below the treatment group and did not differ from the no-treatment control.

The impact of placebo conditions has been more thoroughly examined in behavioral interventions focusing on psychological conditions such as depression, anxiety, and phobias. In these studies, placebo controls have shown greater effects when the intervention involved fewer sessions, when the participant was suffering from a less severe level of the condition, such as depression, and the episode of the condition was brief (Brown, 1994). Findings appear to be mixed from the few studies comparing the long-term maintenance of placebo effects versus treatment effects (Jacobson & Hollon, 1996; Rush, 1994).

C. Component Control Condition

A control condition that is considered by some researchers to avoid not only many of the limitations of the placebo control but also to provide additional valuable theoretical information regarding the nature of the treatment effect has been referred to as a *component comparison* approach (Hazlett-Stevens & Borkovec, 1999). An essential aspect of this approach is the ability to identify and to implement independently the various components of the total intervention package. The component approach is a between-group design in which various components of the intervention are implemented separately and in combination. This approach examines which specific elements of the intervention are responsible for particular observed changes by providing some participants with the entire intervention and some with only selected components.

This approach offers considerable flexibility to examine the optimal packaging of the treatment. Components can be combined in different sequences to examine whether ordering of components enhances magnitude of outcome effects. Moreover, the component control design can be achieved in two different ways. The researcher can approach the issue by systematically adding components until the whole intervention package is examined. Alternatively, the researcher can begin by first examining the entire intervention package and then selectively dismantling or deleting components. Again, this approach is only possible if the components of the intervention package are

well specified and can be easily componentialized (Basham, 1986).

The component approach avoids some of the limitations of the placebo or nonspecific control condition. Each treatment group includes one or more components that is considered to be efficacious and thus should lead to more equivalent expectancies across groups and enhance the credibility of each treatment group. Nonspecific factors are represented in all treatment components. This approach allows for increased specificity in studying cause and effect relationships, addressing specifically what aspects of the intervention are causing a certain outcome (Neely & Bäckman, 1995). In addition, there are less ethical concerns because all components are potentially efficacious.

This approach is represented in cognitive training studies in which imagery instruction and/or training in relaxation techniques were administered in combination with memory training on the method of loci (Gratzinger, Sheikh, Friedman, & Yesavage, 1990; Hill, Sheikh, & Yesavage, 1989; Kotler-Cope & Camp, 1990; Yesavage, 1990; Yesavage, Rose, & Bower, 1983; Zarit, Cole & Guider, 1981). In successive studies, imagery instruction, relaxation techniques, and both were paired with traditional method-of-loci training. Another componential approach has compared memory efficacy training alone or in combination with training on self-generated memory strategies (Lachman, Weaver, Bandura, Elliot, & Lewkowicz, 1992).

D. Parametric Design

This approach addresses the question of whether quantitative changes in one specific aspect of the intervention can increase the effectiveness of the intervention (Kazdin, 1992). An example of the parametric approach is systematically varying the dosage or number of training sessions to examine changes in the magnitude of training effects. The investigator can hold all experimental factors constant except for the one that is being manipulated (e.g., dosage, number of sessions). A unique strength of the parametric approach is that multiple levels of the variable can be examined. Quantitative increments in one specific aspect of the intervention may not be associated with linear changes or increases in the magnitude of treatment effects. For example, as the number of training sessions increases, there may be a diminishing return in terms of increases in treatment outcome.

There have been relatively few studies involving a parametric design approach in cognitive training research, and further utilization of this approach is needed. On the negative side, execution of this design is expensive in terms of factors such as the number of subjects needed and the number of treatment groups required to thoroughly examine incremental approaches in a treatment variable, such as number of training sessions. Hence, the cost and labor of conducting research using this design can be high. On the other hand, such a design has the potential to address some of the most salient cost–benefit and pragmatic issues in intervention research. This approach can address practical questions such as, What is the minimal number of training sessions required to produce an effect of a given magnitude? Can the treatment be administered as effectively in varying sized groups versus one-on-one? Over how long an interval is there maintenance of the outcomes? In the long term, such a systematic, incremental approach can provide important information on the relative cost and labor involved in an intervention when issues related to dissemination or generalization of the intervention arise.

Prior training studies have taken an incremental or parametric approach

primarily to various issues dealing with time and with size. A parametric approach has been used to address factors, such as number of training or practice sessions (Hofland, Willis, & Baltes, 1981); increased efficacy from delayed booster sessions (McDougall, 1999; Willis & Nesselroade, 1990; Willis & Schaie, 1994); varying the number of participants in training sessions, small group versus one-on-one (Rebok et al., 1997); variations in latency or response time (Baltes & Kliegl, 1992; Kliegl, Smith, & Baltes, 1989, 1990); long-term durability of treatment effects (Neely Bäckman, 1993; Scogin & Bienias, 1988; Willis & Nesselroade, 1990; Willis & Schaie, 1994); breadth of training transfer and number of measures used to assess effects (Willis & Schaie, 1986); and variations in level of competence and training of the intervener (Quayhagen & Quayhagen, 1989).

E. Comparative Design

A final type of design for intervention research involves contrasting two or more distinctly different intervention approaches to the same problem (Borkovec & Castonguay, 1998; Kazdin, 1986). The two or more interventions are assumed to represent different conceptual or theoretical views of the nature of the problem and of the optimal mechanisms for intervention. The key question is which of the existing approaches is more effective?

This approach is based on the implicit assumptions that there are two or more intervention approaches to the same problem and that initial research has shown both interventions to have some promise of efficacy. In clinical trial research on psychological disorders, the comparative design has been used to compare the efficacy of pharmacological versus behavioral interventions. This is sometimes described as an intervention with two arms—pharmacological and behavioral. Examples of this approach include clinical trials involving comparisons of the efficacy of pharmacological versus behavioral interventions for conditions such as depression (Jacobson & Hollon, 1996; Klein, 1996), panic disorder (National Institutes of Health, 1991), and insomnia (Morin, Culbert, & Schwartz, 1994).

At first glance, this might be thought to be one of the most common and useful types of intervention designs, but a number of problems and limitations in utilizing this design have been noted (Borkovec & Castonguay, 1998; Hazlett-Stevens & Borkovec, 1999; O'Leary & Borkovec, 1978). First, when comparing two established intervention programs, they often vary in so many ways that there are few variables that can be held constant across interventions. The interventions may vary in factors such as dosing (number of sessions), method of delivery of the program (e.g., delivered by person, computer, pill), and degree of social contact and/or therapeutic alliance. Although the two intervention programs may be modified to equate factors such as number of sessions, the question then arises of whether the modified program is truly comparable to the original version of the intervention that was shown to be effective. Eysenck (1994) has argued that the differences among intervention approaches in terms of treatment parameters and dependent measures stem from very important and often irreconcilable differences in scientific paradigms and methods of acquiring and interpreting knowledge. Such differences are reflected in important qualitative differences in treatment delivery and in the particular outcome measures considered to be most salient.

A second challenge to the comparative design focuses on internal validity issues arising from the need for comparable expertise and therapeutic allegiance of the trainers for each intervention approach (Borkovec, 1994; Jacobson & Hollon, 1996). Are the individuals who administer the interventions comparable in their

understanding and expertise across all interventions and does their allegiance vary across interventions? Administering each intervention at only one site may limit interpretation of results due to staff or sampling differences or problems. If each intervention is only administered at one site by the originators or adherents of the intervention, differences between interventions may be due to factors other than the efficacy of the intervention. Likewise, if each intervention is administered at only one site, sampling variations in the study population may make interpretation of results difficult. Ideally, each of the interventions to be compared should be administered at multiple sites by staff equated in level of expertise or allegiance across intervention approaches. Quality checks need to be made for variations in program implementation across the intervention trial (Hazlett-Stevens. & Borkovec, 1999).

A third challenge to the comparative approach is the development of a common assessment battery across interventions. Different approaches to the same problem often vary in the specific outcome measures deemed to be of particular salience. One approach may be more concerned with affective or self-report measures of the phenomenon, whereas another approach may focus on behavioral outcomes. *Reach* is a clinical trial currently in progress that involves the same core assessment battery (as well as site-specific measures) to compare different intervention approaches for caregivers of demented patients (Coon et al., 1999)

The comparative design approach has been less common in cognitive aging intervention research, particularly, with nondemented elderly, than might be expected. This may be due, in part, to the strong focus on strategies as the primary mechanism studied in cognitive intervention research. Because strategies are considered quite specific to a given cognitive ability or skill (Salthouse, 1991), comparison of distinctly different cognitive strategies for the same cognitive ability or process is less likely. It might be suggested that alternative interventions could be developed focusing on different mechanisms for cognitive deficits, such as limitations in perceived self-efficacy versus strategy use. However, these different approaches have more typically been viewed as separate components in a total intervention package (Lachman et al., 1992) and thus examined within a componential rather than comparative design approach.

As the range of pharmaceutical options for cognitive deficits increases, there may be an increase in outcome or evidence-based trials, with cognitive factors as outcomes that involve pharmaceutical and behavioral intervention arms. However, again, it can be argued that a componential approach involving a combination of pharmaceutical and behavioral interventions may be more efficacious rather than a comparative either–or approach. That is, it might be hypothesized that the optimal approach is to examine the efficacy of a pharmaceutical agent to stabilize cognitive functioning or to retard decline followed by a behavioral intervention to maximize the patient's potential to utilize the existing cognitive resources.

IV. Understanding the Mechanisms for Intervention Effectiveness

A primary goal of intervention studies is to produce a desired outcome—a change in behavior. An equally important objective, but one that is often not addressed adequately, is understanding how and *why* a certain outcome is achieved. A strong theoretical or conceptual framework for the intervention is critical in examining the processes or mechanisms underlying a change in behavior. The

conceptual framework specifies in detail the processes or mechanisms through which a given outcome is to occur. In a recent chapter on intervention research with the elderly, Schulz and Martire (1999) identified one of the most common shortcomings of existing intervention studies to be the failure to articulate a theoretical model that specified the mechanisms for achieving intervention outcomes. An associated deficit in intervention studies has been the lack of appropriate measurement of hypothesized mechanisms.

A. Cognitive Strategies

In cognitive training research, instruction on some form of cognitive strategy is hypothesized to be one of the primary mechanisms or processes by which change in cognitive behavior occurs (Charness, 1985; Kliegl, Smith & Baltes, 1989; Salthouse, 1991). A strategy can be defined as one of several alternative methods for performing a particular cognitive task (Salthouse, 1991). Verbal memory training studies have focused on strategies including the method of loci (Kliegl et al., 1989, 1990; Rebok & Balcerak, 1989; Yesavage, 1990), organization, visualization or imagery (Hill, Sheikh, & Yesavage, 1989; Yesavage, 1990; Zarit, Cole, & Guider, 1981), and formation of associations (Dunlosky & Hertzog, 1998a, in press). Camp's intervention work with demented elderly has utilized a technique or strategy known as spaced retrieval (Camp, 1999). Charness and colleagues have taught adults strategies for squaring two-digit numbers mentally (Charness & Campbell, 1988). Training on spatial orientation ability has focused on strategies facilitating mental rotation of objects, including identifying two salient features of the object and naming of abstract objects (Schaie & Willis, 1986). Our training research on inductive reasoning has involved strategies for identifying a serial pattern, including saying the pattern aloud, underlining repetitions in the pattern, and marking skips in a pattern (Willis & Schaie, 1986).

Study of the mechanisms underlying an outcome involves not only a conceptual framework that specifies the particular processes or mechanisms of interest but also developing procedures for independent assessment of the mechanisms (e.g., strategies; Dunlosky & Hertzog, in press; Saczynski, Willis, & Schaie, in press). Theory-guided intervention research requires assessing both whether use of the processes or strategies becomes more proficient or frequent as the intervention progresses, and whether increased usage of the strategies is associated with enhanced performance on training outcome (Saczynski et al., in press). For example, if utilization of the method of loci is the strategy hypothesized to enhance list learning, then (a) increased usage and/or improvement in the ease with which the loci strategy is implemented must be demonstrated, and (b) increased frequency or improvement in strategy use should be shown to be associated with recall of a greater number of words on the list.

A number of different questions regarding strategy usage may be important to address, depending on the nature of the strategy and outcome variables. Probably the most common strategy variable is frequency of usage of the strategy. Other aspects of strategy usage include the proficiency or speed with which the strategy is employed. Speed or proficiency of usage is important if the task is timed or speeded or if processing of significant information is involved.

B. Criteria for the Study of Strategies

Salthouse (1991) has specified a number of criteria for the study of strategies in cognitive aging. These criteria may be stated in the following manner when applied to intervention research. First, strategies are assumed to be specific to the outcome

that is the target of training. Second, all participants involved in the intervention are assumed to be capable of learning and executing the strategy. Third, the evidence provided to demonstrate use of strategies must be distinct from the measure of the outcome variable. Fourth, it is assumed that differences in performance on the outcome will be associated with strategy use. In training research, it is important first to show an increase in appropriate strategy usage from baseline to posttraining for the treatment group when compared with control groups. In addition, the increased strategy usage must be shown to account for significant variance in enhanced performance on the related outcome measure.

These assumptions regarding the strategies or mechanisms underlying the effectiveness of the intervention have important implications for intervention study design and methodology. The assumption that a strategy or mechanism is specific to a particular cognitive outcome has implications for specification of the hypothesized pattern of training transfer and for the measures selected to assess training outcomes. The issue of training transfer will be discussed in a later section of this chapter. It is important to note here, however, that the assumption that a particular strategy is specific to a given task implies an ability-specific model of training transfer. Training a particular strategy should result in improvement only on the cognitive ability or process with which the strategy is assumed to be associated.

There is increasing pressure to include outcome measures in cognitive training research that involve "real-world" problems that are often cognitively complex or that involve physical as well as cognitive processes. For example, medication compliance may be proposed as an outcome for a memory training program focusing on cognitive strategies and mnemonics. Park and colleagues (Park & Jones, 1997), however, have shown that medication compliance is a cognitively complex task that involves a variety of distinct memory processes, including working memory, verbal memory, and prospective memory. Moreover, for some individuals, medication compliance involves sensory and manual processes required to read the label or open the medicine bottle.

The research literature indicates that distinct memory strategies are associated with each type of memory process. The memory training program may be effective only for those aspects of medication compliance that involve the memory processes associated with the strategies trained. Thus, when specific strategies or mechanisms are the focus of training and hypothesized to account for intervention outcomes, it is critical that the outcome measures map carefully on the strategies and cognitive processes trained. Weaker training effects are to be expected for complex outcome measures involving cognitive processes that are not directly trained in the intervention.

The second and third criteria that all participants being compared are capable of learning and using the strategy and that level of performance on the outcome is associated with strategies also have implications for training programs. A strong interpretation of the second and third assumptions is the "strategy-as-cause" position, assuming that all participants are capable of learning and using the strategy regardless of their ability level (Salthouse, 1991). Performance on the outcome variable is moderated by strategy usage. Some researchers, however, have suggested that certain strategies place heavy demands on memory or other cognitive resources and hence may be beyond the capabilities of some participants (Charness, 1981, 1985; Finkel & Yesavage, 1989). In a related vein, investigators have suggested that strategy utilization may be associated with factors such as motivation,

efficacy beliefs, and social constraints (Cavanaugh, Krammer, Sinnott, Camp & Markley, 1985; Dunlosky & Hertzog, 1998a, in press). Yesavage and colleagues have found personality characteristics such as openness to experience related to learning of strategies associated with face–name recall (Gratzinger et al., 1990). An individual-differences approach to training would in this case be called for, with screening of participants for the required level of ability or other person characteristics deemed necessary to learn and use the strategy.

C. Factors Facilitating Learning of a Strategy and Strategy Usage

Cognitive aging research has found that many older adults do not spontaneously use appropriate strategies (Kausler, 1994). This finding may suggest that older adults find certain strategies difficult to learn and to use and thus need additional assistance in mastering a particular strategy. Alternatively, older adults may question the utility of the strategy or doubt their ability (i.e., efficacy) to use the strategy successfully. Finally, older adults may have difficulty determining the problems or contexts in which a particular strategy would be useful.

Some training studies have employed pretraining components to facilitate learning of strategies shown to be particularly difficult for older adults. The method of loci and forming associations through imagery are two strategies that have been shown to be highly effective in learning unrelated words, yet are difficult for older adults to master (Gratzinger et al., 1990; Hill et al., 1989). Two forms of pretraining have been used to facilitate the learning of these strategies; each form of pretraining focuses on a different hypothesis regarding the difficulty in learning the strategy. The first form of pretraining focuses on enhancing imagery skills. There is some support for a decline in imagery processes with age; moreover, some older adults question the utility of forming images, particularly fanciful images, or find it stressful (Verhaegen & Marcoen, 1994, 1996). Practice in imagery has been administered prior to memory training in work by Yesavage and colleagues (Hill et al., 1989; Yesavage, 1990). Similarly, pretraining in relaxation techniques has been employed prior to training (Hayslip, 1989; Yesavage, 1990). Relaxation techniques are employed to reduce the stress of using the imagery procedure or of learning a difficult strategy, such as method of loci. In a related vein, Bandura (1989) has suggested that affect is important in developing and maintaining self-efficacy. Induced positive mood has been shown to enhance perceived self-efficacy, whereas despondent mood diminishes perceived self-efficacy (Hertzog, McGuire & Lineweaver, 1998; Kavanagh & Bower, 1985; Berry, West, & Dennehey, 1989).

If the treatment outcome variable is cognitively complex or involves multiple components, the participant may be required to determine for which outcome measures the strategy trained is appropriate and useful (Dunlosky & Hertzog, 1998b). For example, if the list of words is unrelated, then the method of loci may be a more productive strategy. In contrast, with a list of related words, formation of meaningful categories may be more effective. The intervention procedure in this case may need to train not only on the specific strategies but also give guidance in determining in what instances a given strategy is likely to be most useful. The intervention would then involve not only training on a specific strategy but enhancing executive or metacognitive processes (see also Salomon & Perkins, 1989), which are hypothesized to affect the selection and monitoring of strategies over a wider range of cognitive tasks. The intervention would require training of metacognitive strategies or skills, on

which there is much less empirical research, and also provide practice not only on problems for which the strategy is relevant but also problems for which the strategy is not relevant (McKeough, Lupart, & Marini, 1995).

These higher order skills and metacognition have been discussed as a part of self-regulated learning and memory monitoring (Dunlosky & Hertzog, 1998a). Bandura (1989) has suggested that development of self-regulatory capabilities requires instilling a resilient sense of efficacy as well as imparting skill in using a given strategy. If the elderly are not fully convinced of their personal efficacy, they rapidly abandon the strategy they have been taught when they fail to get quick results or it requires bothersome effort.

The mnemonic training research of Rebok and Balcerak (1989) on the method of loci found use of the strategy improved the memory performance of older adults but did not raise their beliefs in their memory efficacy. The lack of an association between the strategy and self-efficacy may explain why only 39% of participants used method of loci during generalization tests of memory for digits. In contrast, younger adults whose self-efficacy increased as a result of mnemonic training spontaneously used the loci aid in generalization memory tasks.

Bandura argues that training in cognitive strategies can produce more generalized and lasting effects if self-efficacy beliefs are increased and participants see an association between strategy use and increased control of their memory. Bandura (1989) sees direct mastery experiences as a particularly effective way of building efficacy beliefs. Participants perform memory tasks with and without mnemonic aids and compare the results. Evidence of better memory performance with mnemonic aids provides participants with persuasive demonstrations that they can exercise some control over their memory by enlisting cognitive strategies. Such efficacy-validating trials not only serve as efficacy builders, but also put on trial the value of the techniques being taught.

D. Measurement of Strategy Usage

As noted by Schulz (Schulz & Martire, 1999), prior behavior intervention research has often lacked not only specification of the process or mechanisms underlying the intervention, but also distinct measurement of the targeted strategies. There have been several common procedures for determining strategy usage, each having limitations (Dunlosky & Hertzog, 1998a). The most common and simplest procedure for determining strategy usage has been to ask participants to report on the strategy used after they have completed the task (Cohen & Faulkner, 1986; Rice & Meyer, 1986). The validity of these reports is unknown, as there is often no means of verifying the accuracy of the self-reports. In a recent study, Dunlosky and Hertzog (in press) found that retrospective reports were not completely consistent with concurrent reports, suggesting that the validity of retrospective reports is somewhat diminished by forgetting, particularly in older adults. Other procedures for assessing strategy usage include thinking aloud and then analyzing recordings for indications of strategy use. Analyzing time allocated to each portion of a sequential task (Salthouse & Prill, 1987) can also provide evidence of strategy use. The distribution of these times forms a profile of processing durations, which can be considered a reflection of the strategy used.

In our training research on inductive reasoning ability, the strategies trained provide an objective record of strategy use. Participants are taught to mark the patterns in reasoning problems with specific types of markings to indicate pattern repetitions, pattern skips, and pattern replications (Willis & Schaie, 1986).

Reliable instances of pattern usage at pre- and posttest were coded. Significant pre–posttest increases in pattern usage were shown for the reasoning training group compared to participants trained on spatial orientation. Moreover, increases in strategy usage accounted for significant variance in factor scores of reasoning ability performance (Saczynski, et al. in press).

V. Levels of Outcome and Mediators

As stated at the beginning of this chapter, a primary goal of evidence-based interventions or of clinical trials in the behavioral sciences is to examine whether treatment results in a greater improvement for the intervention group on the outcome measures than for the comparison groups. Although treatment effects are often stated in terms of improvement, the goal of treatment may alternatively involve a reduction in negative outcomes (e.g., memory complaints, incidents of forgetting) or even an outcome of stability, maintenance, or consistency (e.g., maintaining a certain level of accuracy and speed in a driving task; no increase in visits to emergency rooms or need for home health services).

Most behavioral interventions, particularly those with some commitment to real-world problems, will consider multiple domains of participants' functioning in assessing outcomes. First, real-world problems rarely involve only a single domain of an individual's functioning. Second, psychosocial phenomena (e.g., cognition, personality, efficacy, affect) that are of interest to social scientists and that are likely to be the target of psychological behavioral interventions are most adequately assessed in multiple domains (at the neural or physiological level, through behavior, and through perceptions). Third, an intervention may affect not only the cognitive ability or skill that is the target of the intervention but may result in transfer to other domains of the individual's life or other aspects of behavior and perception. Thus, measurement of intervention outcomes often involves multiple levels of outcome and involves outcomes that vary in how close they map on the process, skill, or ability that was the specific target of the intervention.

The importance of a strong theoretical or conceptual framework to guide all major aspects of a behavioral intervention program is a crucial principle stressed throughout this chapter and in the work and writings of other researchers in the field (Baltes, 1987; Borkovec, 1994; Camp, 1999; Lerner, 1986; Schulz & Martare, 1999). An adequate conceptual framework often begins with a basic theory of some aspect of life span development and then is extended and enriched by decades of descriptive and intervention research. The conceptual framework needs to be viewed as dynamic, a work in progress, rather than a theory established by a former giant in the field (or at an earlier stage in the researcher's own professional career) and set in stone. Ideally, a developmental theory or framework *describes* the life span developmental trajectory of a phenomenon, articulates possible *explanatory mechanisms* for development and change in the phenomenon, and based on these components of description and explanation, offers insight regarding the plasticity or *modifiability* of the phenomenon, either across the life span or at specific developmental periods (Baltes & Willis, 1977; Bandura, 1989; Lerner, 1986; Schaie, 2000).

At least three major levels of outcomes are generally recognized in behavioral intervention research: proximal, primary, and distal outcomes. The remaining part of this section will deal with a discussion of these three levels of outcomes.

A. Proximal Intervention Outcome

The proximal outcome in cognitive training research is the key cognitive process, skill, or ability that is the target of training. The proximal outcome represents the nearest level of training transfer. In our intervention research on fluid intelligence, the abilities of figural relations, inductive reasoning, or spatial orientation have been the proximal outcomes (Willis et al., 1981; Willis & Schaie, 1986). In memory intervention research, the proximal outcomes have been variables, such as recall of long word lists, recall of face–name pairs, or recall of text material (Kliegl et al., 1990; Yesavage, 1990). Most cognitive training research has focused almost exclusively on training effects and training transfer at the level of proximal outcome measures. There has been relatively little attention given to the remaining two levels of intervention outcomes, primary and distal outcomes, as discussed below.

An issue that has been examined with regard to proximal outcomes in cognitive training research has been breadth of transfer within the proximal outcome domain. Cognitive training studies have been subject to criticism of "teaching the test" when a single measure of proximal outcome was employed. Training studies have increasingly employed multiple measures of the proximal outcome (Willis & Schaie, 1986). If training effects are shown for multiple measures of the proximal outcome or if the proximal outcome is represented in terms of factor scores, then it can be argued that training effects have been demonstrated at the latent construct level. That is, the shared variance among multiple measures of the proximal outcome has been impacted by the intervention (Schaie, Willis, Hertzog & Schulenberg, 1987).

The proximal outcome must be specified in a consistent, congruent manner in at least three places in the reporting of a study (a) the statement of the objectives of the intervention, (b) the tests or measures used to assess the target of training, and (c) the intervention protocol.

An unfortunate common mistake in interventions is a lack of congruence and specificity in description of the proximal outcome in the aims, measures, and treatment protocol. The aims or goals of the intervention are often stated in much more general or broad terms than what is actually measured by the tests or questionnaire employed. Likewise, a careful examination of the treatment protocol or training manual often indicates that relatively little time is spent focusing directly on the target ability or cognitive skill—or that the treatment protocol is not specified in the detail required to determine whether the majority of the time is spent in practice or training on the target skill or ability.

1. *Mechanisms and Cognitive Strategies: The Mediators of Proximal Outcomes*

As discussed in the previous section on cognitive strategies, it is critical that the conceptual framework for the intervention specify the types of mechanisms that are hypothesized to change performance on the proximal outcome or training target. In much cognitive training research, the key mechanisms are cognitive strategies. In training a skill, a series of skill components may be specified. An important point is that, at least in cognitive interventions, the mechanisms, strategies, or skill components are likely to be highly specific to the ability or skill (proximal outcome) being trained. Unique strategies or skill components have usually been identified for a given ability or skill. Thus, mechanisms (e.g., strategies) and proximal outcomes go hand-in-glove.

Another common mistake in cognitive training research is that the particular mechanisms to enhance the proximal

outcome are not clearly specified and are not the primary focus of the training protocol. To a large extent, the strategies, skill components, or mechanisms taught during training define the proximal outcome. If the proximal outcome is complex, then multiple strategies or skill components may need to be taught and proportional time devoted to each strategy in accord with its hypothesized salience to the proximal outcome as a totality.

The mechanisms or strategies specified to underlie the proximal outcomes then become a critical mediator of change in the proximal outcome. Enhancement in performance on the proximal outcome should occur as a function of an increase in frequency of strategy usage and/or increased expertise in applying or executing the strategy. Thus, competence with and use of the key mechanism or strategy should account for significant variance in individual differences in improvement on the proximal outcome (Dunlosky & Hertzog, 1998, 1998b; Saczynski, Willis, & Schaie, in press).

B. Primary Intervention Outcomes

The second level of intervention outcomes is often known as primary outcomes. Primary outcomes should be a product of the proximal outcome, share significant common variance with the proximal outcome, or both. At least two very significant criteria must be met in order for there to be the possibility of significant training transfer at the primary outcome level: (a) there must be evidence of considerable shared variance between the proximal and primary outcomes, and (b) a significant training effect should be demonstrated at the level of the proximal outcome. That is, in order for a training effect at the level of the primary outcome to be ascribed to the treatment, a training effect must first be demonstrated for the proximal outcome that was the target of the intervention (Salomon & Perkins, 1989).

For example, in our cognitive training research on fluid abilities, one possible primary outcome would be enhanced performance on cognitively demanding tasks of daily living (comprehending medication labels, interpreting phone bills, etc). Prior research has shown a significant association between fluid abilities such as inductive reasoning and cognitively demanding tasks of daily living (Willis, 1996). Moreover, in lagged analyses, fluid and crystallized abilities measured at the first occasion have been shown to predict performance on cognitively demanding tasks of daily living 7 years later (Willis, Jay, Diehl, & Marsiske, 1992). In addition, it would need to be shown that participants who showed significant training on fluid ability were more likely to show reliable improvement on measures of cognitively demanding tasks of daily living.

C. Distal Intervention Outcomes

The third level of intervention outcomes are distal or secondary outcomes. In terms of training transfer, these would be referred to as far transfer. These are the least studied of the three levels of intervention outcomes. It is generally argued that if distal intervention outcomes occur, they are the result of "spill over" from treatment-related change in primary and proximal outcomes. For example, enhanced performance on cognitively demanding tasks of daily living might lead to distal outcomes, such as improved medication adherence or maintenance of the current level of medication adherence as the individual ages. Theoretical rationales for how distal intervention outcomes would occur are not well developed and need considerable further study. Bandura's (1986) theory of the role of self-efficacy at multiple phases in an intervention is one exemplar of the type

of theory and research that needs to be done in this area. Also, the discussion of far transfer by Salomon and Perkins (1989) reviewed in the next section of this chapter may provide important insights.

VI. Issues in Training Transfer

Transfer is a central concept in learning theory and in training research that has been studied and debated over much of the past century (Detterman & Sternberg, 1993). Discussion regarding the construct is usually traced back as far as the writings of Thorndike and Woodworth (1901). Although issues related to transfer have been occasionally addressed in cognitive training studies with the elderly (Donaldson, 1981; Fisk, Rogers, Cooper, & Gilbert, 1997; Willis & Baltes, 1981; Willis, 1987, 1990), discussion of the history and assumptions regarding transfer have received relatively limited attention in cognitive aging. However, transfer is becoming a critical issue, as large-scale behavioral clinical trials targeting psychological and social constructs become more frequent in gerontology. These trials often examine whether intervention into the cognitive and social constructs studied by gerontologists have implications for maintaining (i.e., transferring to) competence in activities associated with the health and independence of the elderly (Coon et al., 1999; Jobe et al., in press).

This section begins with a brief review of the broader transfer literature (Cormier & Hagman, 1987; Detterman & Sternberg, 1993; McKeough et al., 1995; Voss, 1990), followed by a more in-depth discussion of the literature on mechanisms to foster transfer and the specification of a continuum of transfer. It is important to note that although transfer has been extensively studied and debated for over a decade, there is still often limited consensus on critical aspects of the construct (Detterman & Sternberg, 1993).

A. Defining Transfer and a Brief History

The first task is to define transfer as a construct. Transfer is said to occur when learning in one context enhances performance in a somewhat different context (Salomon & Perkins, 1989), or transfer occurs when prior learned knowledge and skill affect the way in which new knowledge and skills are learned and performed (Cormier & Hagman, 1987). Detterman (1993) has defined transfer as the degree to which a behavior will be repeated in a new situation. Transfer refers to a recognition that various terms and entities of one set can be mapped onto those of another set (Ceci & Ruiz, 1993). In job training, transfer has been defined as the application in the workplace of the knowledge, skills, and attitudes learned in training (Yelon, 1992). Although definitions of transfer differ, there is a common often unstated but critical assumption: A prerequisite for transfer is that some form of initial learning has occurred—there must be at least a minimal mastery of information or skills for the opportunity for transfer to exist (Salomon & Perkins, 1989)

The concept of transfer is closely tied to that of learning, and hence as various theories of learning have come and gone over the past century, definitions of transfer have evolved and changed (Voss, 1990). Ferguson (1956) argued that if improvements occur from learning trial to learning trial, then transfer had occurred between trials; this would be a very near form of transfer. At the time of Thorndike's and Woodworth's (1901) early research on transfer, the dominant theory of human learning was associationism. Early explanations of transfer were, thus, typically based on the theory of identical elements and couched in stimulus–response (S–R) language (Cormier &

Hagman, 1987). Thorndike concluded that when transfer occurs it occurs because of common elements in the two situations. The amount of transfer that occurs can be predicted from the proportion of common elements shared by two situations. Given this orientation, Thorndike concluded that only near transfer is likely to occur.

On the other hand, throughout the history of the study of intelligence, a hallmark of intelligence and of an intelligent person has been the ability to think abstractly and to derive general principles from concrete exemplars. The core of intelligence, *g*, from Cattell's notion of fluid intelligence (Cattell, 1963) to Sternberg's triarchic theory (Sternberg, 1985) has been described in terms of linear reasoning, thinking abstractly, and making inductions. Intelligence involves the ability to adapt to new situations. A major adaptive mechanism of the human species is the ability to profit from experience. Humans learn to abstract relevant knowledge and skills from prior experiences to their advantage in new or novel situations. They *transfer* the essence of knowledge, skills, and principles acquired in prior contexts to how they think andbehave in sometimes dramatically new situations.

These two different views regarding transfer are reflected in two classic theories of education (Ceci & Ruiz, 1993; Detterman, 1993) that continue to have followers into the present day. Early in the 20th century the doctrine of formal disciplines was a dominant educational approach. This approach held that training in one discipline enabled one to think more effectively and hopefully abstractly in many other disciplines. For example, specific training in Latin or chess was regarded as an exercise that fostered the development of logical reasoning in general. Learning how to reason in one context was thought to transfer to reasoning in other contexts. Recent expressions of this approach can be found in research that touts the benefits of learning computer programming languages to develop rigorous thinking, to learn the use of heuristics, and to teach the process of problem solving (Salomon & Perkins, 1989). The research of Thorndike arguing for only near transfer was in response to the widespread acceptance of the formal disciplines approach in the early 20th century.

The second educational approach has focused on domain-specific learning—explicitly teaching information, principles, and strategies within each domain. The individual can learn to transfer knowledge and principles to multiple situations *within* a given substantive domain. However, training in a specific domain such as Latin is not likely to result in enhanced reasoning and understanding in another domain such as physics (Detterman, 1993).

B. Four Aspects or Dimensions of Transfer

Major issues studied and debated throughout the 20th century regarding transfer can be summarized in terms of four aspects or dimensions: (a) The *how* or *mechanisms* of transfer; (b) *what* is transferred; (c) the *amount* of transfer; and (d) the *distance* or *breadth* of transfer.

1. The How or Mechanisms of Transfer

A major question regarding transfer, even near transfer, has focused on the mechanisms by which transfer occurs. Salomon and Perkins (1989) have proposed that transfer can occur by different routes or mechanisms or combinations of mechanisms. They hypothesize two major internal mechanisms by which transfer can occur: low-road transfer and high-road transfer. Because of the unintentional "negative transfer" associated

with the term *low-road transfer,* in this discussion, the two mechanisms are referred to as *transfer due to automaticity and transfer due to mindful abstraction.*

Transfer due to automaticity involves the processes of (a) extensive practice in varied contexts, (b) stimulus control, and (c) automaticity. A skill or other cognitive element is learned and practiced in a variety of contexts until it becomes automated. If practiced in a variety of somewhat related and expanding contexts, execution of the skill or cognitive element can become increasingly flexible. Automatization occurs as a result of extensive practice (Shiffrin & Schneider, 1977). The processing and behavior becomes fast, effortless, and relatively unlimited by processing capacity. The behaviors and cognitions are stimulus controlled. One's cognitive system automatically applies the learned behavior whenever it identifies situational cues it takes to be prototypical of a particular category of situations.

A plus of this form of transfer is that it usually increases the efficiency of the behavior; the behavior is performed fast, effortlessly, and with reduced processing demands. The limitation of this mechanism for transfer is that automaticity inhibits analytic reflection. Conscious control and analytic awareness is reduced by automaticity.

In an extensive, ongoing, well-designed program of research, Fisk, Rogers, and colleagues are conducting training studies with older adults that address many issues related to automaticity as a mechanism for transfer (Fisk, Lee, & Rogers, 1991; Fisk, Rogers, & Giambra, 1991; Fisk et al. 1997).

Transfer due to mindful abstraction is the second mechanism of training transfer proposed by Salomon and Perkins (1989). The processes involved in this mechanism of transfer are (a) mindful, deliberate, deep processing of information (Langer, 1989) and (b) the abstraction or decontextualization of cognitions. Abstractions generally take the form of a rule, principle, schema, or prototype. Abstraction is the principle by which transfer occurs; abstraction provides a bridge from one context to another. Formation of the abstraction is a mindful, deliberate process that typically occurs during the learning process (although see the backward-reaching process next).

Salomon and Perkins (1989) suggest two alternative ways in which mindful abstraction may lead to transfer. The first type is called *forward reaching.* In this case, the participant mindfully abstracts basic elements during the initial learning situation in anticipation of making a later application. The information is learned or encoded as a general principle in the initial learning, and new applications of the general principle occur almost spontaneously in later situations (transfer). In the second type called *backward reaching,* the abstraction actually occurs in the transfer situation rather than in the initial learning situation. One is faced with a new situation and deliberately searches for relevant information from prior situations that might be applicable in the new transfer context. The principle is learned originally for a context-specific purpose, but the individual at a later occasion is able to reformulate the information or principle to a higher level of abstraction.

Transfer due to mindful abstraction probably better characterizes the training process conducted in training on fluid abilities (Schaie & Willis, 1986) and perhaps in application of some memory strategies (Dunlosky & Hertzog, in press; Hill et al., 1989; Rasmussion Rebok, & Brandt, 1999). The elder is trained on general rules or strategies (e.g., looking for certain types of patterns in inductive reasoning problems, applying method of loci to list of words) and is given practice in applying (transferring) these rules or strategies to

new instances of the problem. As discussed in the section on cognitive strategies, the participant must not only learn to apply the rule or strategy to the next instance of the problem, but also must determine whether or not a given strategy applies or which strategy applies when faced with problems varying in the specific cognitive strategy needed to solve a problem.

2. *What Is Transferred*

As noted above, the concepts of learning and transfer are closely intertwined. *What* is hypothesized to be transferred will depend on the researcher's or interventionist's theories of learning and of intelligence or cognition. In an associative approach to learning, transfer is viewed as being narrow or near to the initial learning and context. *What* is transferred is defined in S-R terms, in terms of identical or similar elements between the learning and the transfer situation. In contrast, theories of intelligence and learning approaches that focus on concepts such as learning-to-learn, procedural knowledge, and expertise are concerned with the ability of the individual to form abstractions, develop hierarchies of knowledge, and identify general principles. This approach should lead to broader transfer across contexts. *What* is transferred is more likely to be described in terms of a subroutine, learning strategy, overarching principle, or generalized skill.

Moreover, Salomon and Perkins (1989) have argued that *what* is transferred interacts with the *mechanisms* of transfer. What is transferred through the mechanism of automaticity and practice is hypothesized to involve behavior that is unintentional, implicit, based on modeling, and driven by reinforcement. These activities are involved in processes such as socialization, acculturation, and experience-based cognitive development.

In contrast, transfer via mindful abstraction is more likely to occur during explicit instruction that is aimed at teaching or provoking the learner to identify a generalization. Transfer results from the *mindful* generation of an abstraction developed during the learning process (Langer, 1989). Teaching strategies and mnemonics to mildly retarded children was effective only when skills of mindful attention and metacognition were taught along with the strategies and mnemonics (Brown & Kane, 1988; Campione, Shapiro, & Brown, 1995).

3. *Amount of Transfer*

Amount of transfer refers to how much improvement results in the transfer context from attaining some level of performance in the learning context. Salomon and Perkins (1989) hypothesized that the extensive practice involved in transfer due to automaticity affects mainly the *amount* of transfer — practice leads to the automatic activation of whole "bundles" of interrelated responses. Because of bundling, amount of transfer is likely to involve the entire set of skill components or the entire knowledge set, rather than selected components of a skill or fragments of the knowledge set. For example, in transferring driving skill from one car to another, the entire repertoire of driving skills is transferred, even though certain skill components (e.g., reaching for the clutch on a car with automatic transmission) might not be needed or appropriate.

What about amount of transfer from a mindful abstraction approach to transfer? The goal of this type of transfer is to form an abstraction or generalization that can be applied in new situations. However, the amount of transfer may depend in part on the fit or match between the level of abstraction at which a principle is learned and the level of abstraction required in a particular transfer situation. If

abstraction occurs at too high a level in the initial learning, then it may interfere with applying the principle to a transfer situation. The principle is not encoded at the level of concreteness that would most easily facilitate transfer to a particular context. For example, a traveler learns a precise algebraic formula for converting Euros to dollars and vice versa. However, during a particular interval when one Euro is almost equivalent to one dollar, applying the precise algebraic formula to determine the price of an inexpensive item may be overkill.

Addressing the issue of the amount of transfer involves two questions: at which of the levels of intervention outcomes, discussed previously, is amount of transfer most directly studied and upon what factors might the amount of transfer depend? Amount of transfer refers to *how much improvement* results in the transfer context as a result of the level of mastery attained in the learning context. In terms of the previous section on levels of outcome, the *primary outcomes* are of particular concern with respect to transfer. The central question is whether pre–posttest gain in the primary outcomes for the treatment group exceeds the pre–posttest gain in the primary outcomes for the comparison groups.

The second question that focuses on what factors might be associated with variability in the amount of transfer attained is complex. It depends in part on factors such as the investigator's theory of learning, the mechanisms of transfer, and perhaps participant characteristics. Such factors can only be briefly considered here. Based on the discussion above, Salomon and Perkins (1989) suggest that automaticity as the mechanisms of transfer may lead to a high incidence of transfer, but that the breadth of transfer will be more specific or narrow. That is, once a skill is automated, there should be a high likelihood that the skill will be performed in a transfer situation if the stimulus cues are present, and that it will be performed nearly flawlessly.

With regard to amount of transfer and characteristics of the trainee, there is ongoing debate on the relationship between transfer and intelligence (Ceci & Ruiz, 1993; Detterman, 1993; Singley & Anderson, 1989; Sternberg & Frensch, 1993). The debate appears to focus primarily on far (general) transfer, rather than on near transfer. There seems to be more agreement that lower cognitively functioning individuals are less likely to show near transfer. This finding has demonstrated primarily with young mildly retarded individuals (Brown & Kane, 1988), although some evidence with older adults (compared with younger adults) is also reported by Fisk and associates (1997). The more debatable issue across the life span appears to be the extent to which far transfer can be trained and the extent to which successful training in far transfer would reduce broad cognitive deficits. Several educational programs based on this premise are ongoing with young or retarded children (Campione, Shapiro & Brown, 1995) and with young adults (McKeough et al., 1995).

4. Distance or Breadth of Transfer

Distance of transfer concerns how far learning transfers and to which tasks: very similar tasks, somewhat related ones, or even quite remote tasks. Transfer can be conceptualized as a continuum of situations progressively more different from the original learning experience. The major debate in training centers on how broad or far across the continuum transfer should be expected. Two general types of transfer are often distinguished; these two types of transfer have been described in terms of near/far, specific/nonspecific, and deep/surface structure (Cormier & Hagman, 1987; Detterman, 1993; Mayer & Greeno, 1972).

Near transfer occurs when the skills or abilities acquired in training are demonstrated in situations that are similar to the original learning situation except for a few important differences. The more different the original versus new situations, the more likely transfer is to be called *far* transfer.

Salomon and Perkins (1989) have made some hypotheses regarding near versus far transfer when transfer involves mechanisms of automaticity versus mindful abstraction. In transfer associated with automaticity, the individual does not consciously analyze a new situation in terms of similarities or differences compared with the initial learning situation. Similarities of prototypical cues across contexts are detected automatically. Thus, transfer to new situations occurs primarily if these situations activate these response clusters because of automatically detected similarity of prototypical cues (Schneider & Fisk, 1984). Transfer to situations more remote from the learning context are less likely because identification of similarities would require intentional (conscious) examination of the similarities. As discussed under the section on amount of transfer, partial components of a skill or fragments of knowledge that might be relevant in a new situation are not likely to be transferred, because the complete skill or knowledge set is bundled and activated as a totality.

What of near versus far transfer is associated with mindful abstraction? While far or distant transfer appears less likely via the mechanism of automaticity, transfer to more distant contexts is hypothesized to be more likely to occur via the mechanism of mindful abstraction. The higher the level of generality at which a principle is abstracted, the broader or more distal the range of situations to which it might be possible to apply the principle.

Although the near–far distinction is most common in discussions of the breadth or distance of transfer, several other types of contrasts have been noted. In *specific* transfer the learner transfers the contents of learning to a new situations. In *nonspecific* transfer, general skills or principles transfer to the new situations. More recently a distinction has been made between the *deep structure* and *surface structure* in a situation. An example is that all car dashboards give similar information but their dial configurations are different. Deep structure is the same but surface structure is different. On the other hand, airplane dashboards contain dials similar to a car's, but the information presented by the dials is different. For the car and plane dashboards, there is similar surface structure but different deep structure.

In terms of the three levels of training outcomes discussed in the previous section, proximal outcomes would be considered to represent very near or specific transfer. *Primary* outcomes would be considered to include moderately far transfer further along the continuum involving the transfer of either general skills or principles or transfer to markedly different situations. Distal outcomes would represent the furthest form of transfer. In training, the greatest interest has been in whether moderately far (i.e., primary outcomes), nonspecific, or deep structure transfer could be achieved and what factors enhance or impeded such transfer. Whether transfer of general principles can occur between different situations, as examined in primary outcomes, is particularly important in training research.

VII. Summary and Future Directions

The background for this chapter lies in the significant increase in behavioral intervention research with the elderly that has occurred over the last three decades.

Behavioral interventions with the elderly have evolved from experimental research conducted by a single investigator in his or her laboratory to large-scale multisite studies (Schulz, Maddox, & Lawton, 1999). Behavioral intervention research with the elderly is increasingly represented in the clinical trial arena (Coon et al., 1999; Jobe et al., in press). Journal articles and book chapters reporting the substantive findings from these studies are beginning to appear in the literature. An aim of this chapter was to review and discuss selected methodological issues with respect to behavioral intervention research with the elderly. The cognitive training research literature provides many of the exemplars cited throughout the chapter.

A plea is made throughout the chapter that strong conceptual and theoretical frameworks should continue to guide the future evolution of behavioral intervention research. In the first generation of behavioral intervention research, the single investigator typically had considerable knowledge of the existing descriptive and explanatory models of the psychology of aging and grounded the intervention protocol and measurement system within these models. Thus, not only was the issue of invention outcomes addressed, but the training research made significant contributions to the broader psychological aging literature. A reciprocal relationship existed; gerontological behavioral interventions were deeply rooted in developmental and experimental theories of aging, and intervention findings contributed to further theory development.

Although the recent and ongoing large-scale, multisite behavioral intervention studies will make significant contributions to issues regarding the generalizability and representativeness of the earlier training literature, there is also the concern that the reciprocal contributing relationship between gerontological descriptive and experimental research and behavioral interventions may diminish. It is essential that the strong, almost singular emphasis on outcomes and treatment effectiveness that is characteristic of outcome intervention research and in clinical trials be complemented and even tempered by an equally strong commitment to the grounding of the intervention in a conceptual framework—with theory-based assumptions and hypotheses required and made explicit. Moreover, the role of the conceptual framework must be reviewed at each stage in the development and implementation of the intervention. Furthermore, it is important that the methodological procedures and assumptions upon which outcome interventions and clinical trials are based, which are often rooted in pharmacological intervention research, be reexamined and their applicability reevaluated as they are extended to behavioral interventions with the elderly. Behavioral intervention research must not lose its roots and heritage in the psychology of aging, both substantively and methodologically. Each branch of research must continue to maintain a dialogue and contribute to their mutual development.

In the first section of this chapter, the major between-group designs employed in behavioral intervention research were reviewed and their relative utility examined. The between-group perspective is based on examining the treatment efficacy of various types of control or comparison groups, compared to the intervention group. These control groups have included no-treatment control, a nonspecific effects group, component approach, alternative treatment groups, and a parametric approach. Although each alternative has important strengths and limitations, the component approach appears at this time to be the most flexible approach and to be most likely to provide further theory-based knowledge regarding the intervention. Many behavioral interventions began, in honesty, with the "kitchen sink" approach, including

a wide array of factors that might contribute to a positive training effect. Having shown significant effects with the total-package intervention, it is now important to systematically examine the relative contribution of each component. In addition, the parametric approach, although labor intensive and costly in subjects, has the potential to provide critical information in the form of a cost–benefit analysis of the relative gain achieved by increasing or decreasing the amount or intensity of various components of the intervention.

The second section of the chapter considered the essential role of mechanisms, processes, and strategies in behavioral interventions with the elderly. The field of cognitive aging is noteworthy for the theoretical and research efforts made to identify and understand the mechanisms underlying cognitive functioning (Baltes, 1987; Craik & Salthouse, 1992; Kausler, 1994). The early cognitive training research is to be commended for grounding the intervention protocols in conceptual frameworks derived from the cognitive aging literature. Prior training research has been successful in teaching older adults that a particular strategy is useful for a particular cognitive task and how to use the strategy. It is becoming evident that knowing how to use a skill in a highly structured training session does not equate with older adults accepting the efficacy of the skill or being able to determine in what contexts in the real world such a strategy is appropriate and useful. The next generation of training studies will need to take into account to a greater extent the role of context and factors such as motivation and efficacy. Perhaps even more challenging is the likelihood that facilitating the older adult's use of a particular cognitive strategy in appropriate real-world contexts may require some form of executive or metacognitive training in addition to strategy-specific instruction.

Concern with the possible need for training at the executive or metacognitive level led to the final section on training transfer. It bears repeating that transfer is one of the most challenging phenomena in the field of learning and training and one for which there is relatively little consensus. Cognitive aging research can make a unique and very valuable contribution to research on transfer. Many of the major concepts discussed in the transfer literature, including automaticity, prior knowledge, and abstraction, are areas of important current research in cognitive aging. Moreover, there appear to be quantitative and/or qualitative changes in these phenomena with age that may well provide insight into the role of transfer in learning and training.

In summary, the field of behavior intervention research with the elderly is relatively young, but it has experienced a period of near exponential growth in the last two decades, from single-investigator studies involving less than 100 participants to large-scale studies involving several thousand older adults. Given the movement toward larger-scale behavioral intervention studies in aging, it appears that rapid growth of this literature will continue. Perhaps behavioral intervention research in aging is now in its adolescence. It is hoped that as it experiences the growth spurts, mood swings, and unlimited horizons characteristic of adolescence that it will continue to seek guidance from and affiliation with its parent disciplines in the psychology of cognition and aging, particularly with regard to theory and methods.

References

Appel, L. J., Espeland, M., Whelton, P. K., Dolecek, T., Kumanyika, S., Applegate, W. B., Ettinger, W. H., Jr., Dostis, J. B., Wilson, A. C. & Lacy, C. (1995). Trial of nonpharmacologic intervention in the elderly (TONE): Design and rationale of a blood pressure con-

trol trial. *Annals of Epidemiology, 5,* 119–129.

Ball, K., & Owsley C. (2000). Increasing mobility and reducing accidents in older drivers. In K. W. Schaie & M. Pietrucha (Eds.), *Mobility and transportation in the elderly.* New York: Springer.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology. On the dynamics between growth and decline. *Developmental Psychology, 23,* 611–626.

Baltes, P. B., & Danish, S. J. (1979). Gerontological interventions based on a life-span developmental psychology: Problems and concepts. *Zeitschrift Fuer entwicklungspsychologie und Paedagogische Psychologie, 11,* 112–140.

Baltes, P. B., & Kliegl R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology, 28,* 121–125.

Baltes, P. B., & Willis, S. L. (1977). Toward psychological theories of aging and development. In J. Birren & K. W. Schaie (Eds.), *The handbook of the psychology of aging* (pp. 128–154). New York: Van Nostrand-Reinhold.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology, 25,* 729–735.

Basham, R. B. (1986). Scientific and practical advantages of comparative design in psychotherapy outcome research. *Journal of Consulting and Clinical Psychology, 54,* 88–94.

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

Berry, J. M., West, R. L., & Dennehey, D. M. (1989). Reliability and validity of the Memory Self-Efficacy Questionnaire. *Developmental Psychology, 25,* 701–713.

Borkovec, T. D. (1994). Between-group therapy outcome research: Design and methodology. In L S. Onken & J. D. Blaine (Eds.), *Behavioral treatments for drug abuse and dependence. NIDA Research Monograph #137* (pp. 249–289). Rockville, MD: National Institute of Drug Abuse.

Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology, 66,* 136–142.

Brown, W. A. (1994). Placebo as a treatment for depression. *Neuropsychopharmacology, 10,* 265–269.

Brown, A. L. & Kane, L. R. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology, 20,* 493–523.

Camp, C. (1999). Memory interventions for normal and pathological older adults. In R. Schulz, G. Maddox, & M. P. Lawton (Eds.), *Annual review of gerontology and geriatrics: Focus on interventions research with older adults,* (pp. 1–16). New York: Springer.

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

Campione, J. C., Shapiro, A. M., & Brown, A. L. (1995). Forms of transfer in a community of learners: Flexible learning and understanding. In A. McKeough, J. Lupart, & A. Marini (Eds.), *Teaching for transfer: Fostering generalization in learning.* (pp. 35–68) Mahwah, NJ: Erlbaum.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54,* 1–22.

Cavanaugh, J. C., Krammer, D. A., Sinnott, J. D., Camp, C. J., & Markley, R. P (1985). On missing links and such: interfaces between cognitive research and everyday problem solving. *Human Development, 28,* 146–168.

Ceci, S., & Ruiz, A. (1993). Transfer, abstractness and intelligence. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp 168–191). Norwood, NJ: Ablex.

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 467–476.

Charness, N. (1985). Aging and problem solving performance. In N. Charness (Ed.), *Aging and human performance* (pp. 225–259). Chichester: Wiley.

Charness, N. & Campbell, J. I. (1988). Acquiring skill at mental calculation in adulthood: A task decomposition. *Journal of Experimental Psychology: General, 117*, 115–129.

Clark, F., Azen, S., Zemke, R., Jackson, J., Carlson, M., Mandel, D., Hay, J., Hosephson, K., Cherry, B., Hessel, C., Palmer, J., & Lipson, L. (1997). Occupational therapy for independent living older adults. *JAMA, 278*, 1321–1326.

Cohen, G., & Faulkner, D. (1986). Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology, 4*, 187–190.

Coon D. W., Schulz, R., Ory, M. G., & The REACH study group (1999). Innovative intervention approaches for Alzheimer's disease caregivers. In D. E. Biegel & A. Blum (Eds.), *Innovations in practice and service delivery across the lifespan* (pp. 295–325). New York: Oxford University Press.

Cormier, S. M., & Hagman, J. D. (Eds.). (1987). *Transfer of learning: Contemporary research and applications*. San Diego, CA: Academic Press.

Craik, F. I. M., & Salthouse, T. A. (1992). *Handbook of aging and cognition*. Hillsdale, NJ: Erlbaum.

Detterman, D. K. (1993). The case for the prosecution: Transfer as an epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp 1–24). Norwood, NJ: Ablex.

Detterman, D. K., & Sternberg, R. J. (Eds.) 1993 *Transfer on trial: Intelligence, cognition, and instruction*. Norwood, NJ: Ablex.

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

Dunlosky, J., & Hertzog, C. (1998a). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging, 13*, 597–607.

Dunlosky, J., & Hertzog, C. (1998b). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker & J. D. Dunlosky (Eds.), Metacognition in educational theory and practice. The educational psychology series (pp. 249–275). Mahwah, NJ: Lawrence Erlbaum.

Dunlosky, J., & Hertzog, C. (in press). Measuring strategy production during associative learning: The relative utility of concurren versus retrospective reports. *Memory an Cognition*.

Eysenck, H. J. (1994). The outcome problem i psychotherapy: What have we learned *Behavior Research and Therapy, 32*, 477-495.

Ferguson, G. A. (1956). On transfer and th abilities of man. *Canadian Journal o Psychology, 10*, 121–131.

Finkel, S. I., & Yesavage, J. A. (1989). Learning mnemonics: A preliminary evaluation of a computer-aided instruction package for the elderly. *Experimental Aging Research, 15*, 199–201.

Fisk, A. D., Lee, M. D., & Rogers, W. A. (1991). Recombination of automatic processing components: The effects of transfer, reversal, and conflict situations. *Human Factors, 33*, 267–280.

Fisk, A. D., Rogers, W. A., Cooper, B. P., & Gilbert, D. K. (1997). Automatic category search and its transfer: Aging, type of search and level of learning. *Journal of Gerontology: Psychological Sciences, 52B*, P91–P102.

Fisk, A. D., Rogers, W. A., & Giambra, L. M. (1990). Consistent and varied memory/visual search: Is there an interaction between age and response-set effects? *Journal of Gerontology: Psychological Sciences, 45*, P81–87.

Gratzinger, P., Sheikh, J. I., Friedman, L., & Yesavage, J. A. (1990). Cognitive interventions to improve face-name recall: The role of personality trait differences. *Developmental Psychology, 26*, 889–893.

Hayslip, B. (1989). Alternative mechanisms for improvements in fluid ability performance among older adults. *Psychology and Aging, 4*, 122–124.

Hazlett-Stevens, H., & Borkovec, T. D. (1999). Experimental design and methodology in between-group intervention outcome research. In R. Schulz, G. Maddox, & M. P. Lawton (Eds.). *Annual review of gerontology and geriatrics: Focus on interventions research with older adults*. (pp. 17–47). New York: Springer.

Hertzog, C., McGuire, C. L., & Lineweaver, T. T. (1998). Aging, attributions, perceived control, and strategy use in a free recall task. *Aging, Neuropsychology, and Cognition, 5*, 85–106.

Hill, R. D., Sheikh, J. I., & Yesavage, J. A. (1989). Pretraining enhances mnemonic training in elderly adults. *Experimental Aging Research, 14,* 207–211.

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.

Jacobson, N. S., & Hollon, S. D. (1996). Cognitive-behavior therapy versus pharmacotherapy: Now that the jury's returned its verdict, it's time to present the rest of the evidence. *Journal of Consulting and Clinical Psychology, 64,* 74–80.

Jobe, J. B., Smith, D. M., Ball, K., Tennstedt, S. Marsiske, M., Rebok, G., Morris, J. N., Willis, S. L., Helmers, K., Leveck, M. D., & Kleinman, K. (2000). *ACTIVE: A cognitive intervention trial to promote independence in older adults.* Washington, DC: National Institutes of Health.

Kastenbaum, R. (1968). Perspectives on the development and modification of behavior in the aged: A developmental-field perspective. *Gerontologist, 8,* 280–283.

Kausler, D. (1994). *Learning and memory in normal aging.* San Diego, CA: Academic Press.

Kavanagh, D. J., & Bower, G. H. (1985). Mood and self-efficacy: Impact of joy and sadness on perceived capabilities. *Cognitive Therapy and Research, 9,* 507–525.

Kazdin, A. E. (1986). Comparative outcome studies of psychotherapy: Methodological issues and strategies. *Journal of Consulting and Clinical Psychology, 54,* 95–105.

Kazdin, A. E. (1992). *Research design in clinical psychology* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Kliegl, R., Smith, J., & Baltes, P. B. (1989). Testing-the-limits and the study of adult age differences in cognitive plasticity of a mnemonic skill. *Developmental Psychology, 25,* 247–256.

Kliegl, R., Smith, J., & Baltes, P. B. (1990). On the locus and process of magnification of age differences during mnemonic training. *Developmental Psychology, 26,* 894–904.

Klein, D. F. (1996). Preventing hung juries about therapy studies. *Journal of Consulting and Clinical Psychology, 64,* 81–87.

Kotler, S., & Camp, C. J. (1990). Memory interventions and aging. In E. Lovelace (Ed.), *Aging and cognition: Mental processes, self awareness, and interventions.* Amsterdam: North Holland.

Lachman, M. E., Weaver, S. L., Bandura, M., Elliot, E., & Lewkowicz, C. (1992). Improving memory and control beliefs through cognitive restructuring and self-generated strategies. *Journal of Gerontology: Psychological Sciences, 47,* P293–P298.

Langer, E. J. (1989). *Mindfulness.* Reading, MA: Addison-Wesley.

LaRosa, J. C, Applegate, W., Crouse, J. R., Hunninghake, D. B., Grimm, R., Knopp, R., Eckfeldt, J. H., Davis, C. E., & Gordon, D. J. (1994). Cholesterol-lowering in the elderly: Results of the cholesterol reduction in seniors program (CRISP) pilot study. *Archives of Internal Medicine, 154,* 529–539.

Lerner, R. M. (1986). *Concepts and theories of human development.* New York: Random.

Mayer, R. E., & Greeno, J. G. (1972). Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology, 63,* 165–173.

McDougall, G. J. Jr. (1999). Cognitve interventiosn among older adults. In J. J. Fitzpatrick (Ed.), *Annual review of nursing research, vol. 17* (pp. 219–240). New York: Springer.

McKeough, A., Lupart, J., & Marini, A. (Eds). (1995). *Teaching for transfer: Fostering generalization in learning.* Mahwah, NJ: Erlbaum.

Mohs, R. C., Ashman, T. A., Jatzen, K., Albert, M., Brandt, J., Gordon, B., Rasmusson, X., Grossman, M., Jacobs, D., & Stern, Y. (1998). A study of the efficacy of a comprehensive memory enhancement program in healthy elderly persons. *Psychiatry Research, 77,* 183–195.

Morin, C., Colecchi, C., Brink, D., & Abstruc, M. (1995). How "blind" are double-blind placebo-controlled trials of benzodiazepine hynotics? *Sleep, 18,* 240–245.

Morin, C. M., Culbert, J. P., & Schwartz, S. M. (1994). Nonpharmacological interventions for insomnia: A meta-analysis of treatment efficacy. *American Journal of Psychiatry, 151,* 1172–1180.

Schneider, W., & Fisk, D. (1984). Automatic category search and its transfer. *Journal of Experimental Psychology: Learning Memory and Cognition, 10,* 1–15.

Schulz, R., Maddox, G., & Lawton, M. P. (Eds). (1999). *Annual review of gerontology and geriatrics: Focus on interventions research with older adults.* New York: Springer.

Schulz, R., & Martire, L. M. (1999). Intervention research with older adults: Introduction, overview, and future directions. In R. Schulz, G. Maddox, & M. P. Lawton (Eds.), *Annual review of gerontology and geriatrics: Focus on interventions research with older adults* (pp. 1–16). New York: Springer.

Scogin, F., & Bienias, J. L. (1988). A three-year follow-up of older adult participants in a memory skills training program. *Psychology and Aging, 3,* 334–37.

Shapiro, A. K. (1971). Placebo effects in medicine, psychotherapy, and psychoanalysis. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change.* New York: Wiley.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill.* Cambridge, MA: Harvard University Press.

Smyer, M. A., & Gatz, M. (1986). Intervention research approaches. *Research on Aging, 8,* 536–558.

Smyer, M. A, Zarit, S. H., & Qualls, S. H. (1990). Psychological intervention with the aging individual. Chapter in J. E. Birren & K. W. Schaie (Eds.), Handbook of the psychology of aging, 3rd ed., (pp. 375–403). San Diego, CA: Academic Press.

Sternberg, R. L. (1985). *Beyond IQ: A triarchic theory of human intelligence.* New York: Cambridge University of Press.

Sternberg, R. J., & Frensch, P. A. (1993). Mechanisms of transfer. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 25–38). Norwood, NJ: Ablex.

Strayhorn, J. (1987). Control groups for psychosocial intervention outcome studies. *American Journal of Psychiatry, 144,* 275–282.

Thorndike, E. L., & Woodworth, R. S. (1901) The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8,* 247–261.

Tinetti, M. E., Baker, D. I., Garrett, P. A., Gottschalk, M., Koch, M. L., & Horwitz, R. I. (1993). Yale Ficsit: Risk factor abatement strategy for fall prevention. *Journal of the American Geriatric Society, 41,* 315–320.

Verhaeghen, P., & Marcoen, A. (1994). The production deficiency hypothesis revisited: Adult age differences in strategy use as a function of processing resources. *Aging and Cognition, 1,* 32–338.

Verhaeghen, P., & Marcoen, A. (1996). On the mechanisms of plasticity in young and older adults after instruction in the Method of Loci: Evidence for and amplification model. *Psychology and Aging, 11,* 164–178.

Voss, J. F. (1990). Learning and transfer in subject matter learning: A problem-solving model. In P. J. D., Drenth & J. A. Sergeant (Eds.), *European perspectives in psychology, Vol. 1: Theoretical, psychometrics, personality, developmental, educational, cognitive, gerontological* (pp 607–621). Chichester, UK: John Wiley.

Wilkins, W. (1983). Failure of placebo groups to control for nonspecific events in therapy outcome research. *Psychotherapy: Theory, Research and Practice, 20,* 31–37.

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

Willis S. L. (1990). Current issues in cognitive training research. In E. A., Lovelace (Ed.), *Aging and cognition: Mental processes, self awareness, and interventions* (pp. 263–280). Amsterdam: Elsevier.

Willis, S. L. (1996). Everyday cognitive competence in elderly persons: conceptual issues and empirical findings. *Gerontologist, 36,* 595–601.

Willis, S. L., & Baltes, P. B. (1981). Derivation of gerontological training research from the Gf-Gc theory of intelligence: A reply to Donaldson and some critical observations. *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. C., & Baltes P. B. (1983). Training in research in aging: Attentional processes. *Journal of Educational Psychology, 75*, 257–270.

Willis S. L., Jay G. M., Diehl, M., & Marsiske M. (1992). Longitudinal change and prediction of everyday task performance in the elderly. *Research on Aging, 14*, 68–91.

Willis, S. L., & Nesselroade, C. S. (1990). Long term effects of fluid ability training in old age. *Developmental Psychology, 26*, 905–910.

Willis S. L., & Schaie, K. W. (1986). Training the elderly on the ability factors of spatial orientation and inductive reasoning. *Psychology and Aging, 1*, 129–247.

Willis, S. L., & Schaie, K. W. (1994). Cognitive training in the normal elderly. In F. Forette, Y. Christen, & F. Boller (Eds.), *Plasticité cérébrale et stimulation cognitive* (pp. 91–113). Paris: Foundation National de Gérontologie.

Yelon, S. (1992) M.A.S.S.: A model for producing transfer. *Performance Improvement Quarterly, 5*, 13–23.

Yesavage, J. A. (1990). Age-associated memory impairment: Conceptual background and treatment approaches. In *Challenges in aging: The 1990 Sandoz Lectures in gerontology*. (pp. 53–72). London: Academic Press.

Yesavage, J. A., Rose, T. L., & Bower, G. H. (1983). Interactive imagery and affective judgements improve face-name learning in the elderly. *Journal of Gerontology, 38*, 197–203.

Zarit, S., H., Cole, K. D., & Guider, R. L. (1981). Memory training strategies and subjective complaints of memory in the aged. *The Gerontologist, 21*, 158–164.