THE PREDICTION PROCESS IN METAMEMORY: A Distinction Between General and Task-Specific Self-Efficacy

by

Leslie Carol McDonald-Miszczak

B.A., University of Alberta, 1987

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF ARTS

in the Department of Psychology

We accept this thesis as conforming
to the required standard

[redacted]

Dr. D.F. Hultsch, Supervisor (Dept. of Psychology)

[redacted]

Dr. M. Hunter, Departmental Member (Dept. of Psychology)

[redacted]

Dr. E. Gallagher, Outside Member (Dept. of Nursing)

[redacted]

Dr. P. Baker, External Examiner (Dept. of Sociology)

University of Victoria

Supervisor: Dr. David F. Hultsch

ABSTRACT

Previous research has suggested that younger and older adults may differ in their memory monitoring ability (e.g., global performance predictions). More recently, investigations have begun to focus on the prediction process itself suggesting that it is comprised of adults' general memory self-efficacy beliefs, task-specific self-efficacy, and appraisal of the specific memory task. The present investigation examined the influence of normative task information and general memory self-efficacy beliefs on the prediction and recall of word lists. Normative task information was manipulated by having the 36 younger (20-30 yrs.) and 36 older (65-75 yrs.) subjects give their predictions and recall after receiving (a) no normative information, (b) self-generated normative information, or (c) accurate age-graded normative information. A repeated measures (3 trials) design was used. The results indicated that: (a) the elderly predicted their performance more accurately than the younger subjects, (b) subjects in the accurate normative information condition predicted their performance most accurately, and (c) both age groups seemed to monitor their predictions accurately. Hierarchical regression analyses showed that general memory self-efficacy beliefs influenced first trial predictions, whereas subsequent predictions were influenced by previous

that performance predictions are part of a process that is comprised of general memory self-efficacy beliefs, task specific self-efficacy, and task appraisal.

Examiners:

[redacted]

Dr. D.F. Hultsch, Supervisor (Dept. of Psychology)

[redacted]

Dr. M. Hunter, Departmental Member (Dept. of Psychology)

[redacted]

Dr. E. Gallagher, Outside Member (Dept. of Nursing)

[redacted]

Dr. P. Baker, External Examiner (Dept. of Sociology)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge those people who contributed to the preparation of this thesis. I would like to thank my committee members: Dave Hultsch, my supervisor, for his level-headed advice and support throughout this project; Mike Hunter for his statistical/methodological expertise; and Elaine Gallagher for her thoughtful comments and encouragement.

Next, I give special acknowledgement to Tomasz Miszczak who completely supported me and my work, and who put his dreams on hold so that I could realize mine. To my friends Anita, Odette, and Heather, thank you for listening to my concerns and giving me great advice throughout this project.

Finally, I wish to thank Hugh and Martha McDonald for believing in me and teaching me that I could do anything that I put my mind to.

Chapter One

INTRODUCTION

Since the 1970's, the concept of metamemory has been used to investigate aspects of memory development. Metamemory may be defined broadly as the knowledge, beliefs, and affect associated with one's own and others' memory functioning and capabilities (Hultsch, Hertzog, Dixon, & Davidson, 1988). Results from adult cognitive research suggest that beliefs about one's memory are clearly different for younger and older adults. Generally, most questionnaire measures of metamemory show negative age differences or decline with age (Gilewski & Zelinski, 1986). Since the elderly tend to hold more negative beliefs about their memories, it is possible that such beliefs indirectly impair memory performance (i.e., by decreasing one's effort or impairing one's judgement regarding the suitability of a certain mnemonic strategy).

Some empirical research in the adult metamemory literature has focused on the measurement of metamemorial knowledge. Such knowledge may be best represented by facts about one's own memory and the human memory system in general. For example, most people are aware that a list of concrete objects is easier to memorize than a list of abstract ideas. Intuitively, it might be assumed that the more memory facts that are known, the better one's expected memory ability. However, given that the correlations obtained between various measures of metamemory and performance on memory tasks are relatively low, this assumption is not entirely accurate (see Gilewski & Zelinski, 1986; Herrmann, 1982).

Some researchers have concentrated on other aspects of metamemory (Hultsch et al., 1988). Specifically, these investigators have argued that perceptions regarding one's own memory capabilities may be an important aspect of the metamemory concept. Sehulster (1981a, 1981b, 1982) proposed that such beliefs constitute a self-theory of memory that is based on direct (i.e., actual task participation) and indirect (i.e., opinions and observations) task experience. Once a self-theory of memory is formed, it can either be revised with additional experiences or remain unchanged. Accordingly, it is possible that certain individuals may revise their self-theories based on inaccurate interpretations or exaggerations of certain situations, or may resist revision when it is actually necessary. Such individuals would subsequently hold beliefs that do not coincide with their actual memory behaviour.

Memory self-efficacy may be defined as beliefs about one's memory capability in a variety of memory situations. The memory self-efficacy construct has been used to examine adults' memory beliefs. It has been shown that memory self-efficacy is an aspect of metamemory that is separate from memory knowledge (Hertzog, Dixon, Schulenberg, & Hultsch, 1987).

Global performance predictions are another metamemory measure previously assumed to directly measure memory knowledge. The prediction paradigm requires subjects to make a judgement about their performance before actually experiencing the entire memory task. For example, subjects might be asked to indicate the number of words that will be correctly recalled from a list of 50 words. Traditionally, researchers using this paradigm assumed that subjects who accurately predict their performance understand their memories to a greater extent than subjects who do not accurately predict their performance. The prediction literature has shown contradictory results with respect to adult age differences. Some researchers have found that the elderly predict their performance as accurately as the young (Lachman, Lachman, & Thronesbury, 1979; Perlmutter, 1978; Rabinowitz, Ackerman, Craik, & Hinchley, 1982), whereas other researchers have found that the elderly do not predict their performance as accurately as the young (Bruce, Coyne, & Botwinick, 1982; Lovelace & Marsh, 1985; Murphy, Sanders, Gabriesheski, & Schmitt, 1981).

Hertzog, Dixon, and Hultsch (1990a) proposed that the prediction process is composed of (a) assessment of one's general memory self-efficacy (MSE), (b) assessments of normative task difficulty, and (c) mapping one's self-efficacy onto the assessed task performance distribution. Recent investigations have indicated that general MSE is important to predicted performance when the subject has not had direct task experience (i.e., on the first trial) (Hertzog, Dixon, & Hultsch, 1990b; McDonald-Miszczak, Hultsch, & Hunter, 1990; Saylor, 1990). However, following task experience (i.e., the second trial), performance predictions are most highly related to previous task experience (i.e., prediction and recall on trial 1) rather than general MSE. In short, these studies offer a multidimensional view of memory self-efficacy in a particular context (i.e., a specific memory task).

The primary objective of past studies has been to demonstrate that multiple dimensions exist within the prediction process (e.g., general vs task specific memory self-efficacy). In contrast to the recent investigations, there has not been a direct attempt to study the influence of accurate age-graded normative task information on the multidimensional aspects of memory self-efficacy. The present study, then, is designed to examine the relationship between age-graded normative task information, general MSE, and performance prediction accuracy in adults of different ages.

Chapter Two

LITERATURE REVIEW

The following literature review will discuss empirical findings that are relevant to the investigation of age differences in the prediction process. The first section will examine the definition of the metamemory concept, distinguishing between metamemory as memory knowledge and metamemory as memory beliefs. The second section will discuss the prediction paradigm as a measure of metamemory. The last section will introduce the objectives and hypotheses of the present investigation.

The Metamemory Concept

Metamemory was originally defined as the "intelligent structuring and storage of input, of intelligent search and retrieval operations, and of intelligent monitoring and knowledge of these storage and retrieval operations" (Flavell, 1971, p. 277). However, the concept of metamemory has changed greatly since it was first used in the early 1970's. The concept has evolved from a specific focus on memory knowledge to a more general focus that includes memory beliefs and memory self-efficacy.

Metamemory as Memory Knowledge

Despite the fact that Flavell's original definition of metamemory incorporated several components, the construct was first empirically conceptualized as one's knowledge about one's own and others' memory (Flavell, 1971; Flavell & Wellman, 1977).

Such knowledge involves awareness of facts about memory. For example, when asked if a longer word list is harder to memorize than a shorter list, most people know that the longer list is more difficult. This type of memory knowledge, or awareness, has been studied extensively by metamemory researchers.

Developmental research on children of various ages has shown that as individuals grow older, they become more aware and knowledgeable about their own memories (Flavell, 1971). According to Flavell's (1971) argument, not only do memory strategies develop with age, but so does knowledge and intelligent monitoring of encoding and retrieval processes. For example, a study by Kreutzer, Leonard, and Flavell (1975) showed that although children in kindergarten and the first grade did have some basic knowledge about mnemonic processes, older children in the third and fifth grades were superior in the number and complexity of strategies they could generate. Another study by Flavell, Friedrichs, and Hoyt (1970) found that nursery school and kindergarten aged children were less accurate at predicting their own memory span and at knowing when items were memorized to a point where they could be recalled without errors than older children in grades 2 and 4. Based on such research findings, metamemory, conceptualized primarily as knowledge and intelligent usage of memory processes, was thought to be a critical factor in memory development (Flavell, 1971).

Flavell and his colleagues later refined the metamemory concept and proposed a taxonomy for metamemory factors (Flavell & Wellman, 1977). The two main categories were called, (a) sensitivity, and (b) variables. These two main types of metamemory factors shaped future research interests.

The first category includes "sensitivity to the objective need for efforts at present retrieval or at preparation for future retrieval" (Flavell & Wellman, 1977, p. 6). This type of metamemory demonstrates a person's increasing understanding of a response to memory situations that require mnemonic strategies. Children develop the ability to treat a set of items differently when asked to memorize them than when asked to process them for other purposes (Appel, Cooper, McCarrel, Sims-Knight, Yussen, & Flavell, 1972). A child learns planful and goal directed memory behaviours that are used to enhance memory performance (Flavell & Wellman, 1977). Thus, part of memory development itself includes such memory enhancing behaviours.

The second metamemory category, 'variables', includes "knowing what variables interact in what ways to affect the quality of performance on a retrieval problem" (Flavell & Wellman, 1977, p. 10). According to Flavell and Wellman

(1977), these variables consist of three main types: (a) person variables, (b) task variables, and (c) strategy variables. Person variables include all aspects of one's self that are relevant to one's memory (Flavell & Wellman, 1977). As an individual experiences different memory situations, knowledge is obtained regarding one's own and others' memory properties, abilities, and limitations. Task variables include knowledge of factors that render some memory situations more difficult than others. Some memory tasks are more cognitively demanding than others. Strategy variables include knowledge and use of mnemonic strategies. These memory behaviours (e.g., written notes, imagery, organization of information) are used to facilitate memory performance. Flavell and Wellman (1977) proposed that a person who is a sophisticated metamemory user does not treat these three classes of metamemory variables as distinct from one another. These variables provide the best results when they are utilized in a complex interaction by providing the most cognitive support to memory performance.

In short, Flavell (1971) proposed that metamemory plays a pivotal role in memory development, arguing that the intelligent use of mnemonic strategies requires metamemory. Without an adequate understanding of one's memory system, and the type of memorizer one is in a particular task situation, it would be difficult to employ mnemonic strategies to enhance one's memory performance.

It is important to note, however, that the proposed indirect link between metamemory and memory performance also has implications for aging. Thus far, metamemory has been discussed in terms of positive memory development in children. Alternatively, the information acquired from the child literature can be used as a descriptive tool for poor memory performance displayed by certain individuals (i.e., the elderly). From this perspective, one possible explanation for at least some aspects of age differences in adults' memory performance may stem from inadequate metamemory.

Awareness of Metamemory

Metamemorial awareness is an aspect of metamemory that may be useful for understanding differences in adults' memory performance. The logic of metamemorial awareness asserts that the more aware people are of their metamemorial processes, the more likely they will utilize important metamemorial information. Individuals' awareness of memory knowledge has been specifically referred to as systemic awareness (Cavanaugh, 1989). According to Cavanaugh and Perlmutter (1982), systemic awareness is equivalent to the traditional definition of metamemory (i.e., metamemory as memory knowledge). The notion that individuals are aware of their memory knowledge is the basis for questionnaire measures of metamemory. Metamemory questionnaires measure an individual's conscious understanding of various aspects of their memory functioning.

Several metamemory questionnaires have been developed to address cognitive aging issues (for examples see Gilewski & Zelinski, 1986; Herrmann, 1982). The original intention for using metamemory questionnaires was to directly measure individuals' everyday memory knowledge without the complications of naturalistic observation (Herrmann, 1982). Most questionnaires typically use Likert-type ratings on statements that address different memory dimensions (Gilewski & Zelinski, 1986). For example, subjects might be asked about (a) the kinds of strategies they use, (b) the difficulty of certain memory tasks, (c) the amount of memory improvement or decline they have experienced, or (d) how they emotionally react to performing different memory tasks. Many questionnaires have been developed and used in cognitive aging studies (Bennett-Levy & Powell, 1980; Dixon & Hulstch, 1983; Herrmann & Neisser, 1978; Perlmutter, 1978; Zelinski, Gilewski, & Thompson, 1980).

Generally, most questionnaire measures show negative age differences or decline in metamemory with age. Particular findings, however, show some discrepancy. For example, Bennett-Levy and Powell (1980) found that elderly individuals' perceptions of their memory abilities were higher than young individuals' perceptions. Other researchers studying the same aspect of metamemory found mixed results (Chaffin & Herrmann, 1983), or that the young indicated that their memory abilities were greater than the elderly's abilities (Dixon, & Hultsch, 1983). These discrepant findings may have been the result of a number of factors (i.e., sampling differences, different measures of the construct) (Dixon, 1989). The questionnaire studies were also quite variable in reporting reliability and validity information (Dixon, 1989).

Despite the recent popularity of incorporating metamemory questionnaires in cognitive aging research, the relationship between metamemory and memory performance has not been strong. Studies have found that questionnaire responses are reliable, but correspond only moderately to a person's memory performance (Gilewski & Zelinski, 1986; Herrmann, 1982). For example, a series of studies by Schlecter and Herrman (cited in Herrman, 1982) found no significant correlations between the Short Inventory of Memory Experience (SIME) forgetting factors and reported forgetting frequencies recorded in open-ended diary form. Results have been obtained that show only small correlations between perceived aspects of memory and laboratory task performance (Perlmutter, Metzger, Nezworski, and Miller, 1981; Zelinski et al., 1980). Apparently, questionnaire responses regarding one's memory performance are stable, but not very accurate (Herrmann, 1982). The small correlations do not support a strong relationship between questionnaire measures of memory awareness (i.e., memory knowledge) and memory performance.

Summary. Metamemory was first conceptualized as facts about memory and primarily researched within child development. Researchers interested in adult cognitive development adopted the metamemory concept to study indirect effects on adult memory decline. Metamemory questionnaires were developed to measure multidimensional aspects of adults' metamemory. Disappointing and contradictory results were obtained using the questionnaire method.

Metamemory as Memory Beliefs

Knowledge regarding one's own memory and the human memory system may not necessarily determine how capable one believes one is in different memory situations. Some researchers have incorporated the notion of memory self-efficacy into their investigations. Memory self-efficacy can be broadly defined as one's self-efficacy beliefs about one's own memory. Perhaps Herrmann (1982) was one of the first to refer to metamemory questionnaire responses as beliefs rather than as knowledge in an early review. The distinction between memory knowledge and memory beliefs may be particularly relevant to cognitive aging research. If the metamemory system becomes unreliable with increasing age, then elderly individuals may lose confidence when memory failures occur and subsequently behave differently in memory demanding situations (Berry, 1989). Such differences in behaviour may magnify normal memory decline. Memory self-efficacy may be a good tool for investigating the previously disappointing relationship between questionnaires and memory performance (i.e., memory beliefs may not necessarily be congruent with memory performance). Thus, memory self-efficacy should be viewed as another facet of the metamemory concept separate from memory knowledge (Berry, 1989).

Beliefs about one's memory capability have been referred to as memory self-efficacy. Specifically, memory self-efficacy can be defined as beliefs about one's own capability to use memory effectively in a variety of situations (Hertzog et al., 1990a). Obviously this memory concept may be related to a person's general sense of self-efficacy. Self-efficacy refers to a person's beliefs regarding motivation, cognitive resources and capacity, and mobilization of actions to exercise control over a task (Bandura, 1986). Memory self-efficacy refers to an individual's self-efficacy within the specific domain of memory tasks. Such beliefs are not thought to be passive indicators of the individual's opinion regarding a task. A person's memory self-efficacy is thought to lead that person to behave in a particular manner when confronted with a specific memory task.

The notion that metamemory consists of a belief system was originally proposed by Sehulster (1981a, 1981b, 1982).

Sehulster proposed that statements such as "I have a good memory for names" reflect an aspect of an individual's theory of self. A self-theory is part of a person's overall world view and is based on evidence collected from past experiences that are then generalized across time. Evidence may come from either direct or indirect experience. For example, an elderly individual who experiences problems trying to remember names at a few cocktail parties, may consequently believe that he or she is not good at remembering names. Another elderly individual may believe a friend's comments regarding his or her forgetfulness. Thus, the basis of one's self-theory may be formed by direct experience with a memory task (e.g., remembering names) or by indirect experience (e.g., believing a friend's criticisms).

Memory self-efficacy reflects perceptions about one's ability in a memory situation. These perceptions, however, are not static. Once a self-theory has been formed, new experiences may be incorporated into the existing self-theory. Cognitive dissonance theory is an example of how such revisions might occur (Sehulster, 1981a).

On the other hand, one's self-theory does not always accurately reflect memory experience. An individual may resist revision of his or her self-theory. Experience may be denied or distorted so that a revision does not occur. For example, an individual may continue to believe that his or her memory for dates is excellent even though some dates cannot be recalled on demand. Thus, a person's memory self-theory may not always coincide with actual memory behaviour.

Sehulster, (1981a) proposed that a view of metamemory as a memory self-theory expanded traditional conceptions about metamemory in three ways. First, if memory beliefs are viewed as a subset of a person's beliefs about oneself, then findings that have arisen from metamemory studies may be utilized to describe and explain the effect of beliefs on a person's overall cognitive style (Sehulster, 1981a). This broadened definition of metamemory is especially important when the subjects (e.g., the elderly) hold beliefs that may play a salient role in cognition (Hultsch et al., 1987). Second, the concept of memory beliefs has suggested that certain behaviours may prove to be more highly correlated with memory questionnaire data than memory task performance (Sehulster, 1981a). For example, individuals with negative memory beliefs might avoid various memory situations even though they perform quite adequately in such situations. In this case, avoidance behaviour may not correlate highly with actual memory performance, but would correlate with perceived memory ability. This second point highlights the importance of acknowledging the belief component in any measure of metamemory. Third, the study of metamemory as a belief system may allow researchers to discover the aspects of memory that are affected by, and in turn, affect performance. For example, a belief that one has a poor memory for short stories may hamper one's choice and use of mnemonic aids rendering one's recall performance below the norm. In view of findings that have demonstrated little relation between memory questionnaire responses and memory performance (see Gilewski & Zelinski, 1986; see Herrman, 1982), Sehulster's self-theory of memory is theoretically attractive and empirically probable.

Measuring Memory Self-Efficacy

One questionnaire that has been valuable for defining and measuring memory self-efficacy is the Metamemory In Adulthood (MIA) questionnaire (Dixon & Hultsch, 1983). This instrument originally had 120 items consisting of 8 different subscales. Each subscale represents a dimension of metamemory. The original a priori scales were (a) strategy: the subject's reported use of memory strategies of various types, (b) task: the subject's fundamental knowledge of generic memory tasks and process differences, (c) capacity: the respondent's perception of his or her own memory capacities regarding recall performance in designated situations, (d) change: one's perceived developmental process of remembering capacities for one's self and for others, (d) activity: the subject's reported activities that may affect one's memory abilities, (f) anxiety: the respondent's emotional state and the influence of that state when confronted with a memory task, (g) achievement: one's rated importance of memory ability and performance, and (h) locus: the respondent's perceived control over memory skills (Dixon, 1989).

Internal consistency reliability estimates (Chronbach's alpha) obtained from five independent samples (total N = 1108) ranged between .71 and .93 for seven of the eight original MIA subscales (Dixon & Hultsch, 1983). Subsequently, the activities supportive of memory subscale was dropped due to relatively low internal consistency (.28 to .76) (Hultsch et al., 1988). Conceptually, the low alpha scores for the activity subscale are intuitively reasonable. Different activities need not correlate highly with a single performance on a specific memory task. Thus, the MIA has proven to be an encompassing and internally valid measure of metamemory that may serve as a base to conduct further refining analyses on the metamemory concept.

Hertzog et al. (1987) tested the hypothesis that the seven subscales of the MIA are markers of three higher order factors. The hypothesized higher order factors are (a) Memory Knowledge, (b) Memory Self-Efficacy, and (c) Memory Related Affect dimensions. A multiple groups confirmatory factor analysis was conducted on data from six separate studies involving a total of 750 subjects that were divided to yield two half-samples for cross validation analyses. Each half-sample was divided into younger- (overall $\underline{M}$ = 27 years), middle- (overall $\underline{M}$ = 47 years), and older- (overall

M = 68 years) aged groups to view the factor structure at different ages. The first half-sample was used to develop the model and the second half-sample was used to validate it. The models did not fully cross-validate, but evidence was found for at least two higher order factors in the MIA. The first was tentatively labelled Memory Knowledge (MK), and included combined knowledge about memory and affect related to memory. The second was labelled Memory Self-Efficacy (MSE) and consisted of an individual's beliefs regarding one's competence in memory demanding situations.

The researchers found that factor loadings on the MK dimension did not vary according to age. The Strategy, Task, Achievement, and Anxiety subscales consistently loaded on the MK dimension. According to the mostly stable factor loadings on the MK factor, older individuals seem to be as knowledgeable about their memory ability as younger individuals.

On the other hand, the factor loadings on the MSE dimension varied as a function of age. Generally, the MSE dimension was defined by Capacity, Change, Anxiety, Achievement, and Locus subscales. Although all of the above subscales loaded on the MSE factor, it was the Capacity, Change, and Locus subscales that were most highly related to MSE (Hertzog et al., 1987). The Change, Anxiety, and Locus subscales, however, had higher loadings in the old than in the young groups thereby accounting for most of the age differences in this factor. Age differences in the relationship between the Change, Locus, and Anxiety subscales were thought to show that people who believe that their memories have changed greatly are also more anxious about their general memory ability and believe that they have little control over their memory. Moreover, from the age differences apparent in the MSE factor loadings, older individuals seem to perceive themselves as less competent than younger individuals. Thus, from these results, it is possible that past findings that showed age differences in metamemory were primarily caused by differential beliefs regarding memory ability rather than differences in actual memory knowledge.

Discriminant Validity

In order for the role of the memory self-efficacy concept to be understood in metamemory research, it must have discriminant validity. A series of multivariate analyses were conducted by Hultsch et al. (1988) that examined the discriminant validity of the MSE factor from related constructs. The MSE factor was tested for its relation to a personality measure (Jackson Personality Inventory and Jackson Personality Research form), a generalized measure of locus of control (Levenson Internal/External Locus of Control scale), and two measures of psychological well being and depression (Veit/Ware scales and the Centre for Epidemiologic Studies - Depression scale). The correlations of the MIA Locus scale ranged from -.004 on the depression scale to .194 on the personality measure of endurance. The correlations of the Capacity scale ranged from -.041 on Levenson's chance scale to .315 on the personality measure of self-esteem. The correlations of the Change scale ranged from -.048 on Levenson's change scale to .255 on the personality measure of anxiety. Although the correlations between the MIA subscale components of the MSE construct (i.e., Locus, Capacity, Change, & Anxiety) and the other psychological measures exhibited a wide range, they were generally quite low (Hultsch et al., 1988). The low correlations showed that the MIA components of the MSE construct were not simply a reiteration of other psychological measures. The MSE construct was found to be comprised of elements that uniquely measure one's beliefs about memory.

Although the MSE factor was comprised of many different MIA subscales, the data collected by Hertzog et al. (1987) have shown that the best single indicator is the Capacity subscale. The Capacity subscale is unique to the MSE factor apart from the MK factor. Unlike other subscales, it does not highly correlate with any other subscale that may link it directly or indirectly to the MK factor. Moreover, over all age groups, the Capacity subscale consistently shows the highest factor loading on the MSE factor. Thus, if a single indicator of MSE is used, it will be best represented by the Capacity subscale.

Convergent Validity

The MSE construct was found to be discriminantly valid from related constructs, but it is also important to show that it has convergent validity. Such analyses should show that the MSE construct is not only an aspect of the MIA, but that it is a general metamemory phenomenon that can be measured with other metamemory questionnaires.

The Memory Functioning Questionnaire (MFQ) is another multidimensional instrument that was developed to measure metamemory (Zelinski et al., 1980). It contains seven scales: (a) general rating, (b) retrospective functioning, (c) frequency of forgetting, (d) frequency of forgetting when reading, (e) remembering past events, (f) seriousness, and (h) mnemonics (Gilewski & Zelinski, 1986). A factor analysis of the MFQ resulted in 64 items that yielded three correlated common factors: (a) Frequency of Forgetting, (b) Seriousness, and (c) Mnemonics and Retrospective Functioning (Gilewski & Zelinski, 1986).

Hertzog, Hultsch, and Dixon (1989) tested the convergent validity between the MIA and MFQ questionnaires. An earlier content analysis of the MFQ suggested that its global rating scale, memory problems, and remote memory scale resemble the MIA contents of the MSE factor (Hultsch et al., 1988). Convergent validity tests were based on the assumption that the MIA Memory Self-Efficacy and MFQ Memory Self-Efficacy higher order factors should display a perfect correlation if they are entirely convergent. Using latent variable modelling on two distinct samples (Victoria BC and Annville Pennsylvania) factor correlations between the two questionnaire measures of MSE were .87 with a standard error of .04 and .99 with a standard error of .03 respectively. However, the hypothesis that convergence of the MSE factors could be addressed with a formal likelihood ratio test by fixing the correlations between the two MSE factors to 1 was not supported. Additional analyses showed that the MFQ contains a component that attenuates observed age differences among the MFQ scales and produces the smaller correlation between the MIA Memory Self-Efficacy and MFQ Memory Self-Efficacy factors. Thus, despite fundamental differences in scale definition, question wording, and format, the Hertzog et al. (1989) analyses showed that both the MIA and the MFQ contain the higher order MSE factor. These results demonstrate that the MSE factor is not solely peculiar to the MIA questionnaire.

Summary

The memory self-efficacy concept has become the focus of much recent metamemory research. Sehulster (1981a, 1981b, 1982) proposed that metamemory consists of a belief system that reflects a person's theory of self. Although a self-theory was thought to result from direct and indirect memory experiences, Sehulster (1981a) claimed that it did not always accurately reflect such experiences. Research on the relationship between metamemory questionnaires and memory performance corroborates Sehulster's claim. A view of metamemory as a self-theory expanded the breadth of the metamemory concept and reduced the expectation of finding a straight forward relationship between metamemory questionnaire responses and memory performance.

Memory self-efficacy has become a useful construct for examining memory beliefs. It was originally found to be a higher order factor of the MIA questionnaire (Hertzog et al., 1987). Further analyses have shown that the MSE factor is both discriminantly and convergently valid (Hertzog et al., 1989; Hultsch et al., 1988). These results have shown that memory self-efficacy is a pervasive concept in metamemory research.

The lack of a relationship between metamemory questionnaire and memory performance still requires further investigation. One important line of research has linked memory self-efficacy, performance predictions, and memory performance. This line of research is pursued in the next section.

The Prediction Paradigm

Recently, metamemory research has focused on the role of memory beliefs in the prediction process (Hertzog et al., 1990b). The prediction process has been used to highlight the role of memory beliefs because it is thought to contain elements of one's memory beliefs along with evaluations of task difficulty (Hertzog et al., 1990b). This section will move toward a multidimensional understanding of the prediction process by first reviewing the traditional definition of the prediction measure, and by discussing relevant research on age differences in prediction accuracy. A review of background material will highlight the parallel development between the general metamemory concept and the prediction paradigm.

The prediction paradigm was originally developed as a measure of memory knowledge, and the paradigm has been frequently used as an operational definition for one's knowledge regarding one's memory ability. The accuracy of a prediction, viewed in terms of recall performance, has been used as a direct index of one's understanding of one's memory capabilities.

Cavanaugh (1989) categorized performance predictions as a specific type of memory awareness called on-line awareness. This type of memory awareness involves a person's understanding of the task and its demands while performing the task. One must be aware of many aspects of the task as one performs it. When beginning a task, one must have certain cognitive expectations that ready one for performance or may even affect one's choice of undertaking the task altogether. These expectations will be quite different when confronted with a difficult task as opposed to an easy task. Different tasks require a different degree of cognitive readiness (e.g., choice of memory strategies, attention allocation, effort, motivation, performance goals). Once a person chooses to undertake a task and begins work on the task, reassessments of initial task impressions and performance expectations must be made. Such reassessments result from direct experience with the task materials and consequently allow one to improve as one performs the task. For example, an individual who chose to undertake the task using strategy A may decide after a short time that another strategy is required to perform the task adequately (i.e., strategy B or strategy C). On-line awareness, then, is one's cognitive awareness of one's memory as one performs a memory task. Such awareness may lead to revisions in a person's method for completing the task. Performance predictions are a measure of on-line awareness (Cavanaugh, 1989).

The Paradigm Defined

The prediction paradigm involves making a judgement about one's performance before actually experiencing the entire task. Often, subjects are shown a few examples of task items, or given a short practice test before being asked to give their predictions. Generally there are two different types of prediction measures used, (a) global predictions and (b) item-by-item predictions.

The first type of prediction measure is referred to here as a global prediction. This type of prediction measure requires subjects to predict their total recall after a brief introduction to the task, but before direct task experience. For example, an individual may be asked to memorize a list of 45 concrete nouns for 4.5 minutes. The subject is required to predict the number of words that will be correctly recalled (e.g., 30 words will be correctly recalled out of a total of 45 words). A global prediction then, requires a subject to predict overall recall performance (i.e., how well subjects predict they will perform on the entire task).

The second type of prediction measure is referred to here as an item-by-item prediction measure. For this type of prediction measure, subjects are required to make predictions while performing the task. Each individual test item will receive a Likert-type rating during study that reflects the likelihood that the specific item will be correctly recalled. For example, a subject might be presented with a list of 45 concrete nouns one word at a time. The subject may be allowed to study each word for 6 seconds and must rate the memorability of the word on a 5 point Likert-type scale (e.g., the scale may range from 1=will definitely not recall, to 5=will definitely recall). The memorability ratings are then compared to actual recall performance (e.g., words that received a high rating should have a greater probability of being recalled than words that received a low rating). The item-by-item predictions can be averaged to provide an overall estimate of the memory task, or can remain subdivided into separate rating categories (e.g., 1 - 5) and then compared on various factors (i.e., item difficulty, recall, word length).

The value of a prediction itself has not been the focus of metamemory work. The important metamemorial aspect has been the accuracy of one's prediction. The more accurate one's prediction is in terms of actual recall, the more knowledgeable one is thought to be about one's memory ability.

The accuracy of a global prediction has usually been determined by the difference between a subject's overall prediction and actual recall performance (e.g., predicted performance minus actual performance). For example, a subject who predicts that 30 words will be correctly recalled and 25 words are actually recalled, then that person will have a prediction accuracy score of +5. According to the above scoring scheme, a positive number reflects an overestimate, a negative number reflects an underestimate, and a score of zero reflects perfect prediction accuracy. Both the size and direction of the prediction accuracy scores determine one's metamemory knowledge. Generally, a positive score reflects a degree of overconfidence about one's memory ability, whereas a negative score reflects less confidence.

Prediction accuracy is determined quite differently for the item-by-item measure. Since item-by-item predictions are based on ratings of item memorability, a single difference score cannot be calculated. Instead, items that were rated as highly memorable should be recalled more readily than items rated as less memorable. For example, if 30 words were recalled correctly, the majority of these words should have been rated as highly memorable. The greater the correlation between rated memorability and actual recall, the greater one's memory knowledge.

Traditionally, both the global and item-by-item prediction measures were thought to be equivalent indicators of memory knowledge. However, prediction studies that included different age groups have shown contradictory results using the two different measures. These contradictory results have led to a distinction between the two different measures, and a greater empirical interest in the prediction process itself.

Age Differences In Prediction Accuracy

Much of the early metamemorial research performed with adults used the prediction paradigm (Bruce et al., 1982; Lachman et al., 1979; Lovelace & Marsh, 1985; Murphy et al., 1981; Perlmutter, 1978; Rabinowitz et al., 1982). Different studies have found contradictory results for age differences in prediction accuracy. Some researchers have found that the elderly are as accurate as the young in predicting their memory performance (Lachman et al., 1979; Perlmutter, 1978; Rabinowitz et al., 1982), whereas others have found that the elderly tend to overestimate their performance (Bruce et al., 1982; Murphy et al., 1981).

Perlmutter (1978) obtained memory prediction measures for word memory and general facts from younger and older samples. No significant age differences in prediction accuracy were found for either type of memory task. Both younger and older age groups predicted actual performance equally well for the recognition and recall measures of the general fact questions. For the word memory task, subjects tended to overestimate recall performance and underestimate recognition performance, but to the same extent in each age group.

Lachman et al. (1979) designed an experiment to tap accumulated world knowledge. The primary thrust of the design was the hypothesis that a fact becomes less accessible to an individual as the memory store grows with age (Lachman & Lachman, 1980). The design permitted measurement of (a) feeling-of-knowing (FOK) during recall, (b) recall search latencies, and (c) multiple-choice confidence ratings. Lachman et al., (1979) did not find any significant age differences on these measures. The authors concluded that metamemory is indeed maintained in later-life, at least in a healthy sample.

Rabinowitz et al. (1982) investigated the effects of age differences on participants' ability to predict the likelihood of recalling individual items. Rabinowitz et al. used a list of 50 word pairs. The word list included eight buffer pairs (four appeared at the beginning of the list and *the other four at the end), and 42 critical pairs which* consisted of 14 high-, moderate-, and low-related word pairs. The authors found that both younger and older adults were sensitive to the differences in the degree of relatedness between word pairs. Morover, a cross-subject conditional analysis (e.g., estimates of prediction accuracy when one person's predictions were used to predict another person's recall) suggested that accurate predictions were the result of both a culturally shared relatedness factor, and the individual's sensitivity to his or her just completed mnemonic processing. Both age groups appeared to utilize these factors to an equal degree. The result was a complete lack of age related differences on an item-by-item prediction measure. Rabinowitz et al., (1982) concluded that younger and older adults are quite comparable in their metamemorial skill.

Thus, some studies that have investigated prediction accuracy in younger and older adults report no apparent age differences in metamemorial skill. In these studies, older adults appeared to predict memory performance as accurately as younger adults (Lachman et al., 1979; Perlmutter, 1978; Rabinowitz et al., 1982), they were as sensitive to differences in test materials as younger adults (Rabinowitz et al., 1982), and monitored the presence or absence of information in storage and adjusted memory search efforts efficiently (Lachman et al., 1979). These studies seem to indicate that metamemorial ability does not decline with age. According to these findings, it is unlikely that declining metamemorial ability enhances typical age related memory decline.

There are, however, contradictory findings. Age differences have been found in a number of other prediction studies. These differences are thought to indicate a loss of metamemorial skill with age. A loss of metamemorial skill may play an important, albeit indirect, role in memory performance decline.

The first of a two-part study by Murphy et al. (1981) was designed to measure memory-span prediction and recall readiness in younger and older adults. Subjects were required to predict their memory span for a set of line drawings. A memory-span test was also administered to determine prediction accuracy. The participants were given lists consisting of a sub-span (i.e., two items less than one's memory span), memory-span (i.e., the same number of items as one's memory span), and super-span (i.e., two items more than one's memory span) number of items. Subjects were given unlimited time to study the 2 lists in each span category. The researchers found the overall accuracy of memory-span prediction to be similar for both the younger and older groups. In general, however, younger adults tended to underestimate, whereas older adults tended to overestimate their memory spans. Another interesting aspect of this study did not involve prediction data per se, but referred to the memory monitoring process. The authors incorporated study-time as a 'procedural' type of memory monitoring measure rather than simply relying upon the more salient, or 'declarative' nature of predictions. Post hoc tests were performed on recall readiness accuracy as a function of difficulty level (i.e., increased number of test items). These tests showed that younger adults significantly increased their study time for each increase in task difficulty. Older adults did not significantly increase study time from the sub-span to memory-span lists, but did increase study time for the super-span condition. Murphy et al. (1981) proposed that these results show that the older adults suffer from a memory monitoring deficit. A deficit of this nature prevented them from realizing more study time was required for the long lists.

Bruce et al. (1982) examined an alternative explanation for the Murphy et al. (1981) findings discussed above. This subsequent study was designed to test the hypothesis that age differences in study-time appeared in the Murphy et al. (1981) study because older adults did not find the study materials interesting. To test their hypothesis, Bruce et al (1982) introduced four word lists that differed with respect to imagery and frequency ratings. According to the memory monitoring deficit hypothesis, older individuals should not display knowledge regarding differential difficulty of the word lists and should not study such lists for significantly different lengths of time. The results showed that the elderly overestimated their recall, whereas the younger subjects predicted their recall accurately. Although overall prediction accuracy did vary with word frequency, the results indicated that the younger adults had more accurate memory knowledge than the older adults. Further, the predicted number of recalled items did not differ significantly, but recall performance did. This finding led the authors to suggest that memory knowledge does not accurately reflect changes in memory performance. Moreover, this study did not find significant age differences in study time.

Contradictory findings were achieved within a single study designed by Lovelace and Marsh (1985). The study tested the hypothesis that the contradictory findings discussed above were the result of different prediction measures. The authors designed a study to assess age differences in prediction accuracy using global and item-by-item prediction measures. They hypothesized that the different age groups might vary in their perceptions of task difficulty. These perceptions might result in poor memory monitoring by the elderly subjects because they (a) perceive the task as easier than it actually is, (b) fail to discriminate the more memorable items from less memorable items, and/or (c) inaccurately evaluate the correctness of their responses. Lovelace and Marsh employed 60 word pairs which were selected to be high frequency words, imageable, and a maximum of six letters in length. No word in the pool was a strong associate of another word. The results indicated the importance of distinguishing between different types of memory monitoring measures. First, monitoring the relative memorability of individual items at the time of study showed no age difference. The item-by-item predictions were equally accurate in each age group. This first finding supported the results of the Rabinowitz et al. (1982) study discussed above. Second, judgements of the overall difficulty of the task (i.e., a global performance prediction) demonstrated that the younger subjects were quite accurate in determining such aspects of the task, whereas the elderly were not. The elderly overestimated the number of items that they correctly recalled. This second result supported the findings of the Bruce et al. (1982) and Murphy et al. (1981) studies. On the basis of their findings, Lovelace and Marsh proposed that the memory monitoring deficit results from either, (a) older persons not incorporating memory decline into global memory performance estimates, or (b) older persons may not realize or incorporate task difficult into global performance predictions.

The studies that investigated age differences in prediction accuracy showed that, under certain circumstances, the elderly did display a metamemorial deficit, whereas under other circumstances they did not. It is difficult to interpret these contradictory findings without considering the entire prediction process. Recently researchers have begun to examine the role of MSE in global performance predictions (Hertzog et al., 1990b; McDonald-Miszczak et al., 1990; Saylor, 1990).

Memory Self-Efficacy and Predictions

According to Hertzog et al. (1990a) one key to appreciating the relationship between memory self-efficacy and performance predictions may be to distinguish general or global capacity beliefs from efficacy judgements in a particular context (i.e., a specific memory task). Since questionnaires attempt to access an individual's beliefs and knowledge regarding everyday memory situations, they may logically be referred to as general memory beliefs due to the breadth of their focus (Hertzog et al., 1990a). Many memory situations are included in a single questionnaire (i.e., the 7 subscales of the MIA questionnaire), such that memory experience on a specific memory task is not measured. On the other hand, task specific measures (i.e., global performance predictions) may be considered an attempt to access knowledge and beliefs specific to a particular memory task (Hertzog et al., 1990a). According to this scheme, one may make the distinction between the memory self-efficacy judgements made on a metamemory questionnaire and the efficacy judgements obtained from performance predictions (Hertzog et al., 1990a).

Hertzog et al. (1990a) proposed that the global prediction process is composed of (a) assessment of one's general memory self-efficacy, (b) assessment of normative task difficulty, and (c) mapping one's self-efficacy onto the assessed task performance distribution. The task situation will determine which component will exert greater influence in the prediction process. For example, in an unfamiliar task situation, it is more likely that individuals will rely on their general memory self-efficacy to make the prediction (e.g., "I do not know how I will perform on this exact task, but in general I believe my memory is quite good"). On the other hand, in a familiar task situation, an individual is more likely to use task specific self-efficacy beliefs (e.g., "The last time I encountered this task I performed very well"). In either case, an individual must incorporate general memory self-efficacy with an appraisal of the task in order to produce a performance prediction. The degree to which questionnaire measures of memory self-efficacy relate to performance predictions will depend on: (a) the extent to which individual differences in appraisal of the task influence performance predictions, and (b) the degree to which questionnaire measures of memory self-efficacy (e.g., responses on the MIA capacity subscale) correspond with self-efficacy beliefs regarding the specific task in question (Hertzog et al., 1990a).

The conceptual scheme developed by Hertzog et al. (1990a) has important implications for understanding the accuracy of performance predictions. These authors suggest that inaccurate performance predictions are the result of, (a) inaccurate assessment of one's general memory self-efficacy, (b) inaccurate appraisal of the specific memory task, and/or (c) inaccurate correspondence between general memory self-efficacy and the suspected performance distribution. Inaccurate performance predictions then, may not result entirely from overconfidence about one's memory ability as some studies have suggested (Bruce et al., 1982; Murphy et al., 1981). Inaccurate predictions may also reflect an inaccurate assessment of average performance on the memory task.

A study by Shaw and Craik (1989) has demonstrated the plausibility of the proposed prediction process. The study examined elderly adults' failure to use optimal encoding strategies because of their proposed metamemory deficit (Bruce et al., 1982; Murphy et al., 1981). The type of processing at encoding was varied and the effects were measured on predictions and subsequent recall for younger and elderly age groups. A list of 60 common concrete nouns was used to construct 3 lists for each type of encoding strategy. The types of encoding cues were, (a) lowest depth of processing: a letter cue (e.g., "starts with ic: ice"), (b) moderate depth of processing: a rhyme cue (e.g., "rhymes with dice: ice"), or (c) a category/descriptive cue (e.g., "something slippery: ice"). The same cues were used in the cued recall test which followed a 5 minute vocabulary test. The results showed that the younger and older adults did not differ in their predictions, but did differ significantly in their recall. Lower predictions were made in both age groups for the letter cue words, but no reliable difference was found for the rhyme and category cues. Both age groups were equally insensitive to changes in performance caused by the type of encoding. However, the type of encoding did have an effect on recall. Although the younger subjects recalled more words than the older subjects on average, age differences were not apparent for the category cued words. This eradication of age differences was due to the depth of processing for these words. Overall, the younger adults accurately predicted their performance, whereas the elderly overestimated by 10 percent. These age differences should not necessarily be interpreted as overconfidence in the elderly. The study indicated that the propensity to underestimate or overestimate is somewhat dependent on task conditions. For instance, both the younger and older adults overestimated when performance was low and underestimated when performance was high. This result showed that task conditions, or task appraisal, can affect one's tendency to underestimate or overestimate performance.

A study by Devolder, Brigham, and Pressley (1990) also demonstrated the importance of task appraisal in the prediction process. Age differences in memory monitoring ability were examined in a series of three studies. The authors hypothesized that the contradictory results of previous prediction studies may stem from, (a) different memory tasks being used in different studies, (b) varied familiarity of memory tasks, and (c) differential performance levels on tasks. They included 7 different memory tasks: (a) recall of 30 concrete nouns, (b) recognition of 30 concrete nouns, (c) recall of underlined words and phrases from a prose passage, (d) face-name learning, (e) keeping appointments to make phone-calls at specified times, (f) a vocabulary test, and (g) a digit span task. Study 1 included 48 younger ($\underline{M}$ = 28 years) and 48 older ($\underline{M}$ = 71 years) subjects who performed the word recall and recognition tasks. Study 2 tested 48 younger ($\underline{M}$ = 28 years) and 48 older ($\underline{M}$ = 69 years) subjects on word recall, word recognition, face-name learning, and appointment keeping. Study 3 included 40 younger (M = 31 years) and 40 older (M = 70 years) subjects who were tested on prose recall, vocabulary, and digit span tasks. Half of the subjects in each age group were asked to make global performance predictions prior to performing the task, whereas the other half were asked to make global performance postdictions after the task had been completed. Age differences were observed in subjects' prediction accuracy scores, but not in their postdictions. Specifically, younger subjects in Study 2 predicted their word recognition and face-name learning performance more accurately than the elderly subjects. Younger subjects in Study 3 predicted their vocabulary performance more accurately than the elderly, whereas the elderly predicted their appointment-keeping performance more accurately than the younger subjects. There was also a general absence of significant age differences between subjects' propensity to overestimate or underestimate their performance. These findings contradict those of other researchers (Bruce et al., 1982; Murphy et al., 1981) who hypothesized that the elderly do not successfully monitor their memory decline. However, these findings support the Hertzog et al. (1990a) contention that task appraisal is an important aspect of the prediction process since all age differences apparent in this study occurred before task experience (i.e., predictions), whereas no age differences were apparent after direct task experience (i.e., postdictions).

A study by Hertzog et al. (1990b) was designed to address the relationships between general memory self-efficacy, performance predictions, and memory performance. The design of this study utilized three age groups (young M = 37 years; middle aged M = 53 years; old M = 67 years), word lists and text memory tasks, multiple recall of tasks, and a correlational analytic approach. This type of design enabled the researchers to investigate their assumption that the prediction process is influenced by a multitude of variables.

First, the authors hypothesized that the correlation between general memory self-efficacy and performance predictions would increase when subjects are provided with normative performance information regarding the specific memory task. Without such normative information, it would be difficult for subjects to understand how their own memory ability might coincide with the difficulty of the specific task. Such a lack of understanding would possibly result in a low correlation between one's general memory self-efficacy and one's predicted performance. Subjects were told that the average performance on the word list task was 15 out of 30 words, and that the average performance on the text recall task was 25 out of 50 ideas. The normative performance values were arbitrarily chosen and did not reflect estimates of expected performance levels. By providing subjects with normative information, the authors attempted to minimize individual differences in assessments of task difficulty.

Second, subjects were hypothesized to utilize on-line awareness of performance over trials to evaluate their performance on the task and to form task specific self-efficacy beliefs. This on-line processing should improve prediction accuracy over trials, but decrease the relationship between global self-efficacy and performance predictions. Thus, before one experiences the specific task it is likely that general memory self-efficacy might be used to predict performance, whereas after task experience it is likely that previous task experience will be used to predict performance.

Multiple measures of memory self-efficacy were obtained. Both the MIA and MFQ metamemory questionnaires were used to assess general memory self-efficacy. Global prediction measures were obtained on each of the three trials for both memory tasks (e.g., word lists and text recall). A recall measure was obtained for each trial from either the word list or text, depending on which was used, so that the accuracy of one's memory self-efficacy could be assessed.

The Hertzog et al. (1990b) study found age differences in predictions when individuals were given normative task information. Older adults predicted poorer recall performance than younger adults. The authors suggested that this pattern of mean age differences in predictions is similar to patterns of age differences in questionnaire measures of MSE.

A salient relationship was found between MSE and the first prediction of word recall performance. Moreover, as an individual performed more trials, the relationship between MSE and performance predictions weakened and was replaced by past predictions and past performance. This finding supports the authors second hypothesis that general memory self-efficacy is more important on the first prediction when the individual has not yet had a chance to directly experience the memory task. Following direct experience, however, subsequent predictions are more influenced by task specific experience.

Overall, the findings have suggested that prediction overestimation by older adults may be a function of inaccurate task assessments. Thus, results of other studies (Bruce et al., 1982; Lovelace & Marsh, 1985; Murphy et al., 1981) were probably not due to higher task-specific efficacy expectations regarding memory performance as previously suggested. Instead, they may have been a result of inaccurate task assessment. The Hertzog et al., (1990b) findings were most robust for word list recall, but not significant for story recall. The authors hypothesized that differential findings for the word list and story were due to the subjects' greater difficulty in evaluating post-recall performance in the story condition.

A study was conducted by McDonald-Miszczak et al., (1990) that examined the influence of differential task experience on the prediction process. They hypothesized that previous task experience may affect subjects' predictions differently depending on whether the initial task was easy or difficult. This hypothesis was tested by having subjects make predictions and their recall on either an easy (e.g., 15 concrete nouns), or a difficult (e.g., 45 concrete nouns) word list prior to a moderate list (e.g., 30 concrete nouns). The multidimensionality of the prediction process was also addressed by having the participants complete the MIA questionnaire so that the role of general memory self-efficacy beliefs could be examined. The participants in this study (N = 128) consisted of equal groups of younger males and females ($\underline{M}$ = 23 years), and elderly males and females ($\underline{M}$ = 66 years). Equal numbers of younger and older males and females were randomly assigned to receive either a difficult word list or an easy word list on the first trial. All subjects were required to provide self-generated performance norms for younger and older adults prior to predicting their own performance. Subjects then made their own global performance predictions, studied the word list, and recalled the list. On the second trial, all subjects received a moderately difficult 30 word list. Subjects again provided self-generated performance norms for younger and older adults, made global performance predictions for the moderate list, studied the word list, and recalled it. Global predictions for the moderate list were compared for differential effects of previous task experience. The results showed that all subjects underestimated their performance, but younger females were the least accurate. Subjects who received the difficult list first gave higher predictions and were more accurate than those who received the easy list. Analyses showed that general memory self-efficacy was an important influence on first trial predictions, whereas second trial predictions were influenced by previous experience. The effects of differential task experience showed that receiving a difficult word list raised subjects' global predictions on a subsequent easier list. The study also replicated the finding of Hertzog et al. (1990b) that naive predictions are based on one's general memory self-efficacy beliefs, whereas informed predictions are based on previous task experience. Thus, the prediction process can be altered by one's task experience (e.g., task knowledge), but such alterations depend on the type of experience one encounters.

Saylor (1990) also investigated the relationship between MSE and predictions. The study examined (a) the role of MSE in performance predictions, (b) the process of upgrading memory predictions to resemble actual performance, and (c) the effect of prior normative information (e.g., 15 out of 30 words) on prediction and prediction accuracy. MSE was measured with the MIA and the MFQ. Memory monitoring was measured by including two memory trials in the design and requiring subjects to make 3 predictions on each trial. Predictions were made (a) prior to list study, (b) following list study but before recall, and (c) after recall. Multiple prediction measures were included in order to examine when predictions were upgraded to resemble actual memory ability, (a) after study of the lists (memory monitoring), or (b) after recall of the lists (task experience). Predictions were either made after receiving normative performance information (the prior information condition), or without such information (no prior information condition). Arbitrarily chosen normative task information was included to decrease the variability of individuals' task assessments and thereby increase the correlation between first predictions and MSE. Saylor obtained some results that were consistent with the Hertzog et al. (1990b) assertion that predictions are task specific MSE judgements.

Measures of MSE were significantly correlated with the first prediction as in the Hertzog et al. (1990b) study. There were significant relationships between subsequent predictions and recall. The relationship between MSE and performance predictions weakened as task experience was gained. This finding shows that general memory beliefs are important for making a prediction when one has not had any task experience. However, as one understands the nature of the task materials one relies on such understanding to make subsequent predictions. Hence, MSE is most important to the prediction process when one has not had any direct experience with the task materials.

Second, significant upgrading of predictions occurred between the first and second prediction on each trial prior to recall experience. The second prediction then, is not entirely based on MSE, but is also based on memory monitoring as a result of studying the word lists. The results clearly show a distinction between memory monitoring and task experience.

Third, contrary to the hypothesis, prior information did not have a significant effect on the relationship between MSE and predictions. It did significantly influence the older adults' first predictions, but this influence disappeared for subsequent predictions. Saylor speculates that although prior information appears to lower the variance of the prediction estimate, it was probably not effective because it was not accurate, or age-graded. The information then, probably affected the age groups differently by lowering estimates given by the young and raising estimates given by the elderly.

Finally, the elderly were found to predict their performance more accurately than the young. This finding is contrary to much of the prediction literature (Bruce et al., 1982; Lachman et al., 1979; Lovelace and Marsh, 1985; Perlmutter, 1978; Rabinowitz et al., 1982). The study also found a slight tendency in the elderly to overestimate and a greater tendency in the young to underestimate. Thus, the elderly group was more accurate in their predictions in both the prior information condition and the no information condition. Saylor noted, however, that prediction accuracy in the no information condition may be an artifact of the elderly subjects choosing the middle number of 15. Since their performance was close to 15 in both conditions this tendency may have led to more accurate predictions.

The next logical step then, is to study the relationship between MSE and predictions when accurate task information is provided to young and elderly groups. A study of this kind is needed to examine the relationship between MSE and individuals' task assessment. In past studies, an average performance estimate has been provided, but this estimate was not based on actual data and was the same for both age groups (Hertzog et al., 1990b; Saylor, 1990). A study that provides accurate and age-graded information might clarify the role of normative task information in the prediction process.

Summary

The development of the prediction process has paralleled the development of the metamemory concept. The prediction paradigm has grown from being primarily used as a direct measure of memory knowledge to include aspects of task specific MSE. A prediction is a judgement about a task, made before one experiences the entire task. There are two primary types of prediction measures used in cognitive aging research: (a) global predictions, and (b) item-by-item predictions. The accuracy of one's prediction has typically been interpreted as the degree of one's memory knowledge. Some researchers have found that the elderly are as accurate as the young in predicting their memory performance (Lachman et al., 1979; Perlmutter, 1978; Rabinowitz et al., 1982), whereas others have found that the elderly tend to overestimate their performance (Bruce et al., 1982; Murphy et al., 1981). A study by Lovelace and Marsh (1985) has attributed these contradictory results to the type of prediction measure used. It was found that global predictions are typically associated with age differences.

Hertzog et al. (1990a) have proposed that global performance predictions are a measure of task specific MSE. The proposed prediction process is composed of, (a) assessment of one's general MSE, (b) assessment of normative task difficulty, and (c) mapping one's self-efficacy onto the assessed task performance distribution. This conceptual scheme has important implications for understanding the accuracy of performance predictions. Inaccurate predictions may not result entirely from overconfidence as has been argued in the past (Bruce et al., 1982; Murphy et al., 1981). Prediction inaccuracy may also reflect an inaccurate assessment of average performance on the memory task (Hertzog et al., 1990b). Recent studies have shown support for the proposed prediction process and the importance of task assessment (Devolder et al., 1990; Hertzog et al., 1990b; McDonald-Miszczak et al., 1990; Saylor, 1990; Shaw & Craik, 1989). Further research is required to examine the effect of age-graded task information on prediction accuracy.

Objectives and Hypotheses

The present study addresses the relationship between MSE, performance predictions, and recall. The intent was to investigate the interaction between general memory self-efficacy beliefs and specific task information when making predictions using the conceptual scheme outlined above by Hertzog et al. (1990a). Such a scheme asserts that the prediction process consists of both general memory self-efficacy beliefs (MSE) and task specific self-efficacy. General memory self-efficacy beliefs are thought to arise from various memory experiences and performances that have been generalized across time. Task specific self-efficacy, on the other hand, is thought to arise from specific knowledge regarding a particular memory task (Hertzog et al., 1990b). As noted above, a manipulation of normative age graded information is needed to fill gaps in past experimental findings and will, therefore, provide the focus of this study.

Age-Graded Normative Task Information Manipulation

Two types of age-graded normative task information were used in this study, (a) accurate age-graded normative information, and (b) self-generated normative information. These two types of information were used due to hypothesized differences in the quality of information that each provided. In short, the accurate information was hypothesized to be much richer in quality than self-generated information.

Accurate normative information was obtained from a similarly designed study conducted by McDonald-Miszczak et al. (1990) which tested similar age groups. The previous sample consisted of younger (19-30 years) and older (54-77 years) males and females. Since the present study only tested females, the relevant younger and older age groups that participated in the previous study ranged from 20 to 30 years ($\underline{M}$ = 23 years) and 54 to 77 years ($\underline{M}$ = 65 years) respectively. The recall scores were used from the previous study's moderate 30 word list to construct the accurate age-graded task information required for the present investigation. These scores provided the most accurate information since subjects were tested on the same words contained in the present study's word lists. Although sampling differences were expected to create some degree of inaccuracy given the specific nature of the point estimate (e.g., the average young female remembers 24 out of 30 words), the McDonald-Miszczak et al. (1990) study offered the most comparable data available.

The second type of normative information was also based on the McDonald-Miszczak et al. (1990) study. Subjects were asked to give their own performance estimates for younger and older females. This information was predicted to invoke conscious consideration of task difficulty that may not have occurred in the No Normative Information Condition. It was proposed that requiring subjects to provide task norms would provide a moderate condition involving both an evaluation of task difficulty and general MSE. Thus, in contrast, the control condition (e.g., No Norm Information) was proposed to primarily invoke subjects' individual differences in task assessment when predicting performance, whereas the accurate information condition was proposed to primarily invoke general MSE. The Self-Generated Norms Condition was specifically included to provide a mix of both objective consideration of task difficulty and general MSE.

Experimental Hypotheses

It was hypothesized that subjects would show improved prediction accuracy with task experience (e.g., across trials). This has been shown in previous research (Hertzog et al., 1990b). However, the rate of improvement should depend on the amount of initial task information subjects' receive. Subjects who receive the least task information should show the greatest improvement since they must rely on task experience to become familiar with the task (e.g., No Normative Information Condition). On the other hand, subjects who receive accurate task information should begin with an adequate understanding of task demands and therefore demonstrate the least improvement over time (e.g., Accurate Norms Condition). Since the Self-Generated Information Condition was intended to be the moderate information condition, subjects in this condition should show less improvement than those who receive no information, but show more improvement than those who receive accurate information.

An interaction between age group and normative task information condition was also hypothesized for prediction accuracy. Different effects were hypothesized depending on the task information condition to which subjects were assigned. These differences stem from contradictory results based on previous research that employed similar task information conditions. First, it was hypothesized that subjects who do not receive task information prior to making their own performance predictions (i.e., No Normative Information Condition) will predict their performance least accurately, but the younger subjects will estimate more accurately than the older subjects (Bruce et al., 1982; Lovelace & Marsh, 1985; Murphy et al., 1981). Second, it was hypothesized that subjects who are requested to provide their own age-graded normative task information (i.e., Self-Generated Information Condition) will predict their performance more accurately than those not provided with any normative information. However, elderly subjects in the Self-Generated Norms Condition have been found to estimate their performance more accurately than the younger subjects (McDonald-Miszczak et al., 1990). The hypothesized effect for the Self-Generated Norms Condition is contrary to the effect hypothesized for the No Norms Condition since requiring subjects to consciously consider normative task performance may change how they map themselves onto the proposed performance distribution. Finally, the smallest or null age effect on prediction accuracy should appear when subjects receive accurate age-graded normative task information (Accurate Norms Information). In this final condition, both age groups should benefit from the normative information that is provided and most accurately map themselves onto the proposed performance distribution.

A primary expectation of this investigation was that general MSE will significantly determine an individual's first prediction, whereas subsequent predictions will be determined by previous task experience. This hypothesis was based on studies by Hertzog et al. (1990b), McDonald-Miszczak et al. (1990), and Saylor (1990). When a task is unfamiliar, one's general MSE will be used to predict performance, because specific task performance information is not available. Subsequent predictions, however, will be guided by past experience.

The predicted effect of general MSE on first trial predictions should depend on the availability of normative task information. This hypothesis stems from the Hertzog et al. (1990a) proposal that the relationship between general MSE and performance predictions will be stronger when individual differences between subjects' assessments of task difficulty are reduced. Thus, subjects who are provided with age-graded normative task information before giving a performance prediction should utilize general MSE to a greater degree for their first trial prediction, compared to subjects who are not provided with accurate age-graded normative task information.

CHAPTER III

METHODS

Design

The study consisted of a 2 (Age) x 3 (Task Information Conditions) x 3 (Trials) factorial design with repeated measures on the last factor. The Age factor consisted of younger and older adult subjects. The normative task information factor consisted of a control condition (No Normative Task Information), a Self-Generated Normative Task Information Condition, and an Accurate Age-Graded Normative Task Information Condition. Within each age group, subjects were assigned at random to one of the three task information conditions. Each subject received three trials on the experimental task.

Subjects

The sample consisted of 36 women aged 20 to 30 years and 36 women aged 65 to 75 years drawn from the urban area of Victoria, BC. Only female participants were used in order to control for potential gender effects. The younger sample was recruited from the undergraduate subject pool at the University of Victoria. The older sample was recruited through appeals to community organizations requesting volunteers for research in memory and aging. The majority of the elderly participants were retired or working as volunteers (91.6%), whereas most of the younger subjects were full-time students or working full-time (83.34%). On a five-point rating scale, 97.2% of the young and 86.1% of the elderly rated their overall health as good or very good. Table 1 shows other sample characteristics for each age group.

Insert Table 1 about here

Materials

Personal Data Questionnaire. This questionnaire requested demographic and health information. The questionnaire asked 20 questions regarding (a) age, (b) marital status, (c) education, (d) past and present employment, and (e) health. The age, education, and health variables were used in the analyses. The marital status and employment variables were excluded from further analyses since they were not directly relevant to the focus of the study.

Depression (CES-D). Since depressive affect might influence subjects' prediction or recall scores, the Center for Epidemiologic Studies-Depression (CES-D) scale was introduced to determine subjects' affective state at the time of the study. This 20 item scale is designed to measure subjects' current depressive symptoms with an emphasis on depressed mood (Radloff, 1977). A 4-point Likert rating scale is used to obtain responses to questions like, "During the past week I had a crying spell." The CES-

D has high internal consistency (an average of .85 using the coefficient alpha and the Spearman-Brown split-halves methods), adequate test-retest reliability, and shows evidence for construct validity.

Vocabulary. Verbal comprehension was measured using Part 1 of V4 from the ETS Kit of Factor Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976). The test consists of 18 multiple choice items. Subjects were required to circle the corresponding number to a word with the same meaning or nearly the same meaning as the underlined target word. Participants were given 4 minutes to complete the test.

Memory Self-Efficacy (MSE). The Capacity subscale from the Metamemory In Adulthood Questionnaire (MIA) (Dixon & Hultsch, 1983) was used to measure general memory beliefs. The Capacity subscale asks subjects to report their perceived ability on a variety of everyday memory tasks (e.g., "I am good at remembering names.") The subscale shows high internal consistency with an average Cronbach's alpha over five samples of .84 (Hultsch et al., 1988). Hertzog et al. (1987) have shown that the Capacity subscale is the best single marker of MSE in the questionnaire.

Normative Age-Graded Task Information. Three levels of normative age-graded task information were used in this study. These levels were designed to represent different amounts of task knowledge. Differing amounts of task knowledge were expected to affect subsequent prediction and prediction accuracy scores.

The No Information Condition did not provide subjects with any information concerning expected performance levels for younger and older females prior to making their own performance predictions. This condition is typical of the circumstances under which global performance predictions have been traditionally obtained (Bruce et al., 1982; Lovelace & Marsh, 1985; Murphy et al., 1981).

In the Self-Generated Norms Condition, subjects were first asked to provide an estimate of normative performance for the average younger woman followed by an estimate for the average older woman prior to making their own performance predictions. Specifically, participants were asked "How many words out of 30 do you think the average woman 20 to 30 years of age would remember?" Then they were asked "How many words out of 30 do you think the average woman 65 to 75 years of age would remember?" The second condition was based on the prediction measure used in a study by McDonald-Miszczak et al. (1990). In the present investigation, it represented a moderate amount of task information since subjects were required to consciously consider task difficulty before making their own performance predictions.

The third condition provided subjects with accurate age-graded normative task information before making their own performance predictions. Subjects were informed of the average performance for the younger and older age groups. Specifically, participants were informed that, "The average woman aged 20 to 30 years would remember 24 words out of 30 words." Then they were informed that, "The average woman aged 65 to 75 years would remember 17 words out of 30 words." The normative estimates were based on younger and older females' mean recall scores obtained in a similar study conducted by McDonald-Miszczak et al. (1990). This condition was specifically designed to give subjects an expectation regarding the average difficulty of the task for their own age group. Since mean performance levels were presented for both age groups, this condition was also intended to give subjects a feel for task difficulty relative to age.

Word Lists. Three lists of 30 nouns, based on lists used in a previous study by McDonald-Miszczak et al. (1990), were used as the memory task. The lists were developed using the Paivio, Yuille, and Madigan (1968) norms. The words for each list were selected on the basis of four criteria: (a) high frequency (F = 50 or greater), (b) imageable and concrete (I and C or 5.0 or greater), (c) a maximum of 10 letters in length (range 3 to 10), and (d) no word on a list was a strong associate with another word on the list. A counter-balancing procedure was used for the three word lists that resulted in 6 different list orders.

Subjects were randomly assigned to each counter-balancing condition such that an equal number were represented in each age group and task information condition.

Procedure

Subjects were tested in small groups consisting of 2 to 7 people. In a few instances, subjects were tested individually. Testing-group membership was not based on age. Some groups consisted completely of elderly females or young females, whereas other groups consisted of members from both age groups. Testing took place in seminar rooms at the University of Victoria. Before the experiment began, a conscious attempt was made by the experimenter to alleviate any apparent nervousness by making the atmosphere congenial. The entire testing session took between 1 and 1.5 hours.

The testing protocol began with informed consent procedures. Subjects were told the purpose of the study, the procedures to be followed, the total time required, and possible risks (i.e., nervousness regarding memory tests) and benefits (i.e., aiding researchers to increase their understanding of memory and aging). Subjects were given an opportunity to ask whatever questions they may have had and all such questions were answered by the experimenter before informed consent was obtained. Following completion of the consent procedures, participants were asked to complete the questionnaire battery that consisted of (a) the personal data questionnaire, (b) the Center for Epidemiologic Studies-Depression Scale, (c) the ETS vocabulary test, and (d) the Metamemory In Adulthood questionnaire. The tasks were administered to the participants in the order that they are listed above. With the exception of the vocabulary test, all tasks were self-paced.

Each trial of the experimental procedure consisted of four elements, (a) manipulation of task information, (b) predictions, (c) study of the word list, and (d) recall of the word list. The experimenter read the memory task instructions and the subjects followed along with their own written copy. Specifically, subjects were informed that they would be asked to study and remember three different lists of common English words, one at a time, for a specified period of time. They were also informed that they would be required to give their predictions about their recall performance before actually studying the words. Subjects were provided with 9 examples of the kinds of words they would be asked to study (e.g., winter, shoes, charm).

The manipulation of the task information directly followed the general task instructions. On a separate page, subjects were asked to read the instructions silently to themselves. All subjects were told that the following page contained a list of 30 words like those in the example and that they would have 3 minutes to study the words. Subjects in the No Information Condition were then asked to provide their own performance prediction. Subjects in the Self-Generated Information Condition were then asked to provide predictions for the average woman aged 20 to 30 years, the average woman aged 65 to 75 years, followed by their own performance prediction. Subjects in the Accurate Task Information Condition were then provided with the average performance for a woman aged 20 to 30 years, a woman aged 65 to 75 years, and then asked for their own predicted performance. When all subjects had finished their performance predictions, they turned to the next page in their booklet that showed the word list. Following the 3 minute study interval, subjects began their recall. They were given a maximum of 5 minutes to recall words from the word list.

The procedure for the second and third trials was virtually the same as the first trial described above. Specifically, the subsequent trials omitted general task instructions, but repeated the manipulation of task information, performance predictions, list study, and recall. No procedural differences were introduced between the trials. After completion of the third trial recall, the subjects were fully debriefed regarding the purpose of the study and were informed about the experimental hypotheses.

Dependent Measures

Free recall. Recall performance was measured by the total number of words correctly recalled on each list. Each subject had a separate recall score for each trial. A lenient scoring procedure was followed for pluralization of list words and for obvious inaccurate spellings.

Predictions. Each subject was required to predict the total number of words that would be correctly recalled. Subjects were reminded during each trial prediction to give a single number out of 30 (e.g., 15) rather than a range (e.g., 14 to 16).

Prediction Accuracy. Subjects' performance prediction accuracy scores consisted of the predicted number of recalled words on a particular trial minus the actual number of recalled words for that trial. Such scores have been used in past prediction studies (Bruce et al., 1981; Lovelace and Marsh, 1985). These difference scores were calculated separately for each trial. A positive difference score indicated an overestimate of performance, whereas a negative difference score indicated an underestimate.

CHAPTER IV

RESULTS

The data were analyzed in three separate stages. The first stage evaluated potential differences among the age groups, task information conditions, and word list order subgroups in such descriptive characteristics as health, education, vocabulary ability, depressive affect, and general memory self-efficacy beliefs. Such differences may influence the interpretation of subsequent analyses. The second stage of analyses focused on age and task information differences in word list recall, prediction, and prediction accuracy scores across the three trials of measurement. The purpose of the second series of analyses was to identify the effect of experience on each set of dependent variables for each age group and normative task information condition. All analyses in this second series were conducted as repeated measures analyses. The third stage was conducted to examine the variables that may influence subjects' predictions before and after task experience. These analyses were performed as a replication of the analyses conducted by Hertzog et al. (1990b).

Stage One

The first stage of analyses was conducted in order to examine the characteristics of the sample used in this investigation. Specific measures such as subjects' self-rated overall health, number of years of education, vocabulary ability, depressive affect scores, and general memory self-efficacy beliefs were obtained for two specific reasons. First, age is a variable of focus in this study. However, the age variable is very general and does not offer any direct explanation as to why the age groups may differ on certain dependent variables (e.g., recall ability). Thus, general information was collected and analyzed as an attempt to account for possible age differences with more meaningful measures. Secondly, sample characteristics were collected to test whether the random assignment of subjects' to each of the three different task information conditions and to the word list order subgroups was performed successfully. Subjects' within each age group should not differ significantly on descriptive characteristics across the conditions and subgroups. Significant variability across conditions and subgroups would lead to difficulty in interpreting any significant effects of the independent variables on the dependent variables.

It should be noted that the primary motive for collecting descriptive data was to test the random assignment of subjects. Although the descriptive measures were also used to describe subsequent age differences, this was not the primary motivation of this study. Thus, there are only a few descriptive variables used in this study and their inclusion was not based on an a priori theory. Instead, the variables were chosen for their perceived general relevance to other variables in this study. All of the sample characteristics were examined in two separate analyses.

Sample Characteristics

The first step in analysis of sample characteristics was to examine correlations between the descriptive variables, independent variables, and the dependent variables. These correlations are shown in Table 2. The results indicate that subjects' self-rated overall health, education and MIA-Capacity relate to all of the dependent variables at least on some trials of measurement. For example, subjects' self-rated health scores are significantly related to subjects' prediction accuracy scores only on the second trial of measurement, but show a strong relationship with both recall and prediction scores across all trials. Subjects' vocabulary scores are only significantly related to third trial global predictions and their CES-D scores are not significantly related to any of the dependent variables. These correlations indicate that some of the descriptive variables should be examined more closely in analyses performed with the dependent variables. In particular, the descriptive variables that correlate with both age and the dependent variables may be used to provide a clearer description of obtained age differences in subsequent analyses (i.e., repeated measures multivariate analyses of variance described in stage two). Thus, the findings described here will be referred to in stage two of the analyses.

Insert Table 2 about here

The second analysis, performed in order to examine the characteristics of the sample, was a multivariate analysis of variance (MANOVA). The MANOVA was conducted to examine age, task information, and word list order differences in descriptive characteristics of the sample. The focus of this second analysis was to test the success of the random assignment of subjects to task information conditions and to the counter-balanced word list subgroups. Age differences on the descriptive variables were also examined. The analysis was conducted on five descriptive variables: self-rated overall health, number of years of education, vocabulary, CES-D scores, and MIA-Capacity scores. The mean for each descriptive variable attained by each age group is shown in Table 1.

The multivariate test of significance for the main effect of age was highly significant (Pillais F (5, 31) = 26.99, $p < .01$). Univariate analyses revealed that the young subjects reported more depressive affect (F (1, 35) = 5.74, p, < .05), higher MIA-Capacity (F (1, 35) = 21.79, $p < .01$), more years of education (F (1, 35) = 18.23, $p < .01$), and rated themselves as slightly better in their overall health (F (1, 35) = 3.95, p = .05) than the elderly subjects. The elderly subjects scored higher on the vocabulary test (F (1, 35) = 14.51, p < .01) than the young subjects. All other multivariate tests were not significant (p > .05). The lack of significant task information condition or word list order subgroup differences shows that the random assignment procedure was successful for all the descriptive variables.

Stage Two

The purpose of the analyses conducted in stage two was to examine age and task information condition differences in the dependent variables. Thus, separate analyses were performed on subjects' recall, prediction, and prediction accuracy scores. Since each dependent variable was measured on three occasions, repeated measures analyses were used.

Age and Task Information Differences in Recall

The second stage of the analyses began with an examination of subjects' recall scores. A repeated measures MANOVA was performed on subjects' recall scores across the three trials of measurement. The purpose of this analysis was to examine whether there were age group or task information condition differences in the number of words recalled and whether the number recalled changed across trials. Specifically, age, task information condition differences, and an Age x Task Information condition interaction were examined. A significant between subjects main effect of age was found (*F* (1, 66) = 90.28, *p* < .01). Overall, young subjects recalled more words than elderly subjects. The mean recall scores are shown as a function of age and trial of measurement in Table 3. No significant effects of change across trials of measurement for either age, task information condition, or Age x Task Information condition interaction were obtained (*p* > .05).

Insert Table 3 about here

Age and Task Information Differences in Predictions

A repeated measures MANOVA was conducted on subjects' predictions. The purpose of this analysis was to examine whether prediction scores differed according to the subjects' age group and task information condition. The interaction term (e.g., Age Group x Task Information condition) was also included in the analysis. The effects of task experience were examined for subjects' global predictions across the three trials of measurement. The between subjects main effect of age was significant (*F* (1, 64) = 48.88, *p* < .01). Overall, the young subjects made higher performance predictions than the elderly. Both young and elderly subjects' mean predictions are shown in Table 3. The other between subjects main effects of task information condition and Age x Task Information condition were not significant (p > .05).

The multivariate test was significant for the Trials x Age interaction (Pillais *F* (2, 63) = 16.02, *p* < .01). Single degree of freedom polynomial tests showed that both the linear trend (*F* (1, 64) = 29.89, *p* < .01) and the quadratic trend (*F* (1, 64) = 8.39, *p* < .01) were significant. The average prediction made by both young and elderly age groups is shown for the three trials of measurement in Figure 1. The data show that the young make higher predictions than the elderly and that their predictions increased across trials, whereas the elderly's predictions slightly decreased across trials.

Insert Figure 1 about here

The multivariate test was also significant for the Trials x Task Information interaction (Pillais *F* (4, 128) = 2.55, *p* < .05). Single degree of freedom polynomial contrasts showed that the linear trend was significant (*F* (2, 64) = 5.06, *p* < .01). Figure 2 depicts the average prediction made by each task information condition across the three trials. Subjects who were initially given accurate normative task information made higher predictions on trials 1 and 2, and slightly decreased their predictions across all three trials than subjects who were not given any normative task information or who gave self-generated information. Subjects in conditions that made lower predictions on trials 1 and 2 increased their predictions on each subsequent trial. By the third trial, subjects in all conditions made similar predictions.

Insert Figure 2 about here

Age and Task Information Differences in Prediction Accuracy

Another repeated measures MANOVA was conducted on subjects prediction accuracy scores across the three trials of measurement. This analysis addressed the primary hypotheses of the study. It was initially hypothesized that there would be a between subjects interaction effect of subjects' age group and task information condition for their prediction accuracy scores. Moreover, it was also hypothesized that the degree of improvement in subjects' prediction accuracy scores across the three trials would differ according to subjects' task information condition. This analysis examined the main effects and interaction effect of subjects' age group and task information condition on their prediction accuracy scores. The between subjects main effects of age (F (1, 64) = 13.16, $p < .01$) and task information (F (2, 64) = 5.61, $p < .01$) were both significant. Overall, the elderly subjects were more accurate than the young subjects, and all subjects in the accurate information condition were more accurate than subjects in the other task information conditions. The hypothesized interaction effect was not significant ($p = .06$). Mean scores for each age group and task information condition are displayed in Table 4. This table shows that although the Age x Task Information effect was not statistically significant it does show a trend towards significance ($p = .06$).

Insert Table 4 about here

The multivariate test for the Age x Trials interaction was significant (Pillais F (2, 63) = 9.84, $p < .01$). The single degree of freedom polynomial contrasts showed a significant linear trend (F (1, 64) = 19.40, $p < .01$). The data in Figure 3 display the age groups' average prediction accuracy for each trial of measurement. The elderly remain quite accurate across all three trials even though there is a slight shift from overestimating their performance on trials 1 and 2 to underestimating their performance on trial 3. The young subjects greatly underestimate their performance especially on trial 1 and to a lesser degree on trials 2 and 3. By the third trial, the age groups' prediction accuracy is quite similar. The hypothesized Task Information condition x Trials interaction was not significant ($p > .05$).

Insert Figure 3 about here

Analysis of age effects

A significant between subjects main effect of age was found in each repeated measures MANOVA performed on the dependent variables. Age, however, is not a meaningful variable to describe group differences. Age differences can stem from a number of variables. Other measures (e.g., health, education, affect) may more clearly describe the obtained age group differences. Thus, further analyses were conducted to test whether the age variable still accounts for a significant proportion of the variance in the dependent variables after relevant descriptive variables have been examined. The purpose of the following analyses was to specify the nature of the general age group differences found for each of the dependent variables using the descriptive variables.

Hierarchical regression analyses were performed on subjects' mean scores across the three trials for each dependent variable for each subject. For example, each subject's recall scores for all trials were summed and divided by the number of trials in order to obtain a single (i.e., best) estimate of overall recall performance. This averaging procedure, performed on all dependent variables, simplified the analyses examining the between subjects main effects of age.

The first block of each hierarchical regression included the descriptive variables that significantly correlated with the dependent variable of interest (see Table 5). This table shows that the correlations obtained between subjects' mean scores on the dependent variables and the descriptive variables are similar to correlations for each trial shown in Table 2. The second block included the age variable. The change in R^2 was then tested to see whether the age variable added a significant amount of variance to the dependent variable beyond the variance already explained by the descriptive variables.

Insert Table 5 about here

Recall Scores. The descriptive variables included in the initial block of the regression procedure all had a significant relationship with both the age variable and subjects' average recall scores (see Table 5). Subjects self-rated overall health, years of education, and MIA-Capacity scores all had correlations with subjects' mean recall scores at the $p < .05$ significance level. The results of the hierarchical regression analysis are shown in Table 6. The descriptive variables explain a significant portion of the mean recall score variance (F (3, 68) =

16.16, $p < .01$). Since the beta weights for education ($T = 4.70$, $p < .01$) and MIA-Capacity variables ($T = 3.41$, $p < .01$) were significant and positive, they reveal that higher levels of education and higher MIA-Capacity scores correlate significantly with higher average recall scores. Table 6 shows that the second stage of the hierarchical analysis resulted in a significant increase in R^2. This significant increase in variance demonstrates that the age variable accounts for a significant proportion of the total variance after variance has been attributed to descriptive variables. The beta weights in the second step revealed that education ($T = 2.46$, $p < .05$) and age ($T = -6.12$, $p < .01$) variables were major contributors to the overall variance. The beta weights for the second step showed that higher education and younger age resulted in higher average recall scores.

Insert Table 6 about here

Prediction scores. A second hierarchical regression analysis was performed on subjects' average prediction scores. The descriptive variables that correlated significantly ($p < .05$) with subjects' average prediction scores were subjects' self-rated health, years of education, and MIA-Capacity scores (see Table 5). The descriptive variables accounted for a significant portion of subjects' average prediction scores (F (3, 66) = 14.18, $p < .01$). An examination of the beta weights showed that subjects' level of education (T = 4.72, $p < .01$) and MIA-Capacity scores (T = 2.85, $p < .01$) were both significant in the initial step of the analysis. Table 6 shows that the addition of the age variable in second step resulted in a significant increase in R^2. Both the education (T = 3.02, $p < .01$) and age (T = -4.02, $p < .01$) variables had significant beta weights in the second analysis. The positive beta weight of education illustrates that subjects' with more years of education made higher average performance predictions, whereas the negative age beta weight shows the young made higher average predictions than the elderly.

Prediction accuracy scores. The correlation matrix depicted in Table 5 shows that the only descriptive variable that correlates significantly with subjects' average prediction accuracy scores is the MIA-Capacity variable ($p < .05$). Thus, MIA-Capacity was the only variable used in the first step of the hierarchical regression analysis. Table 6 depicts the results. The first step resulted in a significant R^2 value (F (1, 68) = 4.19, $p < .05$). The significant beta weight for the MIA-Capacity scores (T = -2.05, $p < .05$) shows that higher MIA-Capacity scores were associated with less accurate predictions. In the second step, age accounted for a significant increase in R^2. The only significant beta weight in the second step was for the age variable (T = 2.56, $p = .01$). The positive beta weight shows that the elderly group was more accurate in their average predictions than the young group.

Summary.

The second stage of the analyses was conducted in order to examine age and task information differences in the dependent variables. Overall, young subjects recalled more words, made higher predictions, and were less accurate in their predictions than the elderly subjects. Young subjects increased their performance predictions and their accuracy with task experience, whereas the elderly slightly decreased their predictions and maintained their prediction accuracy.

Subjects in the accurate task information condition predicted their performance more accurately than subjects in the other task information conditions. However, this condition effect did not significantly interact with subjects' age group. The accurate task information was not more useful to one age group than the other for increasing prediction accuracy although the mean scores indicate a trend. As for the effect of experience, subjects in the accurate task information condition decreased their predictions across trials, whereas subjects in the no information and self-generated task information conditions increased their predictions. By the last trial, subjects in all information conditions were predicting their performance with similar accuracy.

A between subjects main effect of age was apparent for each dependent variable. Since the age variable is not directly meaningful, further analyses were conducted in order to examine the relationship between the descriptive variables and the dependent variables. In each case, age was still found to significantly influence the dependent variable after the relevant descriptive variables were examined.

Stage Three

The third stage of analyses examined the variables that influence subjects' predictions. Separate regression analyses were conducted on each trial in order to examine the effect of specific variables on the prediction process. It was hypothesized earlier that different variables would affect the prediction process before subjects received task experience than after task experience.

Influences on Predictions

The first hierarchical regression analysis was conducted on subjects' first trial performance predictions. It was hypothesized earlier that initial predictions would be affected by subjects general self-efficacy beliefs about their memory ability. The variables most applicable to this hypothesis are the affect laden and belief variables. Subjects' CES-D scores were included in the first step of this analysis since they represent subjects' depressive affect at the time of the study and may directly or indirectly influence subjects' beliefs about their memory ability. MIA-Capacity scores measure subjects' general memory self-efficacy beliefs and are therefore directly applicable to this analysis. The MIA-Capacity scores were entered as the second step of the analysis. The results are shown in Table 7. They indicate that the CES-D variable did not account for a significant portion of subjects' first trial predictions (F (1, 69) = 1.71, p > .05) in the first step of the analysis. The second step of the analysis showed that the total variance contributed by the two variables was almost statistically significant (F (2, 68) = 2.83, p = .066). However, the beta weight for the MIA-Capacity variable was significant (T = 1.97, p = .05). The beta weight for the CES-D variable was not significant (p > .05). The addition of the MIA-Capacity variable, almost resulted in a significant increase in R^2 (critical F_{inc} (1, 65) = 3.99).

Insert Table 7 about here

The first regression analysis was repeated with the addition of an MIA-Capacity x Task Information condition interaction term. This interaction term was added in order to examine whether subjects in the different task information conditions differentially utilized their general MSE to make their first trial predictions. It was hypothesized above that subjects who were less variable in their individual assessments of the task (e.g., subjects who received accurate normative information) would utilize their general MSE more than subjects who were more variable (e.g., subjects in the No Norms and Self-Generated Norms Conditions) for first trial predictions. The MIA-Capacity x Task Information condition interaction term was not significant ($p > .05$). However, correlations between first trial predictions and MIA-Capacity scores as a function of age and task information condition do show a trend. Table 8 shows that younger subjects in the Accurate Norms Condition do exhibit a higher correlation between their general MSE and first trial predictions than younger subjects in the other two task information conditions. Elderly subjects do not exhibit such a clear trend.

Insert Table 8 about here

A second hierarchical regression analysis was conducted on subjects' second trial performance predictions. This analysis also utilized MIA-Capacity and CES-D variables along with subjects' recall and performance predictions on trial 1. This analysis tested whether belief and affect laden variables or previous task experience would contribute a significant amount of variance to subjects' second trial performance predictions. The CES-D variable was entered first in order to control for possible depressive effects on subjects predictions, recall, or general memory self-efficacy. Subjects' first trial predictions and recall scores were then entered as the second step. Lastly, subjects MIA-Capacity scores were entered in order to test whether subjects' memory self-efficacy beliefs account for a significant portion of their second trial predictions after previous task experience has been considered. The results are shown in Table 7. The CES-D variable did not account for a significant portion of the variance in the first step of the analysis (F (1, 68) = 1.94, $p > .05$). The addition of subjects' first trial prediction and recall scores resulted in a highly significant R^2 (F (3, 66) = 79.62, $p < .01$). The beta weights for subjects' prediction (T = 7.62, $p < .01$) and recall (T = 15.51, $p < .01$) scores were both significant. The beta weight for the CES-D variable was not significant ($p > .05$). The addition of the MIA-Capacity variable did not result in a significant increase in R^2. The beta weights for subjects' first trial performance predictions (T = 7.53, $p < .01$) and recall scores (T = 14.19, $p < .01$) were both significant. The beta weights for both MIA-Capacity and CES-D scores were both insignificant ($p > .05$). This analysis indicates that second trial predictions are more heavily based on previous task experience (i.e., first trial performance predictions and recall scores). General memory self-efficacy beliefs (i.e., MIA-Capacity scores) do not account for a significant portion of the variance beyond the variance already explained by previous task experience variables.

The third hierarchical regression analysis was conducted on subjects' third trial performance predictions. This analysis included all variables described in the first and second regression analyses and subjects' second trial performance predictions and recall scores. This analyses tested the same assumption that was outlined in the second regression. This final analysis included (a) CES-D scores, (b) first trial prediction and recall scores, (c) second trial prediction and recall scores, and (d) MIA-Capacity scores. The results are shown in Table 7. Subjects' depression scores did not result in a significant initial R^2 (F (1, 68) = 0.64, $p > .05$). The addition of first trial prediction and recall scores resulted in a significant R^2 (F (3, 66) = 79.62, $p < .01$). The beta weights for subjects' first trial prediction scores (T = 3.66, $p < .01$) and recall scores (T = 12.02, $p < .01$) were both significant. The beta weight for the CES-D variable was not significant ($p > .05$). Next, the inclusion of subjects' second trial prediction and recall scores resulted in a significant R^2 (F (5, 63) = 118.91, $p < .01$). The beta weights for subjects second trial prediction scores (T = 5.79, $p < .01$), recall scores (T = 6.89, $p < .01$), and first trial recall scores (T = -2.04, $p = .05$) were all significant. Lastly, the inclusion of the MIA-Capacity variable did not result in any change in R^2. The beta weights for subjects' second trial performance predictions ($T = 5.74$, $p < .01$), recall scores ($T = 6.42$, $p < .01$), and first trial recall scores ($T = -2.01$, $p = .05$) were all significant. All of the remaining beta weights were insignificant ($p > .05$).

Summary.

The series of hierarchical regression analyses were conducted on subjects' performance predictions for each trial. These analyses indicated that the variables influencing performance predictions are different before experiencing the memory task than after task experience has been acquired. First trial predictions were primarily based on subjects' general memory self-efficacy beliefs (MIA-Capacity scores). Second and third trial predictions were based on subjects' previous task experience (i.e., earlier performance predictions and recall scores). This analysis was performed as an attempt to replicate the findings of Hertzog et al. (1990b) and was successful at this replication. Although subjects' utilization of their general MSE did not differ on the basis of task information, a correlational analysis did show a trend for younger females.

Chapter V

DISCUSSION

The present study addressed the relationship between memory self-efficacy, performance predictions, and recall of word lists. The intent was to investigate the interaction between general memory self-efficacy beliefs and specific task information when making global performance predictions. In the preceding chapter, the experimental hypotheses were addressed in a sample of younger and older females. The present chapter discusses the implications of the results, focusing on (a) age differences in subjects' performance predictions, (b) task information effects on performance predictions, (c) age differences in prediction accuracy scores, (d) task information effects on prediction accuracy, and (e) the multidimensionality of the prediction process. Limitations of the present research and suggestions for future research are also presented.

Implications

Age Differences in Global Performance Predictions

Age differences in subjects' performance predictions were apparent in the present investigation. This finding is contrary to the results of Bruce et al. (1982), and Shaw and Craik (1989). However, these results are consistent with those of Hertzog et al. (1990b) and Saylor (1990). The results of the present study do not support the view that older adults predict their performance as if they did not experience any memory performance decline. Both age groups made predictions that were congruent with their performance level. Hertzog et al. (1990b) suggested that such a pattern of age differences in subjects' predictions resembles the age differences found using questionnaire measures of general MSE.

Analyses showed that both age groups monitored their predictions across trials. The young subjects showed the most change in their predictions, whereas the elderly generally maintained their predictions. Since the young subjects' predictions were most discrepant from their recall scores, it is not surprising that they showed the greatest amount of change across trials. The elderly's predictions were closer to their performance thereby resulting in less change in their prediction scores with experience. Changes in global performance predictions exhibited by both younger and older subjects required monitoring ability.

Task Information Condition Effects on Global Performance Predictions

The task information conditions affected subjects' performance predictions across trials. Specifically, subjects who were given accurate task norms decreased their predictions with experience, whereas subjects who were not given such information increased their predictions. This prediction pattern resulted from subjects in the Accurate Norms Condition making higher first trial predictions than subjects in the other two task information conditions. One possible explanation for this pattern might stem from the findings of McDonald-Miszczak et al. (1990). These authors showed that subjects who received a difficult list prior to a moderate list made higher performance predictions than subjects who initially received an easy list. With this result in mind, it is possible that the initial memory task instructions that described the nature of the word lists and the memory task might have given all participants the impression that the subsequent task would be very difficult. Such an impression may result in lower performance predictions like those observed in the No Information and the Self-Generated Norms Conditions. Alternatively, age-graded normative information might have provided subjects in this condition with the impression that the task will not be as difficult as their initial impression of the task. Thus, for subjects in the Accurate Norms Condition, the memory task was appraised as easier relative to their initial impression and resulted in higher overall predictions. The normative information led subjects to raise their task specific self-efficacy much like the subjects in the McDonald-Miszczak et al. (1990) study who received the difficult word list prior to the moderately difficult list. This proposed explanation for the observed differences between subjects' performance prediction on trial 1 is very speculative and requires further examination. Concerning overall monitoring ability, subjects in each task information condition appeared to monitor their performance predictions. Such monitoring resulted in similar performance predictions on the third trial.

Age Differences in Prediction Accuracy

This study showed that the elderly females predicted their memory performance more accurately than the younger females. This finding contradicts the results of most prediction studies (Bruce et al., 1982; Lachman et al., 1979; Lovelace & Marsh, 1985; Murphy et al., 1981; Perlmutter, 1978; Rabinowitz et al., 1981). However, the present finding does concur with the results of Hertzog et al. (1990b) who found that older females predicted their performance more accurately than younger females, whereas males did not differ in their prediction accuracy. McDonald-Miszczak et al. (1990) found that older males and females were more accurate than younger males and females. Finally, Saylor (1990) found that the elderly showed only a slight tendency to overestimate their performance, whereas the younger subjects greatly underestimated their performance. Moreover, Saylor argued that this finding might be an artifact stemming from the elderly simply choosing the mid-point for their performance prediction. Since mean recall performance for the older subjects was close to the mid-point in the present investigation, such an artifact may be a possible explanation for their accurate memory predictions.

Age differences were also found in prediction accuracy scores across trials. It is not surprising that the most change in prediction accuracy scores was observed in the most inaccurate group (i.e., younger subjects). The elderly females maintained their prediction accuracy across trials. These findings suggest that both age groups monitored their prediction accuracy appropriately. However, the potential prediction accuracy artifact outlined above is also applicable to these results. Thus, an interpretation of age equivalent monitoring ability must be made with caution.

Task Information Effects on Prediction Accuracy

Although subjects in the Accurate Norms Condition predicted their performance more accurately than subjects in either the No Information or Self-Generated Information Conditions, the hypothesized Task Information Condition x Trials effect was not significant. It was originally hypothesized that subjects who were given the least task information would show the greatest degree of improvement in their prediction accuracy scores across the three trials, since they would initially predict their performance most inaccurately. The lack of support for this hypothesis may stem from the elderly's prevailing prediction accuracy regardless of the task information condition to which they were assigned.

It was also hypothesized that there would be a significant Age x Task Information Condition effect. Previously, elderly subjects have predicted their performance quite accurately in circumstances similar to those in the No Task Information Condition (Bruce et al., 1982; Lovelace & Marsh, 1985; Murphy et al., 1981). This typical finding was not replicated in the present study. Thus, the hypothesis that a significant Age x Task Information Condition effect would be found was not supported due to the elderly's unexpected prediction accuracy.

The Multidimensionality of the Prediction Process

Hertzog et al. (1990a) outlined a conceptual scheme that asserts that initial performance predictions would be based on a different aspect of metamemory than subsequent performance predictions. Specifically, it was expected that first trial predictions would be based on general memory self-efficacy beliefs, whereas second and third trial predictions would be based on task specific experience (e.g., previous trial prediction and recall scores). A replication of the Hertzog et al. (1990b) finding was conducted for this investigation to examine whether task information would affect subjects' utilization of their general memory beliefs for initial predictions.

The primary expectation that general MSE would significantly determine an individual's first trial prediction was supported. Subjects' general MSE was used to predict their performance before they had directly experienced the task. Subjects' utilization of their general memory self-efficacy beliefs was apparent in all of the task information conditions.

The hypothesis that the effect of general MSE on first trial predictions would depend on the availability of normative task information was generally not supported although a trend was shown. Hertzog et al. (1990a) originally hypothesized that the relationship between general MSE and performance predictions would be stronger when individual differences in subjects' assessments of task difficulty were diminished (e.g., age-graded normative information provided). The general distinction between task information condition and subjects' use of their general MSE was not apparent in the present investigation. However, a correlational trend exhibited by the younger subjects supported the Hertzog et al. (1990a) hypothesis. Clearly, the age-graded task information provided to younger females increased the relationship between their general MSE and their first trial prediction. The elderly subjects did not show such a clear trend. The lack of such a finding for the elderly females in the Accurate Norms Condition may stem from the normative information used. In comparing the age-graded norms and subjects' recall scores, the normative information given to the younger subjects was quite accurate whereas the information given to the elderly subjects was too high. The elderly subjects may not have believed the norms for their age group and therefore were more skeptical in their appraisal of the task. However, the correlational trend exhibited by younger subjects shows partial support for the effect of age-graded normative information increasing the relationship between subjects' general MSE and their first trial prediction by decreasing individual differences in task assessment.

The hypothesis that second trial predictions would be based on previous task experience was supported. This finding replicated the results of Hertzog et al. (1990b) and Saylor (1990). The first trial prediction and recall scores were used to predict second trial performance. The second trial prediction and recall scores were used to predict third trial performance. Thus, all subjects' used task specific experience to make their predictions after such task specific information was available.

Conclusion

In conclusion, the results of the present study support the idea that performance predictions are a process that involves general memory self-efficacy, task-specific self-efficacy, and task appraisal. The results show that both younger and older females seem to monitor their performance and predictions equally with task experience. Overall, prediction accuracy was higher when subjects received accurate task norms, but such information was not found to be differentially useful to a particular age group. The results also show that initial predictions are based on different metamemorial beliefs (i.e., general MSE) than subsequent predictions (i.e., task-specific self-efficacy). Partial support was found for the Hertzog et al. (1990a) hypothesis that providing subjects with age-graded performance means may decrease their individual differences in assessment of the task and thereby increase the relationship between their general MSE and their first trial predictions. A correlational trend demonstrated that accurate normative information did increase the relationship between subjects general MSE and first trial predictions for younger females.

Limitations

The results of the present study should be evaluated in light of limitations of the design. The task information conditions were included in order to create 3 different levels of task information. The No Task Information Condition was included in order to replicate the test situation used in previous prediction studies and provide a control group for any task information effects. The Self-Generated Information Condition was used to provide a condition that prompted subjects' conscious consideration of task difficulty. The Age-Graded Normative Task Information Condition was designed to provide subjects with direct information regarding task difficulty relative to their own age groups. Unfortunately, the Self-Generated Information Condition was not successful in prompting subjects to predict their performance differently (i.e., more accurately) than subjects in the No Information Condition. Simply asking subjects to provide performance estimates for younger and older females prior to providing their own performance estimates did not result in predictions that reflected increased consideration of task difficulty.

The normative information used in this investigation was based on performance levels of a previous study (McDonald-Miszczak et al., 1990). Performance norms were based on this previous study since it contained the same words that were used to comprise the present study's word lists. Although the information used to construct the accurate age-graded norms was the best available, the age ranges in the previous study were wider for the elderly sample. It is likely that performance norms for the elderly sample were inappropriately high due to the inclusion of younger subjects in the previous sample. Such inaccuracy of the performance norm points to the larger problem of using point-estimates as the normative information. Variations in subject samples can easily render previously collected norms inaccurate. For example, a sample with a very high level of education will probably have a higher performance norm than a sample with a lower education level. Thus, it is recommended that performance norms are based on a large sample and that experimenters take precautions to make sure the sample receiving the norms is similar to the sample from which the norms were collected.

The present investigation used a memory task (e.g., 30 word list) that was similar to both the tasks used by Hertzog et al. (1990b) and Saylor (1990). Since the findings of these three investigations generally concur, such support is suspect due to the similarity of the memory tasks used. Saylor (1990) pointed out that elderly subjects may predict and monitor their performance accurately due to an artifact of the memory task. Older subjects might simply use the mid-point (e.g., 15) as their prediction. Since the performance of the older subjects is close to the mid-point in all three studies, the general support shown between these studies may also be an artifact. Thus, a limitation of the present study was that the memory task used may have produced an artifact prediction point similar to both the Hertzog et al. (1990b) and Saylor (1990) tasks.

The effects of the descriptive variables (i.e., health, education, vocabulary, MIA-Capacity scores, and depression) were examined for each dependent variable in order to account for age group differences with more meaningful measures (i.e., individual differences). However, these descriptive variables did not account for the observed age differences. Since the age variable itself is extremely general and does not directly account for group differences, descriptive variables should be carefully chosen and examined. The descriptive variables in the present investigation were included primarily to test that the random assignment and counter-balancing procedures were successful.

The present investigation was cross-sectional and only included females. Due to the design of the study, it is problematic to interpret the differences between younger and older females as necessarily developmental in nature. The results from Hertzog et al. (1990b) suggest that females may exhibit such age differences in prediction accuracy, whereas males do not. Longitudinal data that include both males and females must be collected in order to investigate whether the apparent age differences between females' performance predictions are developmental in nature and whether such changes are gender related (i.e., not apparent in males).

Suggestions for Future Research

Further studies are required to examine the prediction process. A study should be designed and executed to demonstrate the effect of accurate normative information on increasing the relationship between elderly adults' general MSE and first trial performance predictions. Such an investigation is required in order to demonstrate that the prediction process is similar for younger and older females. Accurate normative information must be collected from a large subject sample tested with the same experimental materials that will be used by the experimental subject sample. Moreover, researchers must make every effort to ensure that their subject sample is similar to the sample that provided the normative information.

A primary pitfall of the present study stemmed from the possibility that older subjects' prediction accuracy resulted from an artifact of choosing the mid-point of performance. It is strongly suggested that this possible artifact is examined in a future investigation. The performance levels for the elderly group could be manipulated by presenting them with tasks in which they vary in their baseline performance. If the elderly females still predict their performance more accurately than younger females, then a developmental shift in females' prediction accuracy might be present.

If it is demonstrated that age differences in prediction accuracy are not due to an artifact, then a theoretical framework will be needed to explore such a developmental shift in future research. Such a framework should include possible reasons for such a shift and for the timing of the shift. Longitudinal research would be particularly useful for measuring the developmental shift in prediction accuracy and for examining when such a shift might occur. Cross-sectional research would be relevant for exploring specific hypotheses regarding the components of the shift (i.e., possible gender differences). A theoretical framework of this kind will provide continuity to research projects focusing on developmental changes and potential gender differences in adults' task-specific self-efficacy.

References

Appel, L.F., Cooper, R.G., McCarrell, N., Sims-Knight, J., Yussen, S.R., & Flavell, J.H. (1972). The development of the distinction between perceiving and memorizing. Child Development, 43, 1365-1381.

Bandura, A. (1986). Social foundation of thought and action: A social cognitive theory. Englewood Cliffs, N.J.: Prentice-Hall.

Bennett-Levy, J., & Powell, G.E. (1980). The Subjective Memory Questionnaire (SMQ): An investigation into the self-reporting of "real-life" memory skills. British Journal of Social and Clinical Psychology, 19, 177-188.

Berry, J.M. (1989). Cognitive efficacy across the life span: Introduction to the special series. Developmental Psychology, 25, 683-686.

Bruce, P.R., Coyne, A.C., & Botwinick, J. (1982). Adult age differences in metamemory. Journal of Gerontology, 37, 354-357.

Cavanaugh, J.C. (1989). The importance of memory awareness in memory aging. In L.W. Poon, D.C. Rubin, & B.A. Wilson (Eds.), Everyday cognition in adulthood and late life, (pp. 416-436). New York: Cambridge University Press.

Cavanaugh, J.C., & Perlmutter, M. (1982). Metamemory: A critical examination. Child Development, 53, 11-28.

Chaffin, R. & Herrmann, D.J. (1983). Self reports of memory performance by old and young adults. Human Learning, 2, 17-28.

Devolder, P.A., Brigham, M.C., & Pressley, M. (1990). Memory performance awareness in younger and older adults. Psychology and Aging, 5, 291-303.

Dixon, R. (1989). Questionnaire research on metamemory and aging: Issues of structure and function. In L.W., Poon, D.C. Rubin, & B.A. Wilson (Eds.), Everyday cognition in adulthood and late life, (pp. 394-415). New York: Cambridge University Press.

Ekstrom, R.B., French, J.W., Harman, H.H., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. Princeton, N.J.: Educational Testing Service.

Flavell, J.H. (1971). First discussant's comments: What is memory development the development of? Human Development, 14, 272-278.

Flavell, J.H., Friedrichs, A.G., & Hoyt, J.D. (1970). Developmental changes in memorization processes. Cognitive Psychology, 1, 324-340.

Flavell, J.H., & Wellman, H.M. (1977). Metamemory. In R.V. Kail Jr. & J.W. Hagen (Eds.), Perspectives on the development of memory and cognition, (pp. 3-33). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Gilewski, M.J., & Zelinski, E.M. (1986). Questionnaire assessment of memory complaints. In L.W. Poon (Ed.), Clinical Memory Assessment of Older Adults, (pp. 93-107). Hyattsville, M.D.: American Psychological Association.

Herrmann, D.J. (1982). Know thy memory: The use of questionnaires to assess and study memory. Psychological Bulletin, 92, 434-452.

Herrmann, D.J., & Neisser, U. (1978). An inventory of everyday memory experience. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), Practical aspects of memory, (pp. 35-51). New York: Academic Press.

Hertzog, C., Dixon, R.A., & Hultsch, D.F. (1990a). Metamemory in adulthood: Differentiating knowledge, belief, and behavior. In T.M. Hess (Ed.), Aging and cognition: Knowledge organization and utilization, (pp. 161-212). North-Holland: Elsevier Science Publishers.

Hertzog, C., Dixon, R.A., & Hultsch, D.F. (1990b). Relationships between metamemory, memory predictions, and memory task performance in adults. Psychology and Aging, 5, 215-227.

Hertzog, C., Dixon, R.A., Schulenberg, J.E., & Hultsch, D.F. (1987). On the differentiation of memory beliefs from memory knowledge: The factor structure of the Metamemory in Adulthood scale. Experimental Aging Research, 13, 101-107.

Hertzog, C., Hultsch, D.F., & Dixon, R.A. (1989). Evidence for the convergent validity of two self-report metamemory questionnaires. Developmental Psychology, 25, 687-700.

Hultsch, D.F., Hertzog, C., & Dixon, R.A. (1987). Age differences in metamemory: Resolving the inconsistences. Canadian Journal of Psychology, 41, 193-208.

Hultsch, D.F., Hertzog, C., Dixon, R.A., & Davidson, H. (1988). Memory self knowledge and self-efficacy in the aged. In M.L. Howe & C.J. Brainerd (Eds.), Cognitive development in adulthood: Progress in cognitive development research, (pp. 65-92). New York: Springer.

Kreutzer, M.A., Leonard, C., & Flavell, J.H. (1975). An interview study of children's knowledge about memory. Monographs of the Society of Research in Child Development, 40(1, serial no. 159).

Lachman, J.L., Lachman, R. (1980). Age and the actualization of world knowledge. In L.W. Poon, J.L. Fozard, L.S. Cermak, D. Arenberg, & L.W. Thompson (Eds.), New directions in memory and aging, (pp. 285-311). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Lachman, J.L., Lachman, R., & Thronesbury, C. (1979). Metamemory throughout the adult lifespan. Developmental Psychology, 15, 543-551.

Lovelace, E.A., & Marsh, G.R. (1985). Prediction and evaluation of memory performance by young and old adults. Journal of Gerontology, 40, 192-197.

McDonald-Miszczak, L.C., Hultsch, D.F., & Hunter, M. (1990, October). Effects of task experience on performance, prediction and prediction accuracy. Paper presented at the meeting of the Canadian Association on Gerontology, Victoria, BC.

Murphy, M.D., Sanders, R.E., Gabriesheski, A.S., & Schmitt, F.A. (1981). Metamemory in the aged. Journal of Gerontology, 36, 185-193.

Paivio, A., Yuille, J.C., & Madigan, S.A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology. Monograph Supplement, 76, (1, pt. 2).

Perlmutter, M. (1978). What is memory aging the aging of? Developmental Psychology, 14, 330-345.

Perlmutter, M., Metzger, R., Nezworski, R., & Miller, K., (1981). Spatial and temporal memory in 20- and 60-year-olds. Journal of Gerontology, 36, 59-65.

Rabinowitz, J.C., Ackerman, B.P., Craik, F.I.M., & Hinchley, J.L. (1982). Aging and metamemory: The roles of relatedness and imagery. Journal of Gerontology, 37, 688-695.

Radloff, L.S. (1977, Summer). The CES-D scale: A self-report depression scale for research in the general population. Applied Psychological Measurement, 1, 385-401.

Saylor, L., (1990). Investigating the relationship between metamemory and memory performance predictions. Unpublished master's thesis, Georgia Institute of Technology, Atlanta, G.A.

Sehulster, J.R. (1981a). Structure and pragmatics of a self-theory of memory. Memory & Cognition, 9, 263-276.

Sehulster, J.R. (1981b). Phenomenological correlates of a self-theory of memory. American Journal of Psychology, 94. 527-537.

Sehulster, J.R. (1982). Phenomenological correlates of a self-theory of memory II: Dimensions of memory experience. American Journal of Psychology, 95, 441-454.

Shaw, R.J., & Craik, F.I.M. (1989). Age differences in predictions and performance on a cued recall task. Psychology and Aging, 4, 131-135.

Zelinski, E.M., Gilewski, M.J., & Thompson, L.W. (1980). Do laboratory tests relate to self-assessment of memory ability in the young and old? In L.W. Poon, J.L. Fozard, L.S. Cermak, D. Arenberg, & L.W. Thompson (Eds.), New directions in memory and aging: Proceedings of the George A. Talland Memorial Conference, (pp. 519-544). Hillsdale, N.J.: Erlbaum.

Table 1

Means and Standard Deviations (in parentheses) for Sample Characteristics as a Function of Age

	Age Group			
	20-30 years		65-75 years	
Chronological Age	23.81	(2.81)	70.11	(3.54)
Years of Education	14.92	(1.05)	13.17	(2.42)
Vocabulary	11.36	(2.94)	14.03	(2.95)
MIA-Capacity	58.97	(7.08)	53.25	(6.52)
CES-D	14.14	(8.44)	9.83	(6.59)
Health*	1.47	(0.65)	1.81	(0.67)

* Rated on a 5-point scale (1 = Very Good)

Table 2

Pearson Correlations Between Sample Characteristics, Independent Variables, and the Dependent Variables by Trials

	Age	Info. Cond.	List Order	MIA-Cap.	CES-D	Vocab.	Educ.	Health
Recall								
1st	-.72[a]	-.13	-.03	.43[a]	.20	-.15	.49[a]	-.30[a]
2nd	-.70[a]	-.07	-.09	.52[a]	.16	-.21	.45[a]	-.38[a]
3rd	-.70[a]	-.12	-.15	.46[a]	.17	-.20	.39[a]	-.27*
Prediction								
1st	-.31[a]	-.03	-.01	.23*	.17	-.07	.31[a]	-.25*
2nd	-.69[a]	-.05	-.09	.40[a]	.17	-.18	.52[a]	-.26*
3rd	-.69[a]	-.01	-.10	.47[a]	.12	-.25*	.53[a]	-.30[a]
Prediction Accuracy								
1st	.48[a]	.11	.03	-.25*	-.06	.09	-.24*	.09
2nd	.27*	.05	.04	-.34[a]	-.05	.12	-.07	.28[a]
3rd	.03	.15	.09	.01	-.08	-.07	.21	-.05

* $p < .05$

[a] $p < .01$

Table 3

Mean Recall and Prediction Scores as a Function of Age and Test Trial

	Recall		Prediction	
	M	SD	M	SD
Young				
Trial 1	23.19	3.59	18.86	5.12
Trial 2	23.22	4.91	21.69	3.67
Trial 3	23.39	4.89	21.94	4.65
Old				
Trial 1	14.39	4.77	15.42	4.32
Trial 2	13.58	4.74	14.23	4.13
Trial 3	14.03	4.46	13.23	4.42

Table 4

Mean Prediction Accuracy Scores as a Function of Age, Task Information, and Test Trial

	Task Information					
	No Norm		Accurate Norm		Self-Gen. Norm	
Age Group	M	SD	M	SD	M	SD
Young						
Trial 1	-7.33	5.42	0.42	4.36	-6.08	5.28
Trial 2	-2.83	5.41	0.83	3.46	-2.58	3.34
Trial 3	-2.83	4.47	0.00	4.51	-1.50	4.62
Old						
Trial 1	-0.17	3.35	2.17	3.56	1.80	5.90
Trial 2	0.17	3.93	1.09	2.12	0.50	3.40
Trial 3	-1.83	4.04	-0.33	2.39	-0.55	4.41

Table 5

Pearson Correlations Between Independent Variables, Sample Characteristics, and Mean-Dependent Variables (Averaged Over Trials)

	Age	Info. Cond.	List Order	MIA-Cap	CES-D	Vocab.	Educ.	Health
Mean Recall	-.74[a]	-.11	-.12	.49[a]	.18	-.20	.46[a]	-.33[a]
Mean Prediction	-.64[a]	-.03	-.09	.42[a]	.17	-.19	.52[a]	-.31[a]
Mean Pred. Accuracy	.37[a]	.13	-.07	-.25*	-.08	.06	-.07	.14

* $p < .05$

[a] $p < .01$

Table 6

Hierarchical Regression Analyses for the Effects of the Sample Characteristics and Age on the Mean-Dependent Variables

Dependent Variable	R^2	ΔR^2	F inc	df	p
Mean Recall					
	.42	.21	37.39	1,67	<.01
Mean Prediction					
	.41	.52	14.51	1,64	<.01
Mean Prediction Accuracy					
	.07	.08	6.69	1,65	<.05

Table 7

Hierarchical Regression Analyses for the Effects of Depression, General Memory Self-Efficacy, and Task Experience on Performance Predictions as a Function of Test Trial

Dependent Variable	R^2	ΔR^2	F inc	df	p
Trial 1					
CES-D	.02	-	-	-	-
MIA-Capacity	.08	.06	3.90	1,68	>.05
Trial 2					
CES-D	.03	-	-	-	-
1st Prediction/Recall	.89	.86	245.92	2,66	<.01
MIA-Capacity	.89	.00	0.00	1,65	>.05
Trial 3					
CES-D	.01	-	-	-	-
1st Pred./Recall	.78	.77	117.71	2,66	<.01
2nd Pred./Recall	.90	.12	44.32	2,63	<.01
MIA-Capacity	.90	.00	0.00	1,62	>.05

Table 8

Pearson Correlations Between Global Performance Predictions and MIA-Capacity Scores as a Function of Age, Task Information Condition, and Trial

	Young			Old		
	Tr. 1	Tr. 2	Tr. 3	Tr. 1	Tr. 2	Tr. 3
No Norm Information	0.32	0.22	0.27	-0.22	-0.35	-0.11
Self-Generated Information	-0.08	0.05	0.22	-0.08	0.18	0.18
Accurate Norm Information	0.72	0.36	0.21	-0.35	-0.26	-0.24

Figure 1
Global Performance Predictions as a Function of Age Group and Trial

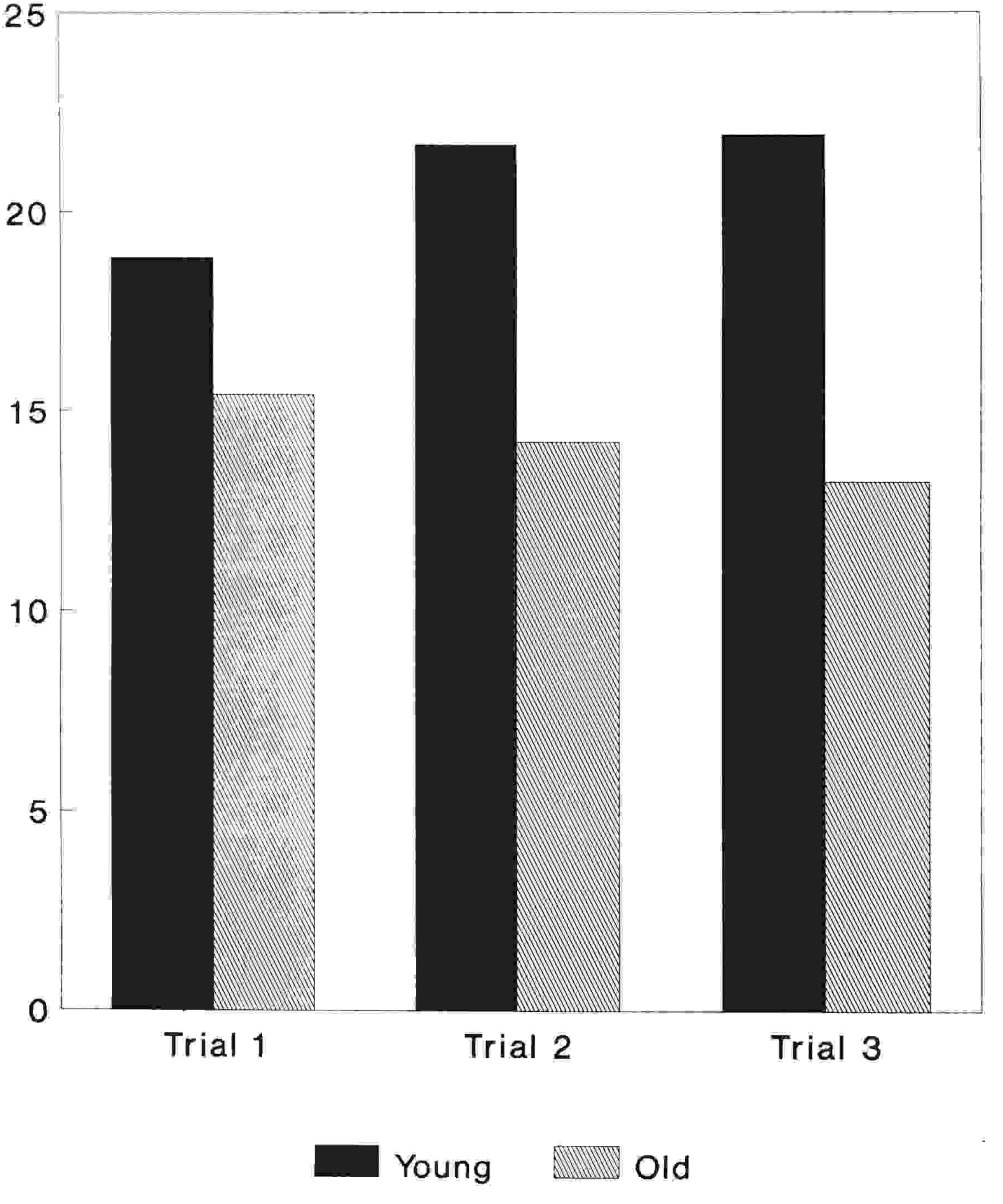


Figure 2

Global Performance Predictions as a Function of Information Condition and Trial

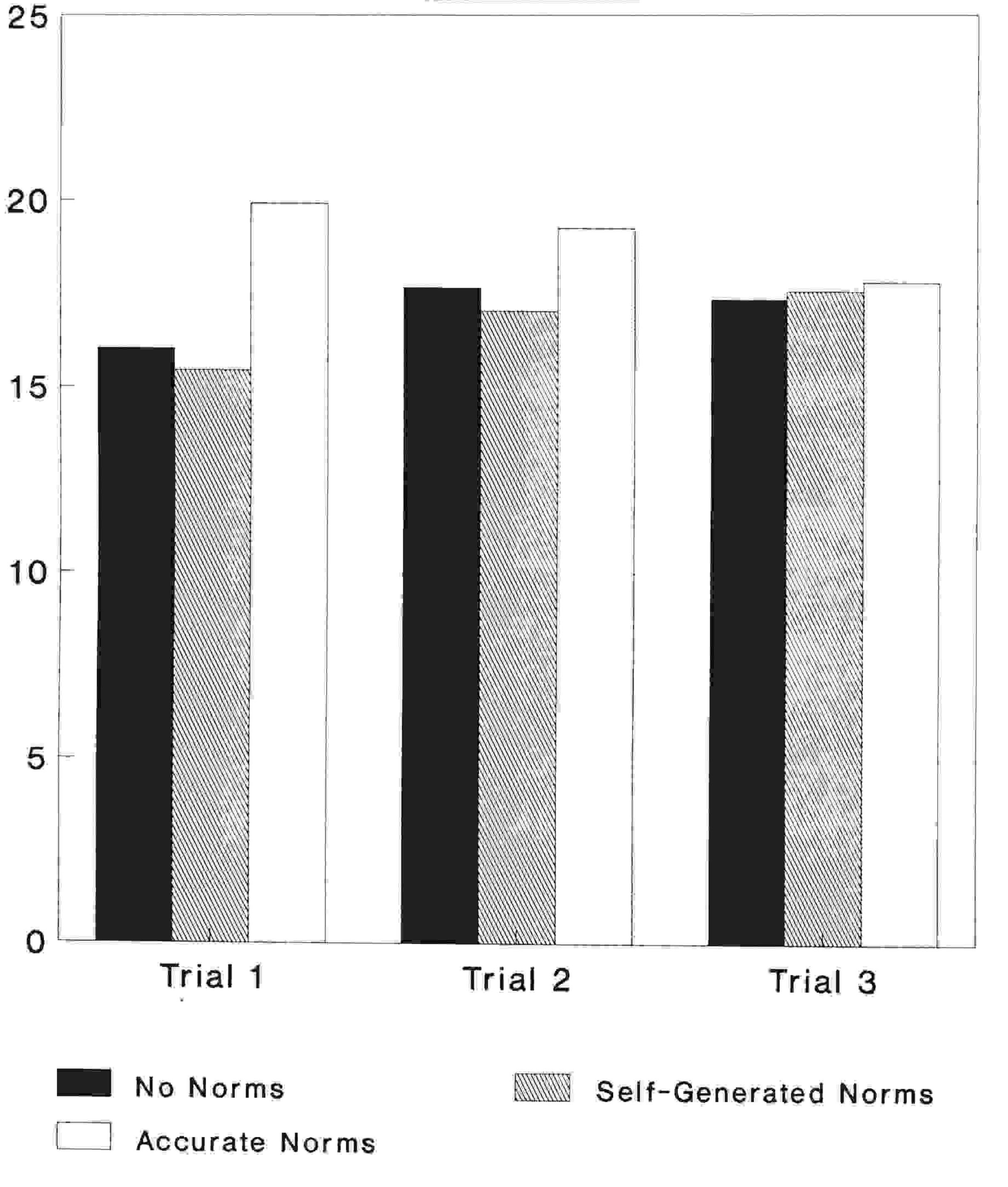
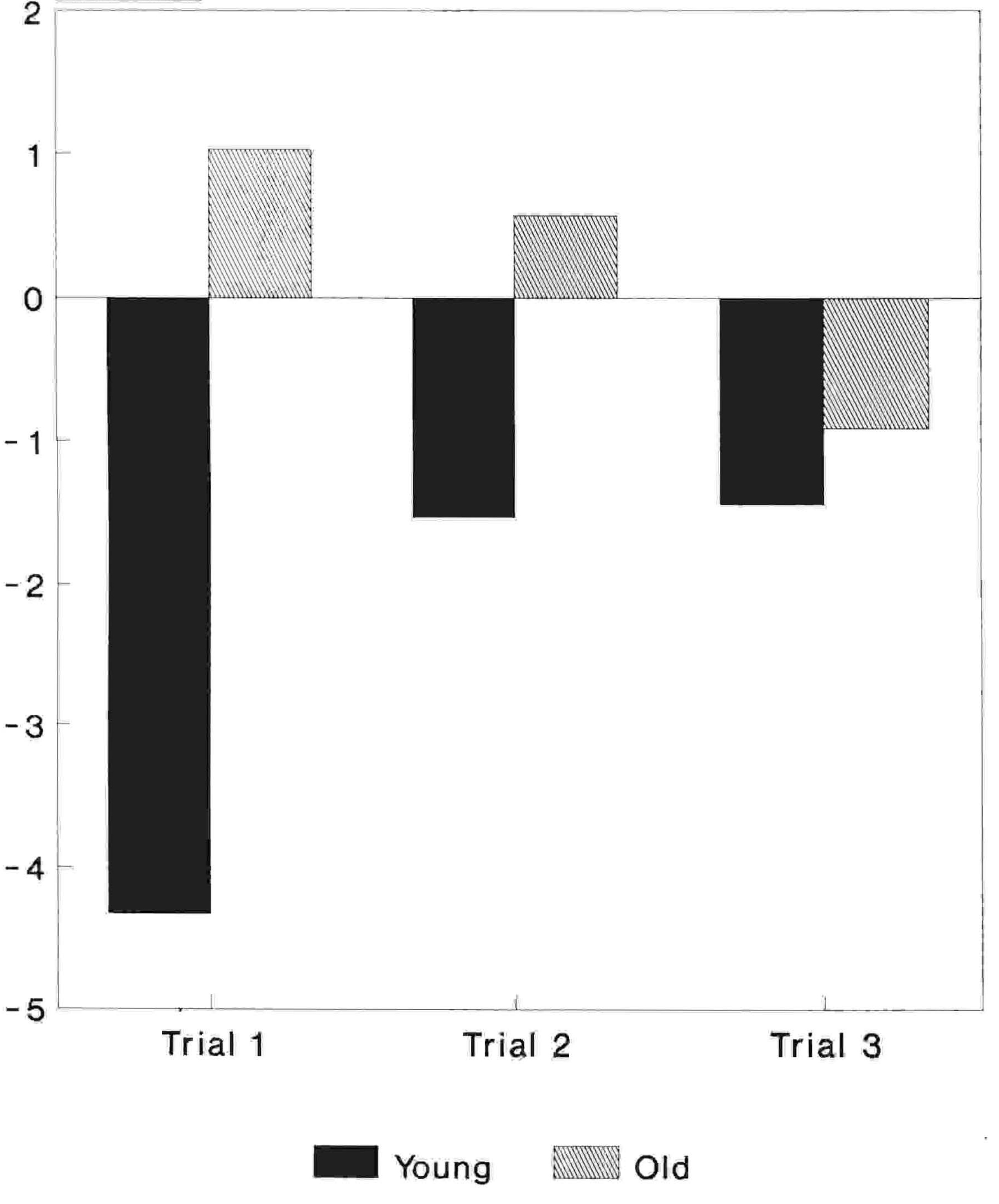


Figure 3

Prediction Accuracy as a Function of Age Group and Trial



VITA

Surname: McDonald-Miszczak Given Names: Leslie Carol

Place of Birth: Calgary, AB Date of Birth: June 14, 1964

Educational Institutions Attended:

University of Victoria	1988 to 1991
University of Alberta	1982 to 1987

Degrees Awarded:

B.A. University of Alberta 1987