1987

Age Differences In Cognitive Monitoring

Margaret Christine Brigham

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ MICROFILMÉE TELLE QUE NOUS L'AVONS REÇUE

AGE DIFFERENCES IN COGNITIVE MONITORING

by

Margaret C. Brigham

Department of Psychology

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
University of Western Ontario
London, Ontario
September 1986

ISBN 0-315-36093-3

ABSTRACT

Metacognitive knowledge regulates and directs cognitive activity. Thus, if metacognitive knowledge deteriorates with age cognitive performance as a whole will suffer. The focus of the present research was on assessing age differences in metacognitive knowledge. The cognitive tasks used were: vocabulary learning with two differentially effective learning strategies, prose recall, vocabulary defining and digit span.

On each task subjects were asked to predict their performance either before study, after study (in the case of the prose task only) or after testing. Following the model of Pressley and associates (e.g., Pressley, Ross, Levin and Ghatala, 1984) the ability to monitor ongoing cognitive activity was assessed by comparing the accuracy of performance predictions made at different points in the study-test cycle. Additional monitoring measures included strategy selection and selection rationale data with the vocabulary learning task and prediction range data with the prose, vocabulary defining and digit span tasks. It was expected that both younger and older adults would provide more accurate performance judgments after testing.

Three groups of subjects were included in this research: older adults (60 years and older), non-student

younger adults (20 to 40 years of age) and first year university students. The two groups of younger adults performed similarily throughout. This finding argues against the hypothesis that age differences are exaggerated when university students form the young adult comparison sample.

Both younger and older adults improved the accuracy of their performance predictions and reduced the size of their prediction ranges from before study to after test. Few age differences in monitoring accuracy were obtained within conditions. There were some age differences in the strategy selection and selection rationale data on the vocabulary learning task which indicated that task familiarity was an important consideration. Older adults were reluctant to select the more potent strategy because it was less familiar. The results are discussed in terms of the importance of ecological validity in age difference research and the appropriateness of a youth-oriented decrement model of aging.

ACKNOWLEDGEMENTS

I would like to thank Dr. Michael Pressley for providing excellent guidance and support throughout this project. I could not have asked for a better supervisor.

Many people, and in particular my parents Jack and Janet Gibson, took an active interest in recruiting volunteers for this study. I would like to thank them for their assistance and enthusiasm. I would also like to thank the volunteers: their co-operation was greatly appreciated.

To my husband Tom, I would like to say **THANKS**.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

Metacognition refers to the individual's awareness of and regulation of his or her own cognitive states and processes (Flavell, 1985). Metacognitions concerning past and ongoing cognitive performances are thought to play a central role in determining the outcome of many cognitive enterprises (Baker and Brown, 1982; Flavell and Wellman, 1977; Pressley, Borkowski and Schneider, in press). Not surprisingly, given the intuitively-obvious relationship between metacognitions and academic performance, much empirical research has focused on metacognitive development in children (e.g., Brown, Bransford, Ferrara and Campione, 1983; Pressley, Borkowski and O'Sullivan, 1985).

Considerably less attention has been directed at metacognitive development across the adult lifespan. There are, however, important reasons for considering this aspect of adulthood. Specifically, if metacognitive processes deteriorate with age, then one's beliefs about one's learning and memory abilities (metacognitions) may come to diverge widely from actual abilities, and the difficulties experienced as a consequence of any age-related decrements in cognitive abilities <u>per se</u> will be compounded. On the other hand, if metacognitive processes remain intact across adulthood, older adults may be able to compensate for minor decrements and, thus, will be able to function independently better, and longer, than might otherwise be expected (Kausler, 1982; Lachman, Lachman and Thronesbery, 1979; Perlmutter, 1978; Sehulster, 1981).

Metacognition

Since Flavell (1971) first used the term "metacognition" to describe the phenomenon of "knowing about knowing", considerable effort has been expended in determining the limits and contents of this domain of knowledge (Yussen, 1985). Flavell and Wellman (1977) distinguished between two aspects of metacognition: metacognitive experiences and metacognitive knowledge. Metacognitive experiences are the conscious feelings and sensations that accompany a cognitive enterprise, the here-and-now reactions to the ongoing cognitive activity (Yussen, 1985). Often, these experiences relate to one's impression of the progress one is making towards a given cognitive goal (Flavell, 1981).

Metacognitive knowledge consists of the individual's store of knowledge concerning factors that influence cognitive activity. Three main classes of interacting factors have been identifed: persons, tasks and strategies. The persons variable encompasses the beliefs that people hold concerning the mnemonic abilities of themselves and others. The task variable pertains to knowledge concerning what factors affect the acquisition and retrieval of information, and includes knowledge concerning when a situation calls for one versus another kind of processing (Flavell, 1985). The strategy variable pertains to knowledge concerning mnemonic activities one can undertake to enhance acquisition and retrieval. A specific item of metacognitive knowledge typically entails an integration of the person, task and strategy variables. For example, "You might believe that you (unlike your brother) should use Strategy A (rather than Strategy B) in Task X (as contrasted to Task Y)" (Flavell, 1979, p. 907).

Baker and Brown (1982) have further clarified the concept of metacognitive knowledge by distinguishing between static metacognitive knowledge and strategic metacognitive knowledge. Static knowledge refers to the declarative aspects of what an individual knows about cognition. For example, the individual might be able to state rules concerning how the person, task and strategy variables affect cognitive performance. Strategic knowledge, on the other hand, "consists of those steps that individuals actually take to regulate and modify the progress of a cognitive activity when it actually occurs" (Yussen, 1985, p. 256). These strategic activities may be specific to the particular cognitive enterprise, or generally applicable to all cognitive enterprises. Generally applicable strategic activities include such behaviours as planning the activity, predicting the outcome, guessing at the solution and monitoring one's progress.

Borkowski, Pressley and associates (e.g., Borkowski, Carr and Pressley, submitted; Pressley, Borkowski and Schneider, in press; Pressley, Borkowski and O'Sullivan, 1985) have integrated both static and strategic components into their model of metacognitive knowledge. According to this model, the sophisticated thinker has available to him or her a variety of cognitive strategies which are useful in pursuing cognitive goals. These strategies include: 1) goal-specific strategies appropriate for use on different cognitive tasks (e.g., rehearsal to aid list memorization, underlining to facilitate reading comprehension); 2) monitoring strategies for use in the regulation of ongoing cognitive activities (e.g., self-testing to see how well one has learned the material); and 3) higher-order sequencing strategies for coordinating the use of goal-specific and monitoring strategies. Higher-order sequencing strategies make complex, cognitive tasks manageable by reducing the amount of information which must be attended to at any one time.

The use of these specific strategies is guided by the metacognitive knowledge the individual possesses concerning when and where in a cognitive task a particular strategy is called for. A crucial aspect of this metacognitive knowledge is the ability to recognize when using a particular strategy has been successful. This motivates future task-appropriate use of the strategy, and contributes to the development of a general tendency to act strategically when warranted by the cognitive task.

The tendency to act strategically forms part of the individual's general metacognitive knowledge. Other aspects of general metacognitive knowledge include the knowledge that success is often enhanced through effort, particularily task-appropriate strategic effort, that specific strategies effective in one learning situation are likely to be effective in other, similar learning situations, and that a good learning environment is conducive to success on cognitive tasks. Thus, a basic premise of this model is that efficient use of task-appropriate learning strategies facilitates cognitive performance and that this usage is motivated and directed by specific and general metacognitive knowledge.

In metacognitive theory declarative aspects of metacognitive knowledge (i.e., rules and beliefs the individual can state concerning memory functioning) interact with procedural aspects of metacognitive knowledge (i.e., the active monitoring of ongoing cognitive activities). For example, the individual's beliefs concerning the efficacy of a particular learning strategy may be altered by how effective the strategy turns out to be in actual use, and the accuracy with which the individual performs the monitoring operations will depend on the knowledge he or she has concerning monitoring strategies. In terms of experimental design, self-report questionaires and performance predictions primarily tap the individual's declarative metacognitive knowledge. On the other hand, measures such as recall readiness judgments and estimates of performance requested during or after the cognitive task has commenced tap on-line process aspects of metacognitive knowledge. In reviewing metacognitive research it is important to keep this distinction in mind.

Age Differences in Metacognition

Older adults typically perform more poorly than younger adults on memory tasks which involve the processing or reorganization of materials (Craik, 1977; Poon, 1985). However, these age differences are ameliorated when the cognitive activities of older adults are altered and directed through such manipulations as strategy training and forced practice (Lachman and Jelilian, 1984; Murphy, Sanders, Gabriesheski and Schmitt, 1981; Rabinowitz, Ackerman, Craik and Hinchley, 1982; Schmitt, Murphy and Sanders, 1981; Scogin, Starandt and Lott, 1985; Yesavage, 1983; Yesavage and Rose, 1984). The observation that older adults do not perform as well as younger adults, despite having the ability to do so when externally guided has lead some researchers to suggest that metacognitive processes may be deficient in older adults (Craik and Rabinowitz, 1984; Perlmutter and Mitchell, 1982). The evidence for and against this hypothesis is examined below.

In one of the earliest studies of age differences in metacognitive functioning, Perlmutter (1978) administered a 60-item metamemory questionaire to 32 younger (20 to 25 years) and 32 older (60 to 65 years) adults. The questionaire assessed knowledge of naturalistic memory situations. Older adults reported more memory problems and expected greater memory deterioration with age than did younger adults. These memory beliefs were tapped by items such as "How often do you forget names?" and "Do you think your memory will get worse as you get older?". No age differences were found in beliefs about factors that influence ease of remembering (e.g., "Do you find it easier to remember organized things than unorganized things?"), or in the mnemonic strategies subjects reported using (e.g., "How often do you write reminder notes?"). Both younger and older adults reported that they used many more external than internal memory aids. Thus, both younger and older adults expressed beliefs about their own memories that reflect general cognitive psychology findings. However, Perlmutter did not compare the questionaire results to any measures of actual memory performance. Without this comparison it is impossible to know whether beliefs about memory performance are representative of actual memory abilities (Herrmann, 1982).

In the second phase of this study, Perlmutter requested her subjects to predict their free recall and recognition scores for two 24-item word lists, and their recognition scores for 24 general information statements. The statements pertained to historic events occurring between 1890 and 1969. Note that this task may have been biased in favour of the older adults, insofar as they would have lived through more of the historical events than had their younger counterparts. These predictions were made before the test, but after studying the actual stimuli. Thus, this task allowed for subjects to self-test, and to the extent that they did so, measured procedural rather than only declarative aspects of metacognitive knowledge.

Performance by the younger adults was superior on the word tasks, while performance by the older adults was superior on the fact task. No age differences were found in the accuracy of predictions. However, both younger and older adults tended to overpredict their recall on the word tasks. Perlmutter concluded that "the knowledge about memory that subjects in both age groups had, as indicated by their responses on the memory questionaire, was comparable", and "age differences in monitoring skills do not contribute to age differences in adult memory" (p. 342).

Similar conclusions were reached by Lachman et al. (1979) in their assessment of age differences in the ability to monitor on-going cognitive activity. In this study, 12 younger (19 to 22 years), 12 middle-aged (44 to 53 years) and 12 older (65 to 74 years) adults were presented with 95 general information questions. For each question, the subjects either gave an answer or indicated that they did not know the answer. Following this test, questions for which the answer had not been known were presented again and subjects were requested to make a 'feeling-of-knowing' judgment based on a four-point scale. The accuracy of these judgments was assessed by presenting the questions again, along with four possible answers for each. Subjects selected the answer they considered most appropriate. No age differences were obtained in performance, or in the accuracy of subject's feeling-of-knowing judgments.

Other researchers have found evidence for age differences in metacognitive functioning, however. Bruce et al. (1982) showed 24 adults in each of three age groups (18 to 31; 60 to 69; 70 to 79 years) four four-item examples of the following word-pair types: high imagery/high frequency; high imagery/low frequency; low imagery/high frequency; and low imagery/low frequency. Subjects were asked to predict how many items of each type they would recall if given unlimited study time and 20-pair lists for each word-pair type. Predictions were followed by self-paced study and a recall test for each list.

Actual recall levels decreased as age increased. However, predicted recall levels were equivalent among groups. In general, older adults overpredicted their recall, while younger adults underpredicted their recall or were accurate. In addition to these differences in declarative knowledge, differences in the ability to monitor performance were implied: there were no differences in self-paced study time, despite the differences in recall. Thus, the recall readiness judgments of younger adults can be considered more accurate than those of older adults.

Murphy et al. (Experiment 1, 1981) assessed age differences in serial picture span prediction accuracy and in recall readiness judgments. Twenty-three younger (17 to 29 years) and 23 older (60 to 80 years) adults were shown sets of line drawings beginning with a two-item series. On each successive trial, a new set of drawings incremented by one was presented and the subjects predicted whether they could recall that many items in order. When maximum serial span was indicated, the accuracy of this prediction was assessed by presenting new stimuli for recall in the same format as had been used in the prediction task. In the second phase of the study, recall readiness monitoring was assessed by giving subjects unlimited time to study subspan, span and superspan sets of line drawings, with a criterion of perfect recall. Recall was then tested.

Actual serial span was greater for younger than for older adults. Both groups made accurate serial span predictions, however the direction of error differed for the two age levels: older adults tended to overpredict their performance, while younger adults tended to make underpredictions. In the second phase of the study, older adults monitored their recall readiness less accurately than did younger adults: they spent less time studying at each level of task difficulty, and increased their study time less in response to increased task difficulty.

Lachman and Jelilian (1984) included a prediction measure in their study of age differences in attributions for intellectual performance. In their study, 37 younger (18 to 22 years) and 48 older (64 to 91) adults completed a letter series task and a vocabulary test. On the basis of sample items, subjects were asked to predict how well they thought they would do on the actual tests.

On the letter series test, younger adults performed better than older adults. Younger adults predicted their performance more accurately than did their older counterparts: older adults tended to overpredict their performance, while younger adults made underpredictions. In contrast, on the vocabulary task, older adults outperformed younger adults, and predicted their performance more accurately. Older adults made accurate predictions, while younger adults tended to underpredict their performance.

This study included a monitoring component in addition to the prediction measure. Subjects provided after-test estimates for their performance. The sequence of before-study recall predictions, testing and after-test recall predictions was repeated twice. Both younger and older adults became more accurate in their self-assessments by the second set of after-test predictions. Monitoring accuracy differences on the vocabulary task disappeared, while younger adults were still more accurate on the letter series task.

Devolder and Pressley (in preparation) presented 48 younger (19 to 40 years) and 48 older (58 to 84 years) adults with two 30-item word lists. Subjects were asked to predict their free recall scores for one list, and their recognition scores for the other. Younger adults outperformed older adults. Older adults tended to overpredict their recall scores, while younger adults made underpredictions. Subjects at both age levels underpredicted their recognition scores. An after-test postdiction component was included in this study. Monitoring performance was similar between younger and older adults after testing.

Lovelace and Marsh (1984) assessed age differences in performance monitoring by requesting twenty younger (17 to 24 years) and 20 older (60 to 75) adults to self-initiate presentation of 60 unrelated paired associates. Upon seeing each stimulus pair, subjects used a five-point scale to indicate how sure they were that they would be able to match those two words correctly on the test. Following study, subjects completed a matching task, in which they were instructed to match all pairs and to rate the accuracy of each match.

Younger adults made more correct pairings. No age differences were found in the accuracy of monitoring judgments made on an item-by-item basis, as indexed by appropriate use of the rating scale, although overall younger adults used the rating scale more accurately than did older adults. (Older adults overpredicted their performance). Note, however, that the derived measure of monitoring accuracy used in this study does not correspond to the solicited monitoring judgments used in the studies reported above.

Rabinowitz et al. (1982) likewise compared monitoring judgments made on an item-by-item basis. They asked 24 younger (17 to 24 years) and 24 older (61 to 71 years) adults to study 50 cue-target word pairs of either high, medium or low relatedness, for 10 seconds each. Subjects were instructed to learn the pairings in one condition and to form an interactive image in the other. Following study of each stimulus item, subjects were asked to indicate using a 10-point scale how sure they were that they would remember the target when given the cue. Cued recall was tested following study.

Younger adults showed superior recall in both conditions, however, no age differences were obtained in monitoring accuracy as measured by use of the rating scale. Both younger and older adults appeared to be sensitive to the beneficial effects on recall of relatedness, but unaware of the benefits of imagery.

Several interesting points concerning age differences in metacognitive knowledge are revealed by this review. First, age differences have been found on tasks with low ecological validity (paired associates: Bruce et al., 1982, Lovelace and Marsh, 1984; letter series: Lachman and Jellilan, (1984); picture series: Murphy et al., 1982), but not on more ecologically-valid tasks (general facts: Lachman et al., 1979, Perlmutter, 1978; vocabulary: Lachman and Jellilan, 1984). Second, age differences have been found in the accuracy of overall recall predictions (e.g., Bruce et al., 1982; Murphy et al., 1981), but not in the accuracy of item-by-item predictions (Lovelace and Marsh, 1984; Rabinowitz et al., 1982). Third, age differences have been found for some measures of metacognitive knowledge, but not for others. Specifically, age differences have been found in performance predictions made before study and in recall readiness judgments (e.g., Bruce et al., 1982; Murphy et al., 1981) but not in performance predictions made after testing (Lachman and Jellilan, 1984; Devolder and Pressley, in preparation). And fourth, when age differences are found, they generally take the form of overprediction on the part of the older adults, while younger adults tend to underpredict (e.g., Bruce et al., 1982; Murphy et al., 1981; Lachman and

Jellilan, 1984).

In summary, the limited amount of research available suggests that age differences are neither dramatic nor pervasive. Rather, age differences have been obtained under specific circumstances only; in particular, when an overall prediction of performance was requested before testing on a laboratory-style task of low ecological validity. Moreover, when age differences have been found, they have typically been represented by differences between younger and older adults in the direction of prediction error, rather than by differences in predictive or monitoring accuracy.

An explanation for this pattern of age differences is suggested by Lovelace and Marsh (1984). These authors suggest that older adults do not so much overpredict their abilities, as underpredict task difficulty, due to their lack of recent experience with cognitive tasks similar to those encountered in typical laboratory experiments. Younger adults, on the other hand, have more recently been in school; hence they do not suffer from this same handicap but may in fact show the reverse effect: they expect a "test" to be more difficult than is usually the case in psychological experiments. Thus, it has been argued (e.g., Poon, 1985; Poon, Kraus and Bowles, 1984) that age differences in metacognitive performance are particularily likely when the performance of older adults is compared to that of university students, as is commonly done in the age difference literature (e.g., Lachman and Jelilian, 1984; Lovelace and Marsh, 1984; Murphy et al., 1982; Rabinowitz et al., 1981).

Thus, age differences in metacognition are not found generally, but rather are found only under specific experimental conditions. Moreover, when age differences are found they can be interpreted readily within a framework which acknowledges the experiential differences between young and old. These observations give rise to the hypothesis that metacognition, like other aspects of cognitive and social functioning (e.g., Labouvie-Vief, 1985; Yussen, 1985) may be structured differently in younger and older adults. That is, most researchers have used a unitary operational definition of metacognitive knowledge and then interpreted any age differences in performance as evidence for deficient metacognitive processing on the part of older adults (Dixon and Hultsch, 1983a, 1983b). It may be, however, that these age differences are found, not because of an age-related decrement in metacognition, but because the nature of the multidimensional relationship between metacognition and cognitive behaviour changes with age.

Dixon and Hultsch addressed this possibility in their development and validation of the Metamemory in Adulthood (MIA) instrument (Dixon and Hultsch, 1984). The MIA was designed (Dixon and Hultsch, 1983a) to measure familiar or theoretically-relevant metamemory activities and beliefs. One hundred and twenty questions, divided among eight dimensions of metamemory are included. Subjects indicate their agreement with each question or the frequency of occurrence of the reported event using a five point scale. The eight metamemory dimensions are as follows: 1) Use of memory strategies (Strategy): this category assesses use of strategies, not knowledge concerning them. A sample question is "Do you make mental images or pictures to help you remember?" 2) Knowledge of memory tasks (Task): this category taps knowledge of basic memory processes (e.g., "For most people it is easier to remember things in which they are more interested than things in which they are less interested"). 3) Knowledge of own memory capacities (Capacity): this category refers to how well an individual expects to perform in a given memory situation (e.g., "I am good at remembering the content of news articles and broadcasts"). 4) Attitudes towards own memory - perception of change (Change): this category focuses on how memory is expected to change with age (e.g., "My memory will get better as I get older"). 5) Activities supportive of memory (Activity): whether the individual regularily engages in cognitive activities which would support memory is tapped by this category (e.g., "How often do you read fiction books?'). 6) Memory and state anxiety (Anxiety): this category assesses both knowledge of how emotional state affects cognitive performance and knowledge of how cognitive demands affect emotional state (e.g., "I have difficulty remembering things when I am anxious"). 7) Memory and achievement motivation (Achievement): the importance to the individual of having a good memory is assessed by this category (e.g., "It is important to me to have a good memory"). 8) Locus of control in memory abilities (Locus): this category petains to the individual's sense of control over remembering abilities (e.g., "No matter how hard a person works on his memory, it cannot be improved very much").

The relationship between performance on the MIA and memory for text was examined using three separate samples of adults: Sample 1 (younger 18 to 31; older 60 to 81 years); Sample 2 (younger 21 to 39; middle-aged 39 to 58; older 60 to 84 years); and Sample 3 (younger 21 to 39; middle-aged 39 to 58; older 60 to 74 years). Each sample completed the MIA instrument and read a series of text articles. Sample 1 read five newspaper articles about a recent news event. Sample 2 read six articles adapted from textbooks on health-related topics. Sample 3 read four personal narrative texts adapted from magazine articles. Reading was self-paced, and self-paced written recall was tested after each text had been read. Verbatim recall was not required., Samples 1 and 2 completed both an immediate recall test and a delayed recall test administered one week after the original session. Sample 3 completed an immediate recall test and two delayed recall tests, one administered one week, and the other four weeks, after the original session.

Some differences in performance were found between the samples and these were attributed to as-yet unspecified differences in text structure. The most important finding, however, was that metamemory and memory behaviour were related somewhat differently for the different age levels:

> for younger adults what is known about retrieval strategies and physical reminders (Strategy), as well as what is known and predicted about performance on given tasks (Capacity) and known about general memory processes (Task), best predicts memory performance. For middle-aged adults what is known about retrieval and reminder strategies (Strategy), what is known about general memory processes (Task), what is known about one's own specific performance abilities (Capacity) and the level of motivation to achieve in memory performance (Achievement) predicts memory for text performance. For older adults the picture is somewhat different: What is known about general memory processes (Task), the level of motivation to achieve in memory performance (Achievement) and the sense of internal-external locus of control of

> memory functioning (Locus) predict memory performance. (p. 694).

Dixon and Hultsch (1983b) concluded that "Whereas memory performance by young adults appears to be best accounted for by knowledge components of metamemory, the performance by older adults is considerably more related to affective dimensions" (p. 694).

These findings suggest that the memory performance of older adults "may be more susceptible than that of younger adults to noncognitive influences" (Dixon and Hultsch, 1983b, p. 694). Borkowski et al. (submitted) argue that non-cognitive factors such as self-esteem, an internal locus of control, and the tendency to attribute success to effort derive from and motivate efficient metacognitive functioning. These non-cognitive factors are more situation-specific and variable than are general beliefs about cognitive ability. Taken together, these observations suggest that when age differences in predictive and monitoring accuracy are found, both the cognitive and non-cognitive aspects of metacognitive knowledge may be contributing factors.

<u>The Present Experiments</u>

In the present study several of the issues raised by the results of the very limited amount of previous research available were addressed. First, the primary concern of the study was whether or not age differences would be found in the ability to monitor ongoing cognitive performance. Previous research indicates that monitoring skills are relatively stable across the adult lifespan.

Second, whether or not age differences would be found in the direction of monitoring error was assessed. Previous research indicates a consistent tendency on the part of older adults to overpredict their performance levels, while younger adults are more likely to underpredict.

Third, the role of ecological-validity in determining whether or not age differences are found was assessed by using discrete-item, laboratory-style materials in Experiment 1 and more real-life discrete-item materials, along with prose materials in Experiment 2. Previous research suggests that age differences are more likely with materials and tasks of low ecological validity.

Fourth, the view that any interpretation of age differences in metacognitive functioning should take into account the non-cognitive components of metacognitive knowledge as well as the cognitive components (Dixon and Hultsch, 1983b) was acknowledged by conducting a semi-structured interview in Experiment 2 designed to assess subject's beliefs about their memory performance.

And fifth, the hypothesis that age differences are more likely to be found when university students form the young adult comparison sample (Poon, 1985; Poon et al., 1980) was addressed by including three groups of subjects: first year university students, non-student community-dwelling younger adults, and community-dwelling older adults. If the students and non-student younger adults performed similarily, the argument that age differences are exaggerated when "test-sophisticated" students form the comparison sample would be diminished. On the other hand, if the students and non-student younger adults did not perform similarily, the argument that using university students as a comparison sample in age difference research introduces a bias would gain support.

Experiment 1

Introduction

One aspect of memory monitoring is the ability to judge how well materials have been learned. It is on the basis of such judgments that an individual decides, for example, whether the learning strategy he or she has been using is appropriate for the task at hand. Shaughnessy (1981, Experiment 2) investigated this aspect of metamemory. University undergraduates were asked to learn four 10-item paired-associate word lists. Two of the lists were studied by reciting the word pairs aloud (maintenance rehearsal), and the other two were studied by forming an interactive visual image using the members of each pair (elaborative rehearsal). In addition, the likelihood that a given word pair would be recalled correctly was rated immediately after study of that pair. The most important result was that although recall with the elaborative rehearsal strategy was superior, subjects expected performance with the two strategies to be equivalent.

Shaughnessy (1981, Experiment 3) used a paired comparison procedure to conduct a more stringent test of this failure to discriminate the relative efficacy of two learning strategies. The paired-associate word lists were studied as in Experiment 2. However, rather than rate the memorability of each individual item, after study of all the items subjects were asked to rate the relative memorability of sets of paired-associate items, where one item in each set had been studied using elaborative rehearsal and the other item had been studied using maintenance rehearsal. Subjects again failed to detect the superior efficacy of the elaboration strategy.

Thus, subjects in Shaughnessy's (1981, Experiments 2 and 3) experiments did not make accurate strategy efficacy judgments while using the strategies to study. In light of this finding, Pressley, Levin and Ghatala (1984) reasoned that testing, rather than study alone, may be necessary in order to develop knowledge concerning relative strategy efficacy. Accordingly, they compared the efficiency with which strategy efficacy judgments were made at different points in the study-test cycle.

The two strategies they used were maintenance rehearsal and an associative elaboration method of learning, known as the 'keyword' method. The keyword method requires the learner to encode the material-to-be-learned, for example, unfamiliar vocabulary words, using keywords, which are proxies that physically ressemble the to-be-learned items and are more familiar than the to-be-acquired content. For example, a good keyword for the word handsel would be hand, because of the acoustic similarity between it and the vocabulary word, and because hand is a word well known to most native English speakers. An elaboration can then be constructed between the keyword and the definition of the vocabulary item. For handsel, which is an old-English word meaning payment, the verbal elaboration might be the sentence "She carried the payment in her hand". Many types of lexical items are learned more effectively by children and young adults with the keyword method than with a variety of alternative techniques (Pressley, Levin and Delaney, 1982).

In a series of five experiments, Pressley, Levin and Ghatala (1984) asked university undergraduates and 10- to 13-year old grade school students to chose either the maintenance rehearsal strategy or the keyword strategy for use in learning English synonyms for unfamiliar Spanish and Latin vocabulary words. The university students made their strategy selections either before study, after study but before testing, or after testing. Recall with the keyword strategy was superior. However, a pronounced preference for this strategy was evidenced only after the subjects had experienced both study and testing with the two strategies.

The grade school students made their strategy selections either before study, after study and testing, or after study, testing, and the provision of explicit feedback on their performance with the two strategies. The keyword strategy lead to superior recall. However, subjects demonstrated a pronounced preference for the more effective keyword strategy only after study, testing <u>and</u> the provision of feedback as to how many items they had recalled with each method.

Pressley, Ross, Levin and Ghatala (1984) hypothesized that the children in Pressley, Levin and Ghatala's (1984) study might have had information in metamemory that they failed to deploy efficiently. They reasoned that if that were so, providing specific performance feedback might not be necessary. Rather, prompting children to compare their relative performance with two strategies before making their strategy selections should activate this knowledge and lead to appropriate strategy selection. Pressley, Ross, Levin and Ghatala (1984) exposed 10- to 13-year old children to two study strategies for learning a 22-item list of unfamiliar old-English vocabulary words: a semantic context strategy and the more potent keyword strategy. The semantic context strategy involved constructing sentences with the new words used correctly.

Subjects were requested to choose between the two strategies either before study (control condition), after study and testing (practice condition), after study, testing and prompting to "think back" and decide which method had enabled them to learn more items (prompt condition), or after study, testing, and the provision of explicit feedback on their performance with the two strategies (feedback condition).

Above chance selection of the keyword strategy occurred in only the prompt and feedback conditions. However, when questioned after testing, over 75 percent of the subjects in the practice condition were aware that recall with the keyword method was superior. This finding is consistent with the hypothesis that children can possess an item of metamemory, but not deploy it efficiently.

Thus, it appears that even mature cognizers (university students) monitor their memory performance inefficiently on the basis of study alone. However, these individuals generally are well aware of their performance levels after taking a test on the materials in question. For individuals with less well-developed metacognitive skills (grade school children), even the experience of testing is inadequate to produce learners who exhibit efficient monitoring skills.

The studies by Pressley and associates (Pressley, Levin and Ghatala, 1984; Pressley, Ross, Levin and Ghatala, 1984) indicate that the study-test experimental paradigm is sensitive to differences in monitoring skill. Accordingly, in Experiment 1 a study-test procedure similar to that developed by Pressley, Ross, Levin and Ghatala (1984) was used to compare the monitoring performance of younger adults to that of older adults. Subjects were asked to indicate their preference for either the keyword strategy or a context strategy, and to estimate their recall performance with each, either before study and/or after study and testing with the 22-item list of unfamiliar vocabulary words used by Pressley, Ross, Levin and Ghatala (1984). Given the previous research findings (e.g., Bruce et al., 1982; Lovelace and Marsh, 1984; Lachman and Jelilian, 1984; Murphy et al., 1981), it was expected that subjects at all age levels would improve the accuracy of their monitoring judgments across the study-test cycle. However, performance between age levels was not expected to be identical; rather, older adults were expected to show the typical tendency towards overprediction, and to be hampered by the low ecological validity of the stimuli. The limited amount of research available argued against formulating these hypotheses more specifically.

Three groups of subjects were included in Experiment 1: older adults, non-student community-dwelling younger adults, and university students. The extent to which performance was similar or dissimilar between the two groups of young adults was expected to shed light on the question of whether university students are an adequate comparison sample in age difference research.

Method

<u>Subjects.</u> Subjects were 40 community-dwelling older adults between 60 and 88 years of age (<u>M</u>=72.6, <u>S.D.</u>=7.4), 40 community-dwelling, non-student younger adults between 24 and 39 years of age (<u>M</u>=33.6, <u>S.D.</u>=4.5) and 40 university students between 18 and 38 years of age (<u>M</u>=19.7, <u>S.D.</u>=1.5). All subjects were tested within a one year period. The university students were enrolled in introductory psychology courses at the University of Western Ontario and received course credit for participating. The community-dwelling older adults and community-dwelling non-student younger adults were residents of Stratford, Ontario and the surrounding rural area. The majority were recruited through personal contacts and the remainder through clubs and organizations.

Design. Six independent groups were formed by the factorial combination of two variables, age (student, non-student younger adult, older adult) and condition (before-study, after-test). Twenty subjects from each age group were assigned to each condition. Within each condition there were two orders of instructional presentation (keyword followed by context, context followed by keyword), with half of the participants in each condition receiving the keyword-context order and the other half receiving the context-keyword order. Within each form of instructional presentation there were two versions of the stimuli (List A and List B). Each version was administered to half of the subjects receiving each form of instructional presentation. The design of Experiment 1 is summarized in Table 1.

Materials. The two word lists used in this study are presented in Appendix A. The first word list consisted of the 22 old-English nouns used by Pressley, Ross, Levin and Ghatala (1984). Two versions of the first word list were prepared, with the order of stimuli the same on each version. On one version of the list (List A), the old-English noun, a synonym from current usage and a keyword were presented for even items (keyword presentation) and only the old-English noun and its synonym were presented for odd items (context

Table 1

Design of Experiment 1

	Condition (n=20)			
Order of Presentation	Keyword-Context (n=10)		Context-Keyword (n=10)	
Version of Word List 1	List A (n=5)	List B (n=5)	List A (n=5)	List B (n=5)

Notes: 1. 2 conditions, 3 age levels: N = 120.

presentation). On the other version of the list (List B), the keyword presentation was used for odd items and the context presentation was used for even items. The second word list consisted of 9 old-English nouns. On one version of the second list, the keyword presentation was used for all nine items, while on the other version, the context presentation was used for all items.

The study items were presented both auditorily and visually at a rate of 25 seconds per item. This rate was used by Pressley, Ross, Levin and Ghatala (1984). Each study item was printed on a separate 20.3 x 12.7 cm flashcard in uppercase letters approximately 2.0 cm high. Keyword presentation flashcards included the old English noun, a keyword and the synonym. The accompanying audiotape presentation was "(word) sounds like (keyword) and means (synonym). Make up a sentence that makes sense that has (keyword) and (synonym) in it." Context presentation flashcards included the old English word and the synonym. The audiotape presentation was "(word) means (meaning). Make up a sentence with (word) used correctly". Corresponding to each study flashcard was a test flashcard, on which the old-English noun was printed alone. Four sample stimuli, procedurally identical to the task stimuli were used to instruct the subjects. The sample stimuli are presented in Appendix B. The sentence rather than the imagery version of the keyword method was used because of the operational similarity between it and the context method and because it can be executed overtly. Thus, it was possible to ensure that the subjects were actually applying the technique.

Three five-point rating scales were used. Possible responses ranged from "I just took a quess" (1) to "absolutely certain" (5). The three scales differed only in that one was worded to refer to confidence in strategy choice, another to confidence in before-study performance predictions and the third to confidence in after-test performance predictions. These scales are presented in Appendix C. The vocabulary subtest of the Weschler Adult Intelligence Scale (WAIS) (Weschler, 1955) and a subject data questionaire were also used. These are presented in Appendices D and E, respectively.

Procedure. Table 2 presents a summary of the procedure followed in Experiment 1. Participants were tested individually. Students were seen in an office at the University of Western Ontario. Non-student younger adults and older adults were given a choice of testing locations. Two older and two younger adults were seen at offices in Stratford, six younger adults were tested at the experimenter's home, and the remaining subjects were tested in their homes.

Table 2

Summary of Procedure for Experiment 1

Condition	
Before Study	After Test
Strategy instruction	Strategy instruction
Strategy monitoring	
Word List 1 study	Word List 1 study
Word List 1 test	Word List 1 test
Strategy monitoring	Strategy monitoring
Word List 2 study	Word List 2 study
Word List 2 test	Word List 2 test
WAIS Vocabulary test	WAIS vocabulary test
Subject data	Subject data
Debriefing	Debriefing

All subjects were informed that their task was to learn some new vocabulary words. Half of the subjects in each condition had the keyword method explained first, followed by the context method; the other subjects learned about the context method first and the keyword method second. The directions in the Keyword-context order were as follows:

> Today I am going to ask you to learn some vocabulary words and I want to tell you about two different ways of doing it. One way to learn the words is to do the following: Consider the word handsel which means payment. The way to remember that handsel means payment is first to note that part of handsel sounds like the English word hand. Then, make up a sentence that makes sense that has hand and payment in it. Can you do that? What is your sentence? Good. Another example is "He held the payment in his hand". Any sentence is good just so long as it is about a hand and a payment doing something together. (Once subjects had produced an adequate sentence, the procedure was repeated for jarvey means driver).

There is another way to learn vocabulary words. It is to do the following: Consider the word <u>casern</u> which means barracks. To remember this word you could make up a sentence that makes sense and that contains the word <u>casern</u>. Can you do that? What is your sentence? Good. Another example is "The troops were confined to their <u>casern</u> because of the bad weather outside". Any sentence is good as long as <u>casern</u> is used correctly. (Once subjects had produced an adequate sentence, the procedure was repeated for <u>piggin</u> means bucket).

Subjects in the before-study condition were asked to choose which of the two strategies they would prefer to use. The order of mention of the strategies corresponded to the order of mention during strategy instruction. Subjects were told:

Now I am going to present a list of vocabulary words for you to learn. What you must remember for each word is the one word meaning. Suppose that you can use either one of the two methods of learning we have discussed. You can either make up a sentence that has a sound-alike word and the meaning in it, or you can make up a sentence that uses

> the word correctly. Which would you prefer to do? Make up a sentence that uses the sound-alike word and the meaning, or make up a sentence that uses the word correctly?

After they had made their strategy choices, subjects were asked to rank themselves on a five-point scale with respect to how sure they were that they had made the right strategy choice. They were asked why they chose that strategy, and which strategy they thought would produce better learning of the meanings. Then they were asked to predict their recall performance, as follows:

> Suppose I was to show you a list of 22 unfamiliar words and I asked you to learn half of the words (11) by using a sound-alike word and the meaning in a sentence, and the other half (also 11) by making up a sentence with the word used correctly in it. How many of the meanings do you think that you would get right on a subsequent memory test? Of that number, how many of the meanings do you think you would have learned by using the sound-alike word and the meaning in a sentence? How many of the meanings do you think you would have learned by making up a sentence with the word used correctly?

Subjects then ranked themselves on a 5-point scale with respect to how certain they were that they would recall more items with the strategy with which they expected to do better.

After subjects in the before-study condition had been instructed in strategy use and had made their monitoring judgments, and immediately after strategy instruction for subjects in the after-test condition, subjects in both conditions were told that they would be presented with a list of vocabulary words to learn, as follows:

> Now I am going to show you a lot of words to learn. For half of them I want you to make up a sentence with the sound-alike word and the meaning doing something together. For the other half of the words, I want you to make up a sentence that has the word in it. Whenever I show you a card laid out like this (keyword example) you should make up a sentence with the sound-alike word and the meaning doing something together. Whenever I show you a card laid out like this one (context example) you should make up a sentence that has the word in it used correctly. For every other card you'll make up a sentence with the sound-alike word and the meaning doing

> something together, and for the other cards you'll make up a sentence with just the word used correctly.

The Word List 1 study items were presented. Immediately following presentation of the last item, recall of the word meanings was tested using the test flashcards. Recall was self-paced and responses were recorded verbatim.

Next, subjects in both conditions were told that they would be presented with another list of words, and that this time they could choose to use either one of the two strategies. Instruction was as follows:

> Now I am going to present another list of vocabulary words for you to learn. What you must remember for each word is the one-word meaning. You can use either one of the methods of learning that we have discussed. You can either make up a sentence that has the sound-alike word and the meaning in it, or you can make up a sentence that uses the word correctly. Which do you want to do? Make up a sentence that uses the sound-alike word and the meaning, or make up a sentence that uses the word correctly?

Confidence in the strategy selection was ranked using a five-point scale, then subjects were asked to explain their choices and to indicate which of the strategies they thought would produce better learning of the meanings. Performance estimates were elicited as follows:

> How many of the meanings do you think you got right on the test? Of that number, how many meanings do you think you learned by using the sound-alike word and the meaning in a sentence? How many words do you think you learned by making up a sentence with the word used correctly?

Subjects ranked themselves on a five-point scale with respect to how confident they were that they had recalled more items with the strategy with which they expected to do better. They then were given their actual score out of 22, and the procedure of estimating performance and providing a confidence rating for that estimate was repeated.

The subject's strategy choice determined which version of the second word was administered. Recall was tested immediately following presentation of Word List 2. The WAIS vocabulary subtest was administered to all participants, followed by the subject data questionaire. All subjects were given an explanation of the purpose of the study and any questions were answered.

Results

Subjects were asked to report their age, sex, educational history and living arrangements. They were asked to rate their health and memory for everyday tasks relative to others their own age (using a five-point scale) and to report the use of any medication known to affect memory functioning. They were also asked to report any difficulties they experienced in seeing or hearing the materials used in the experimental tasks. Vocabulary skill was assessed by administering the vocabulary subscale of the WAIS. Table 3 provides data on subject characteristics by condition for each age group.

Table 3

Subject Data as a Function of Age and Point of Prediction

Variable		Before Study Older	Before Study Younger	Before Study Student	After Test Older	After Test Younger	After Test Student
N		20	20	20	20	20	20
Age	Mean	73.1	33.6	19.0	72.2	32.5	20.3
	S.D.	7.5	6.0	0.7	7.5	4.3	4.3
WAIS Vocabulary							
	Mean	14.2	11.9	12.8	14.1	12.1	12.3
	S.D.	2.3	1.0	1.4	2.2	1.6	1.6
Years of Education							
	Mean	12.3	14.1	14.0	12.2	15.7	14.1
	S.D.	3.0	2.5	0.1	2.7	2.1	0.1
Level of Education							
elementary		4	0	0	2	0	0
secondary		8	8	0	10	4	0
post-secondary		8	12	20	8	16	20
Sex	Males	10	6	5	9	7	6
	Females	10	14	15	11	13	14
Residence							
alone		8	0	1	7	3	0
with Spouse		12	5	1	13	3	0
Spouse,child		0	12	0	0	11	1
with child		0	1	0	0	1	0
with parents		0	2	6	0	0	3
with roomate		0	0	11	0	2	15
room and board		0	0	1	0	0	1
Health							
below average		0	0	0	2	1	0
average		6	7	13	7	7	9
above average		14	13	7	11	12	11
Memory							
below average		0	0	0	2	2	1
average		14	7	10	15	16	13
above average		6	13	10	3	2	6
Medication		0	0	0	0	0	0
Trouble seeing		0	0	0	1	0	0
Trouble hearing		0	0	0	1	0	0

The between subject results are presented first. The monitoring judgments included in these analyses were those made before-study by subjects in the before-study condition and those made immediately after testing by subjects in the after-test condition.

Two within-subject comparisons were also conducted: 1) a comparison of before-study and first after-test performance measures for subjects in the before-study condition, and 2) a comparison of first and second after-test performance measures for all subjects. The within-subject design includes a reactive component: requesting monitoring judgments before study alerts subjects to the fact that the experimenter is interested in this aspect of performance. After-test monitoring judgments may reflect this heightened awareness. A great deal of this reactivity was noted in the pilot phases of Pressley, Levin and Ghatala (1984) and Pressley, Ross, Levin and Ghatala (1984). In the between-subject design, on the other hand, after-test monitoring judgments are made without prior knowledge of the monitoring task.

It should be noted, however, that subjects clearly attempt to do the recall task as requested; they do not merely try to match their performance to their before-study predictions. This is evident from the finding that accuracy is not greater before study than after-test (Pressley, Levin and Ghatala, 1984; Pressley, Ross, Levin and Ghatala, 1984; Experiments 1 and 2 reported here).

Administering the second word list legitimized the request that subjects choose a strategy to "use on the next list". Thus, this second list was kept as short as possible, and was not considered an interpretable measure of performance.

Between-subject comparisons. For each subject there were three aspects to Word list 1 performance: 1) actual recall, 2) predicted recall, and 3) recall prediction accuracy. The actual recall data consisted of a score for recall of keyword items, a score for recall of context items, and a score for the difference between keyword and context recall. The difference score was obtained by constructing a derived variable, keyword recall minus context recall (signed). The predicted recall data likewise consisted of a score for predicted keyword recall, a score for predicted context recall, and a score for the difference between keyword and context predictions. The difference score was obtained by constructing a derived variable, keyword prediction minus context prediction (signed). Two derived variables were constructed for each subject to measure recall prediction accuracy: 1) the absolute value of keyword recall minus keyword prediction, and 2) the absolute value of context recall minus context prediction. Scores on these variables were unsigned since the object was to assess the extent, not the direction, of error in performance monitoring. Table 4 presents means and standard deviations for each of these dependent variables, by age level and condition.

Specific sets of planned pairwise comparisions were conducted for each dependent variable to assess, first, age differences within each condition, and second, differences across the study-test cycle. The Type I error rate for each pairwise comparison was .0167. Thus, the overall error rate for each variable was .15, which is the overall error rate that would be obtained in the corresponding 3x2 analysis of variance with a Type I error rate of .05 per effect. The critical value for each pairwise comparison can be expressed as $\underline{t}(114)=2.43$.

Age differences were assessed within each condition by comparing scores on each dependent variable between age levels. Table 5 presents a summary of the $\underline{t}$ statistics associated with each of these pairwise comparisons. There were no age differences in context recall. Older adults recalled fewer keyword items than either students or non-student younger adults, and non-student younger adults in turn recalled fewer keyword items than students. Keyword performance was superior to context performance for all age levels. This effect was greatest for students and least for older adults. Recall predictions generally paralleled actual recall, especially in the after-test condition. The only age

Table 4

Mean Performance on Vocabulary Learning Task as a Function of Age and Point of Prediction

Age/Point of Prediction	Recall K	Recall C	Recall K-C	Prediction K	Prediction C	Prediction K-C	Accuracy K	Accuracy C
Older/ Before Study	4.70 (2.08)	2.95 (2.42)	1.75 (2.38)	5.50 (3.22)	4.85 (3.17)	0.65 (5.23)	2.80 (1.88)	3.80 (2.71)
Older/ After Test	4.40 (2.41)	2.35 (1.93)	2.05 (2.59)	2.80 (2.38)	2.95 (2.76)	0.15 (3.65)	2.15 (1.46)	1.60 (1.43)
Younger/ Before Study	7.20 (2.46)	3.45 (2.28)	3.75 (2.15)	6.70 (2.74)	5.00 (2.53)	1.70 (4.06)	2.00 (1.65)	2.15 (1.73)
Younger/ After Test	6.60 (2.16)	3.10 (1.65)	3.50 (2.12)	6.05 (2.19)	2.50 (1.54)	3.55 (1.79)	1.35 (1.09)	1.10 (1.12)
Student/ Before Study	8.05 (2.19)	3.40 (1.57)	4.65 (2.39)	7.50 (3.02)	6.40 (3.00)	1.10 (5.61)	2.45 (2.14)	3.90 (2.73)
Student/ After Test	8.55 (2.06)	3.05 (2.68)	5.50 (2.93)	8.45 (2.39)	2.45 (2.82)	6.00 (3.84)	1.10 (1.12)	1.10 (0.97)

Notes:
1. Standard deviations in brackets.
2. K: Keyword method of learning,
   C: Context method of learning.
3. Maximum score possible: 11.
4. Accuracy measure: (|Recall - Prediction|).

difference in monitoring accuracy was that, in the before-study condition, non-student younger adults predicted their context recall more accurately than did either older adults or students.

To assess the impact of the study-test cycle, before-study performance was compared to after-test performance within each age level. Table 6 provides a summary of the *t* statistics associated with each of the pairwise comparisons conducted at each age level. Actual recall did not differ between conditions. Context predictions were lower in the after-test condition than in the before-study condition for all age levels. For older adults only, keyword predictions were also lower in the after-test condition than in the before-study condition. Subjects at all age levels showed a tendency towards more accurate predictions after test than before study. This tendency reached statistical significance for students (both study methods) and older adults (context study method).

Several other aspects of performance were compared using the chi-square test of significance based on a per comparison type I error rate of .05. Age differences in the direction of monitoring error were assessed for each condition. The number of subjects who over- and underpredicted their recall was compared between age levels. (Subjects who predicted their recall accurately

Table 5

Summary of t-Statistics for Between Age Level Comparisons of Performance on Vocabulary Learning Task for Each Condition

Comparison	Recall			Prediction			Accuracy	
	K[a]	C[b]	K-C[c]	K[d]	C[e]	K-C[f]	K[g]	C[h]
Before Study								
Older vs Younger	3.52*	0.75	1.45	1.41	0.18	0.80	1.57	2.70*
Older vs Student	4.72*	0.67	3.77*	2.35	1.82	0.45	0.69	0.16
Younger vs Student	1.20	0.07	1.17	0.94	1.65	0.60	0.88	2.87*
After Test								
Older vs Younger	3.10*	1.12	1.62	3.82*	0.53	2.78*	1.57	0.82
Older vs Student	5.85*	1.04	4.41*	6.65*	0.59	4.62*	2.06	0.82
Younger vs Student	2.75*	0.07	2.60*	2.82*	0.06	1.84	0.49	0.00

*critical $t(114)=2.43$ at $p<.0167$ per comparison
[a]MSE=4.99
[b]MSE=4.53
[c]MSE=6.11
[d]MSE=7.19
[e]MSE=7.23
[f]MSE=17.19
[g]MSE=2.97
[h]MSE=4.90

Table 6

Summary of t-Statistics for Between Condition Comparisons of Performance on Vocabulary Learning Task for Each Age Level

Comparison	Recall			Prediction			Accuracy	
	K[a]	C[b]	K-C[c]	K[d]	C[e]	K-C[f]	K[g]	C[h]
Older								
Before Study vs After Test	0.42	0.90	0.39	3.18*	2.24	0.60	1.27	3.61*
Younger								
Before Study vs After Test	0.85	0.52	0.32	0.76	2.94*	1.41	1.27	1.72
Student								
Before Study vs After Test	0.70	0.52	1.10	1.12	4.65*	3.74*	2.65*	4.59*

*critical $t(114)=2.43$ at $p<.0167$ per comparison
[a]MSE=4.99
[b]MSE=4.53
[c]MSE=6.11
[d]MSE=7.19
[e]MSE=7.23
[f]MSE=17.19
[g]MSE=2.97
[h]MSE=4.90

were omitted from these analyses). The critical value for each comparison can be expressed as $\underline{X}^2$ (2)=5.99. Table 7 presents obtained and expected values and a summary of the chi-square statistics.

Direction of monitoring error did not differ significantly between age levels in the before-study condition. For all age levels, context predictions tended to be too high, while keyword predictions were as likely to be too high as too low. After testing age differences were not found in the direction of monitoring error for keyword predictions; subjects at all age levels tended to underpredict their recall. However, age differences were found in the direction of error for context predictions after the test. Older adults tended to overpredict their context recall while younger adults tended to make predictions that were too low.

Subjects were asked to indicate which of the two methods of learning they would prefer to use. Within each age level, the number of subjects who chose each method before study was compared to the number of subjects who chose each method after testing. The critical value for these comparisons can be expressed as $\underline{X}^2$(1)=3.84. Table 8 presents obtained and expected values and a summary of the chi-square statistics. Students and non-student younger adults showed an increasing preference for the keyword method across the

Table 7

Summary of Chi-Square Statistics for Between Age Level Comparisons of the Number of Subjects Who Over- and Underpredicted Their Performance on the Vocabulary Learning Task for Each Condition

Comparison	Overpredicted	Underpredicted	Chi-Square
KEYWORD			
Before Study			
Older	11 (8.67)	7 (9.33)	
Younger	7 (8.67)	11 (9.33)	1.93
Student	8 (8.67)	10 (9.33)	
After Test			
Older	4 (5.78)	15 (13.20)	
Younger	5 (4.87)	11 (11.13)	1.85
Student	5 (3.45)	6 (7.65)	
CONTEXT			
Before Study			
Older	14 (14.50)	4 (3.46)	
Younger	13 (12.92)	3 (3.08)	0.18
Student	15 (14.50)	3 (3.46)	
After Test			
Older	9 (5.25)	5 (8.75)	
Younger	3 (4.50)	9 (7.50)	6.63*
Student	3 (5.25)	11 (8.75)	

*critical value $X^2(2)=5.99$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

Table 8

Summary of Chi-Square Statistics for Between Condition Comparisons of the Number of Subjects Who Chose Each Strategy on the Vocabulary Learning Task for Each Age Level

Comparison	Keyword	Context	Chi-Square
Older			
Before Study	8 (7.50)	12 (12.50)	
After Test	7 (7.50)	13 (12.50)	0.13
Younger			
Before Study	13 (16.50)	7 (13.50)	
After Test	20 (16.50)	0 (3.50)	8.48*
Student			
Before Study	12 (15.00)	8 (5.00)	
After Test	18 (15.00)	2 (5.00)	4.80*

*critical value $X^2(1)=3.84$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

study-test cycle. Older adults on the other hand, showed a consistent preference for the context method.

A summary of the relationship between the recall, prediction and strategy choice data is presented in Table 9. Table 9 reveals that in both conditions older adults showed much more variability in their pattern of predictions and strategy choices, and in whether these reflected actual recall, than did students and non-student younger adults. Students and non-student younger adults exhibited an increasingly consistent pattern across the study-test cycle, while older adults did not.

After indicating their strategy choice, subjects were asked to explain why they preferred the strategy they had chosen. These reasons were categorized as either task-oriented (i.e., any response which indicated that the subject had made the strategy choice by considering which method would lead to superior recall) or not task-oriented (i.e., any response that indicated that the subject had used some criteria other than superior recall to make the strategy choice). Within each age level the number of subjects who provided each type of rationale before study was compared to the number who provided each type after testing. The critical value for each comparison can be expressed as $\chi^2(1)=3.84$. Obtained and expected values for the number of subjects

Table 9

Summary of Monitoring Performance as a Function of Age Level and Point of Prediction

			Before Study			After Test		
Choice	Predict	Score	Older	Younger	Student	Older	Younger	Student
k	k	k	3	13	12	4	17	18
k	k	c	2	0	0	1	1	0
k	c	k	0	0	0	1	0	0
k	c	c	0	0	0	0	0	0
c	c	c	2	0	0	3	0	1
c	c	k	6	5	6	5	0	0
c	k	c	0	0	0	0	0	0
c	k	k	3	1	1	3	0	0
k	t	t	0	0	0	0	0	0
k	t	k	3	0	0	1	0	0
k	k	t	0	0	0	0	2	0
k	t	c	0	0	0	0	0	0
k	c	t	0	0	0	0	0	0
c	t	t	0	0	0	0	0	1
c	t	c	0	0	0	0	0	0
c	c	t	0	0	1	1	0	0
c	t	k	0	1	0	1	0	0
c	k	t	1	0	0	0	0	0

Notes: 1. k = Keyword
c = Context
t = Tied

2. Choice = Strategy Choice
Predict = Greater recall predicted for that strategy
Score = Greater recall obtained with that strategy

at each age level who made each type of response, and a summary of the chi-square statistics are presented in Table 10.

Few older adults in either condition gave task-oriented reasons for their strategy choices. Slightly more than half of the students and non-student younger adults gave task-oriented reasons in the before-study condition, and almost all students and non-student younger adults gave task-oriented reasons in the after-test condition. Typically, reasons which were not task-oriented focused on the familiarity and/or ease of execution of the selected method.

Subjects used a five-point scale to provide confidence ratings for their strategy choices and recall predictions. Means and standard deviations for these ratings are summarized in Table 11. Confidence ratings for each dependent variable were compared between conditions for each age level. The per contrast Type I error rate was .0167. Thus, the overall error rate for each variable was .05, which is the overall error rate that would be obtained in the corresponding one-way analysis of variance. The critical value for each pairwise comparison can be expressed as t (114)=2.43. Table 12 presents a summary of the t statistics associated with each pairwise comparison. An increase in confidence from before study to after test was displayed

Table 10

Summary of Chi-Square Statistics for Between Condition Comparisons of the Number of Subjects Who Provided Task-Oriented and Not Task-Oriented Strategy Choice Rationale on the Vocabulary Learning Task for Each Age Level

Comparison	Task-Oriented	Not Task-Oriented	Chi-Square
Older			
Before Study	5 (5.00)	15 (15.00)	
After Test	5 (5.00)	15 (15.00)	0.00
Younger			
Before Study	11 (14.50)	9 (5.50)	
After Test	18 (14.50)	2 (5.50)	6.14*
Student			
Before Study	13 (15.50)	7 (4.50)	
After Test	18 (15.50)	2 (4.50)	3.58

*critical value $X^2(1)=3.84$, $p<.05$, per comparison

Notes: 1. Expected values in brackets.

Table 11

Mean Confidence Ratings for Strategy Choices and Predictions on Vocabulary Learning Task as a Function of Age and Point of Prediction

Age/Point of Prediction	Strategy Choice	Prediction
Older/ Before Study	3.50 (1.24)	3.00 (1.38)
Older/ After Test	3.40 (1.10)	3.55 (1.10)
Younger/ Before Study	3.55 (0.76)	3.40 (1.05)
Younger/ After Test	4.20 (0.89)	4.25 (0.85)
Student/ Before Study	3.85 (0.67)	3.60 (0.75)
Student/ After Test	4.55 (0.51)	4.45 (0.69)

Notes: 1. Standard deviations in brackets.
2. Maximum rating possible: 5.

Table 12

Summary of t-Statistics for Between Condition Comparisons of Confidence Ratings on Vocabulary Learning Task for Each Age Level

Comparison	Strategy Choice[a]	Prediction[b]
Older		
Before Study vs After Test	0.35	1.75
Younger		
Before Study vs After Test	2.32	2.07
Student		
Before Study vs After Test	2.47*	2.70*

*critical $t(114)=2.43$ at $p<.05$
[a] MSE=0.80
[b] MSE=0.99

by the two groups of younger adults, but not by the older adults.

Within-subject comparisons. Only subjects in the before-study condition were included in the within-subjects analysis. In the first within-subjects analysis, the impact of test-taking on monitoring accuracy was assessed by comparing the accuracy of predictions made before study to the accuracy of predictions made after testing. For each study method three derived variables were constructed for each subject: 1) the absolute value of actual recall minus before-study prediction, 2) the absolute value of actual recall minus first after-test prediction, and 3) the absolute value of (1) minus (2). Means and standard deviations pertinent to these indices are presented in Table 13.

The Type I error rate for each comparison between before-study and after-test accuracy was .0167. This corresponds to an overall Type I error rate of .05 for each dependent variable, which is the Type I error rate that would be obtained in the corresponding one-way analysis of variance. The critical value for each comparison can be expressed as $\underline{t}(57)=2.47$.

Table 13

Means for Within Subject Comparisons
of Before Study and After Test Accuracy
on Vocabulary Learning Task for Each Age Level

Age	Accuracy
KEYWORD	
Older	1.50 (1.79)
Younger	1.55 (1.43)
Student	1.90 (1.92)
CONTEXT	
Older	2.60 (2.46)
Younger	1.40 (1.54)
Student	3.10 (2.51)

Notes: 1. Standard deviations in brackets.
2. Accuracy measure:
(|Before Study Recall-Estimate|-
|After Test Recall-Estimate|)

Table 14 provides a summary of the $\underline{t}$ statistics associated with each of the pairwise comparisons conducted with each method of learning. Improved monitoring accuracy after testing was observed with both study methods for all age levels.

The effect of test-taking on several other aspects of performance was assessed using the chi-square test of significance based on a per comparison Type I error rate of .05. Age differences in the direction of monitoring error after testing were assessed by comparing the number of subjects at each age level who over- and underpredicted their recall. (Accurate estimates were omitted). The critical value for these comparisons can be expressed as $\underline{X}^2(2)=5.99$. Table 15 presents obtained and expected values and a summary of the chi-square statistics. No age differences in direction of monitoring error were obtained. Subjects at all age levels tended to underpredict their keyword recall and were as likely to underpredict as to overpredict their context recall.

The between-subjects analysis had revealed a shift towards greater preference for the keyword method after testing for younger adults, but not for older adults. Preference shifts were examined within-subjects by comparing the number of subjects who chose each study method before study to the the number who chose each

Table 14

Summary of t-Statistics for Within Subject Comparisons of Before Study and After Test Accuracy on Vocabulary Learning Task for Each Age Level

Comparison	Accuracy
KEYWORD[a]	
Older	2.75*
Younger	2.84*
Student	3.48*
CONTEXT[b]	
Older	4.76*
Younger	2.56*
Student	5.68*

*critical $t(57)=2.47$ at $p<.0167$ per comparison
[a]MSE=2.99
[b]MSE=4.90

Table 15

Summary of Chi-Square Statistics for Between Age Level Comparisons of the Number of Subjects Who Over- and Underpredicted Their Performance on the Vocabulary Learning Task After Testing: Subjects in Before-Study Condition Only

Comparison	Overpredicted	Underpredicted	Chi-Square
KEYWORD			
<u>After Test</u>			
Older	5 (5.15)	14 (13.85)	
Younger	4 (3.25)	8 (8.75)	0.34
Student	4 (4.60)	13 (12.40)	
CONTEXT			
<u>After Test</u>			
Older	8 (6.50)	5 (6.50)	
Younger	5 (7.50)	10 (7.50)	2.70
Student	7 (6.00	5 (6.00)	

*critical value $X^2(2)=5.99$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

method after testing, for each age level. The critical value for these comparisons can be expressed as $X^2(1)=3.84$. Table 16 presents obtained and expected values and a summary of the chi-square statistics. Younger adults showed an increase in preference for the keyword method across the study-test cycle, while older adults did not.

Table 17 presents a summary of the relationship between strategy selections, performance predictions and actual performance within subjects. As in the between-subjects analysis, older adults were much more variable than were their younger counterparts.

A within subject comparison of the number of subjects who did and did not give task-oriented reasons for their strategy choices before study and after testing was conducted for each age level. The critical value for each comparison can be expressed as $\underline{X}^2(1)=3.84$. Table 18 presents the obtained and expected values and a summary of the chi-square statistics. For all age levels, the number of subject's who gave task-oriented reasons was greater after testing than before study.

Confidence ratings for strategy choice and predictions were compared within-subjects as well. A derived variable was constructed for each subject by calculating the absolute value of the before-study

Table 16

Summary of Chi-Square Statistics for Within-Subject Comparisons of the Number of Subjects Who Chose Each Strategy on the Vocabulary Learning Task for Each Age Level: Subjects in Before-Study Condition Only

Comparison	Keyword	Context	Chi-Square
Older			
Before Study	8 (9.00)	12 (11.00)	
After Test	10 (9.00)	10 (11.00)	0.40
Younger			
Before Study	13 (16.00)	7 (4.00)	
After Test	19 (16.00)	1 (4.00)	5.62*
Student			
Before Study	12 (15.50)	8 (4.50)	
After Test	19 (15.50)	1 (4.50)	7.02*

*critical value $X^2(1)=3.84$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

Table 17

Summary of Monitoring Performance as a Function of Age Level and Point of Prediction: Subjects in Before-Study Condition Only

			Before Study			After Test		
Choice	Predict	Score	Older	Younger	Student	Older	Younger	Student
k	k	k	3	13	12	4	19	17
k	k	c	2	0	0	2	0	0
k	c	k	0	0	0	2	0	1
k	c	c	0	0	0	1	0	0
c	c	c	2	0	0	1	0	0
c	c	k	6	5	6	3	1	1
c	k	c	0	0	0	0	0	0
c	k	k	3	1	1	4	0	0
k	t	t	0	0	0	0	0	0
k	t	k	3	0	0	1	0	0
k	k	t	0	0	0	0	0	1
k	t	c	0	0	0	0	0	0
k	c	t	0	0	0	0	0	0
c	t	t	0	0	0	[illegible]	0	0
c	t	c	0	0	0	[illegible]	0	0
c	c	t	0	0	1	1	0	0
c	t	k	0	1	0	1	0	0
c	k	t	1	0	0	0	0	0

Notes: 1. k = Keyword
c = Context
t = Tied
2. Choice = Strategy Choice
Predict = Greater recall predicted for that strategy
Score = Greater recall obtained with that strategy

Table 18

Summary of Chi-Square Statistics for Within-Subjects Comparisons of the Number of Subjects Who Provided Task-Oriented and Not Task-Oriented Strategy Choice Rationale on the Vocabulary Learning Task for Each Age Level: Subjects in Before Study Condition Only

Comparison	Task-Oriented	Not Task-Oriented	Chi-Square
<u>Older</u>			
Before Study	5 (9.50)	15 (10.50)	
After Test	14 (9.50)	6 (10.50)	7.12*
<u>Younger</u>			
Before Study	11 (15.50)	9 (4.50)	
After Test	19 (15.50)	1 (4.50)	11.62*
<u>Student</u>			
Before Study	13 (16.50)	7 (3.50)	
After Test	20 (16.50)	0 (3.50)	8.48*

*critical value $X^2(1)=3.84$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

confidence rating minus the after-test confidence rating, for each dependent measure. Means and standard deviations for these ratings are summarized in Table 19. The Type I error rate for each comparison between before-study and after-test ratings was .0167. Thus, the overall error rate for each variable was .05, which is the overall error rate that would be obtained in the corresponding one-way analysis of variance. The critical value for each pairwise comparison can be expressed as $\underline{t}(57)=2.47$.

Table 20 presents a summary of the $\underline{t}$ statistics associated with each pairwise comparison. Only one significant difference was observed; students were more confident of their predictions after testing than they had been before study.

The second within-subjects analysis compared first and second after-test monitoring judgments by subjects in all conditions. The first set of after-test recall predictions were made without the assistance of the experimenter. The second set were made after the subject had been told how many items in total he or she had recalled correctly. The number of subjects who changed the direction of their relative predictions (i.e., from more keyword items recalled to more context items recalled or vice versa) was assessed. Only two subjects, one older adult and one student, exhibited this reversal;

Table 19
Mean (|Before Study - After Test|) Difference Scores for Confidence Ratings for Strategy Choices and Predictions on Vocabulary Learning Task as a Function of Age: Subjects in the Before Study Condition Only

Age	Strategy Choice	Prediction
Older	0.70 (1.21)	0.40 (1.05)
Younger	0.15 (1.14)	0.60 (1.14)
Student	0.50 (0.88)	0.80 (0.83)

Notes: 1. Standard deviations in brackets.

Table 20
Summary of t-Statistics for Within Subject Comparison of Confidence Ratings for Strategy Choices and Predictions on Vocabulary Learning Task for Each Age Level: Subjects in the Before Study Condition Only.

Age	Strategy Choice[a]	Prediction[b]
Older	2.00	1.25
Younger	0.43	1.88
Student	1.43	2.50*

*critical value $t(57)=2.47$ at $p<.0167$ per comparison
[a]MSE=1.19
[b]MSE=1.03

all others adjusted their first set of after-test predictions to match the given total without changing their expectations as to which method had resulted in superior recall. Confidence ratings for the second set of after-test predictions in most cases were identical to confidence ratings for the first set of after-test predictions.

Discussion

The most important finding of Experiment I was that both younger and older adults were capable of monitoring their ongoing cognitive performance. Subjects at all age levels displayed an improvement in the accuracy of their performance predictions across the study-test cycle. This finding is consistent with the previous research by Pressley and associates (Pressley, Levin and Ghatala, 1984; Pressley, Ross, Levin and Ghatala, 1984), which found that adults used the metacognitive experience of testing to enhance the accuracy of their monitoring judgments. It is important to note that both younger and older adults displayed skill in monitoring their ongoing cognitive activities, despite there being age differences in actual performance levels.

This finding indicates that there may be some age-related deterioration in memory abilities, but this deterioration is not necessarily accompanied by a corresponding deterioration in metacognitive skill. Thus, it may be that older adults are able to compensate for the minor memory decrements they experience by having accurate metacognitive awareness. This hypothesis has been postulated by other researchers to account for the discrepancy between the poor performance of community-dwelling older adults on experimental tests of memory, and their competent functioning in their own environments (Kausler, 1982; Lachman et al., 1979; Sehluster, 1981).

Thus, the present results support an optimistic view of the aging process, insofar as the procedural, or monitoring aspects of metacognition are concerned. This finding is in agreement with the findings of previous age difference research (e.g., Lachman and Jelilian, 1984; Devolder and Pressley, in preparation). With respect to the other aspect of metacognition assessed in the present study, that is, the ability to predict upcoming cognitive performance without the benefit of concurrent testing experience, however, previous research is not so optimistic. It has typically been found that older adults tend to overpredict their abilities when asked to guess how well they will do on an upcoming memory task, while younger adults, who tend to be more accurate, underpredict (e.g., Bruce et al., 1981; Murphy et al., 1982).

The present results revealed the typical tendency towards overprediction by older adults, but no age differences in either accuracy or direction of error for before-study predictions. This finding suggests that characteristics of the task and stimuli, rather than age *per se* may determine the pattern of over- and underpredictions that is obtained. Lovelace and Marsh (1984) have suggested that overprediction is likely to be found when the subject, through some variable such as lack of experience with the task, underestimates task difficulty. In the present study, subjects at all age levels indicated that they thought the context method of learning would be easier than it was, and that they were unaware of the potency of the keyword method. Thus, they provided predictions for recall with the two study methods which were similar. These predictions were closer to actual keyword recall levels than to actual context levels, with the result that context performance was overpredicted for all age levels, and keyword performance was overpredicted for approximately half the subjects in each age level, and underpredicted for the other half.

While the present results did not provide evidence for age differences in the accuracy of performance monitoring or the direction of error in recall prediction, age differences in metacognition were suggested by other performance measures. Subjects were asked to indicate which learning strategy, keyword or context, they preferred. For the majority of subjects at each age level recall with the keyword method was superior. Few subjects had this item of metacognitive knowledge available to them before taking the test. Not surprisingly, then, no dramatic preferences for the keyword method were observed in the before-study condition. If effective monitoring occurred during the task, however, subjects should have detected the superior efficacy of the keyword method and indicated a preference for this method after testing. The two groups of younger adults displayed this pattern. Older adults, on the other hand, did not show an increase from before study to after test in their preference for the keyword method.

This finding implies deficient monitoring on the part of older adults. Borkowski et al. (submitted) have suggested that when a monitoring deficiency is detected it is worthwhile to consider the role played by the motivational components of metacognitive knowledge. The strategy choice data are pertinent to this consideration. Subjects were asked to justify their strategy choices.

The reasons given were classified as either "task-oriented" (i.e.,reasons which made reference to the experimenter-defined goal of recalling as many words as possible), or not task-oriented. Reasons which were classified as not task-oriented typically made reference to the familiarity, apparent logic and/or relative simplicity of the chosen strategy (students: 89%; non-student younger adults: 77%; older adults: 78%). The remaining subjects who gave not task-oriented responses either indicated that they had made a random choice or they did not make a clear choice but rather rephrased one of the strategies.

More than half of the young adults made their strategy selections on task-oriented grounds even before study. After testing, almost all younger adults selected the method they thought would result in better recall. Few older adults, on the other hand, gave task-oriented reasons in either condition.

This difference did not occur because older adults were incapable of giving task-oriented reasons for their strategy choices. In the within-subjects analysis, the number of subjects at all age levels who gave task-oriented reasons for their strategy choices increased substantially from before study to after test. The within-subjects design may have functioned like the "prompt" condition in Pressley, Ross, Levin and Ghatala's (1984) study with children; alerting subjects to the experimenter's interest in relative strategy efficacy judgments (by soliciting monitoring judgments both before study and after testing) may have increased the likelihood that subjects would attend to this aspect in making their strategy choices. Even in the within-subjects design, however, older adults did not show an increase in preference for the keyword method. This pattern occurred even in several cases where older adults acknowledged that their performance had been superior with the keyword method.

While on the surface this failure to prefer the keyword method, especially after testing, implies a metacognitive deficiency on the part of older adults, it may actually represent adaptive cognitive behaviour. Processing resources diminish or become less efficient with age (e.g., Craik and Rabinowitz, 1984). In light of this, it is wise for older adults to focus on the ease and familiarily of the cognitive task, since by doing so they enhance the likelihood that they will achieve the desired cognitive outcome, in spite of their diminished resources. In the present study the potency of the keyword method rendered this focus on efficiency maladaptive; it is likely that on 'real-life' cognitive tasks an emphasis on efficiency to compensate for age-related limitations serves a useful function. In light of this hypothesis, it is not surprising that, unlike their younger counterparts, older adults did not show an increase in confidence in their performance predictions and strategy choices across the study-test cycle.

It has been argued that students are an inappropriate comparison sample for age difference research, insofar as they are more 'test-sophisticated' than older adults (e.g., Poon, 1985). The validity of this criticism was assessed in Experiment 1 by including both students and non-student younger adult samples. The non-students were selected from the same rural-oriented city as were their older counterparts. On average, they were somewhat older than the typical first-year university student. Consequently, even though many of these individuals had attended university or college, in most cases they were several years removed from that experience. Thus, the non-student younger adults and the older adults shared a common community environment, and were similarily unaccustomed to studying and test-taking, relative to the sample of university students.

Nonetheless, performance by the non-student younger adults was much more similar to that of the students than it was to that of the older adults. Therefore, any differences in monitoring performance observed between younger and older adults cannot be dismissed as an artifact of the subject selection procedure.

The results of Experiment 1 supported the view that metacognitive monitoring and prediction skills are relatively stable across the adult lifespan, despite a decrease in recall abilities. It appeared, however, that there were some differences in the way in which younger and older adults approached cognitive tasks. However, whether these differences were adaptive or maladaptive appeared to be tied to the ecological-validity of the experimental task. Accordingly, in Experiment 2 age differences in metacognition were assessed using memory tasks with real-life analogues.

Experiment 2

Introduction

Three concerns raised by the results of Experiment 1 were addressed in Experiment 2. The first was whether age differences in memory monitoring would be found when more ecologically-valid memory tasks were used. In previous research, age differences have not been found on ecologically-valid memory monitoring tasks (e.g., Lachman et al., 1979; Lachman and Jelilian, 1984; Perlmutter, 1978). Furthermore, in Experiment 1 (reported here) older adults were biased against using the keyword strategy because of its unfamiliarity. Thus, in Experiment 2, age differences in memory monitoring skills were examined using tasks with real-life analogues.

There is a potential for circularity in the concept of ecological validity. Specifically, tasks on which age differences are found may be seen, after the fact, as less ecologically valid, while tasks on which age differences are <u>not</u> found are seen as more ecologically valid. To avoid this pitfall, the tasks used in Experiment 2 were designated, a priori, as more ecologically valid than the tasks used in Experiment 1. This designation was based on both intuition and on the conclusions of previous researchers (e.g., Dixon and Hultsch, 1983a, 1983b; Kausler, 1982; Perlmutter, 1978).

The three tasks used in Experiment 2 were a prose recall task, a vocabulary definition task, and a digit span task. The prose task involved monitoring retention of bits of information read in a magazine-type article. Recounting information one has read is a familiar experience for most adults. The other two tasks, a vocabulary definition task and a digit span task, are unlikely to be encountered in 'real life' in the structured format in which they were presented experimentally. However, they do have analogues in the everyday experience of adults. The vocabulary definition monitoring task involved judging whether one had defined words correctly; lexical decisions of this nature are made during the course of reading and conversation, and on the basis of one's judgments, clarification may or may not be sought. The digit span task involved judging whether one had repeated a number series in the correct order. Filling out order forms, dialing telephone numbers and similar daily activities require this sort of judgment.

Given that more ecologically-valid materials were to be used, the second concern of Experiment 2 was to assess age differences in the ability to monitor ongoing cognitive activities. Thus, as in Experiment 1, the study-test paradigm was used.

There is some research on the role test-taking plays in monitoring of prose materials. Glenberg, Epstein and their associates (Glenberg, Wilkinson and Epstein, 1982; Epstein, Glenberg and Bradley, 1984) used a detection technique similar to that developed for use with children by Markman (1977) to assess how accurately subjects judged their comprehension of prose passages which contained major contradictions. These monitoring judgments were made after study, not after testing. Typically, the contradictions were undetected and the subjects provided high ratings for comprehension.

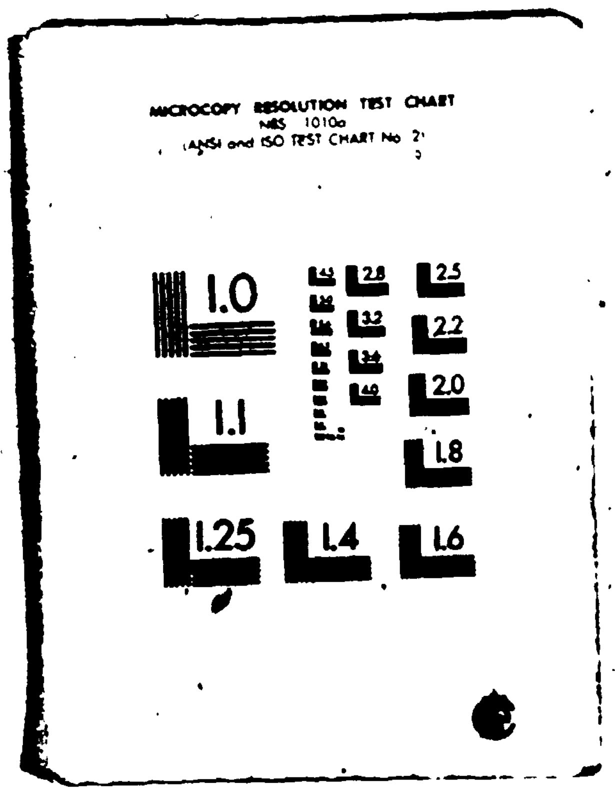
Similar findings were reported by Maki and Berry (1984). These researchers assessed monitoring of prose by university students by having them predict their performance on an upcoming test immediately after studying textbook passages. Low correlations were found between test performance and ratings for the likelihood that a subject would correctly answer questions about what they had read.

Pressley, Snyder, Levin, Murray and Ghatala (submitted) addressed the possibility that the metacognitive experience of test-taking plays an important role in the ability to accurately monitor memory for prose. They assessed university students' awareness of the need to restudy textbook passages in order to obtain a given grade on a test. Subjects predicted how well they would do on a test after one reading of the text, and predicted whether they would have to reread the text in order to score 20, 40, 60 and 80 percent correct on the test. Performance estimates made after the test were more accurate than estimates made before study. Furthermore, need to restudy estimates became increasingly more accurate from before study to after study to after test.

These findings indicate that efficient monitoring of prose by students is enhanced by test-taking. Thus, as was the case with unconnected stimuli, the prose results indicate that mature cognizers are capable of monitoring their ongoing cognitive performance. Study provides for some enhancement of monitoring accuracy, but test experience provides for much more.

Thus, in Experiment 2, changes in monitoring accuracy across the study-test cycle were assessed. With the vocabulary, definition and digit span tasks, before-study monitoring judgments were compared to after-test monitoring judgments, as had been done in Experiment 1. On the prose task, it was possible to compare monitoring at three points in the study-test cycle: before study, after study and after test. It was expected that both younger and older adults would show improved monitoring accuracy across the study-test cycle.

2


MICROCOPY RESOLUTION TEST CHART
NBS 1010a
(ANSI and ISO TEST CHART No. 2)
1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

The third concern of Experiment 2 was to compare subjective variables between younger and older adults in an effort to shed light on monitoring performance differences. Dixon and Hultsch (1983a, 1983b) had developed the Metamemory in Adulthood (MIA) questionaire in an effort to identify differences in metamemory functioning between younger and older adults. The MIA contains ecologically-relevant questions pertaining to eight dimensions of memory performance. Dixon and Hultsch found differences between younger and older adults in that affective dimensions of metacognitive knowledge were more strongly related to the prose recall performance of older adults than to that of younger adults.

Thus, in a semi-structured interview format, subjects in Experiment 2 were asked specific questions about their memory performance on the prose task, as well as some questions about their memory in general. These questions were based on the eight dimensions of memory performance identified by Dixon and Hultsch (1983a, 1983b). There was at least one question designed to assess how the subject felt about his or her memory performance from the perspective of each of the following dimensions: use of memory strategies, knowledge of memory tasks, knowledge of memory capacities, perception of memory change, activities supportive of memory, memory and anxiety, achievement motivation and locus of control. The 128-item MIA itself was not administered since the time required to completed this instrument would have made the experimental sessions too long.

As in Experiment 1, three groups of subjects were included in Experiment 2: community-dwelling older adults, non-student community-dwelling younger adults, and university students. Experiment 1 revealed few performance differences between the two groups of younger adults. It was expected that this pattern likewise would be observed in Experiment 2.

Method

Subjects. Subjects were 60 university students between 17 and 31 years of age (*M*=19.8, *S.D.*=2.9), 60 community-dwelling non-student younger adults between 21 and 40 years of age (*M*=30.8, *S.D.*=5.3), and 60 community-dwelling older adults between 60 and 93 years of age (*M*=70.8, *S.D.*=6.9). All subjects were tested within a 6 month period. Subject selection was the same as in Experiment 1. Individuals who participated in Experiment 1 were not eligible for participation in Experiment 2.

Design. There were two factors in Experiment 2, age and condition. Both factors were assessed between conditions. For the prose task, there were three levels of age, (student, non-student younger adult, older adult) and three conditions (before-study, after-study, after-test). Twenty subjects from each age group were included in each condition.

For the vocabulary and digit span tasks, there were three levels of age, (student, non-student younger adult, older adult) and two conditions (before-study, after-test). Thirty subjects from each age group served in each condition. That is, all subjects who were in the before-study condition for the prose measure, and half of the subjects who were in the after-study condition for the prose measure were in the before-study condition for the vocabulary and digit span measures. All subjects who were in the after-test condition for the prose measure and the remaining half of the subjects who were in the after-study condition for the prose measure were in the after-test condition for the vocabulary and digit span measures. All subjects completed the prose recall task first, with the vocabulary and digit span tasks counterbalanced in second and third position. The design of Experiment 2 in depicted in Table 21.

Table 21

Design of Experiment 2

	Condition (n=20)	
Order of Presentation	Prose Vocabulary Digit Span (n=10)	Prose Digit Span Vocabulary (n=10)

Notes: 1. 3 conditions, 3 age levels: N = 180.
2. Subjects in After-Study Condition were reassigned to either Before-Study or After-Test Condition after completing the Prose Task.

__Materials.__ The prose passage and the cloze test used in this experiment are presented in Appendices F and G, respectively. The prose passage was derived from predictions for the future found in __The People's Almanac Presents the Book of Predictions__ (Wallechinsky, Wallace and Wallace, 1981). The passage was 718 words long (typed on three and a half pages) and was designed to represent the sort of casual magazine reading that might be done by adults of all ages. A passage concerned with the future was chosen to avoid the problem that individuals of different ages may be differentially knowledgable about and/or interested in historical information or current events. One might argue that the young could be more interested in the future since they might live to see it, however, self-reports concerning interest (presented later) defuse this argument. There were 40 underlined words or phrases, approximately equally spaced throughout the passage. Each underlined word or phrase represented a specific concept; for example 'test-tube babies', 'unemployment' and 'people of all ages'. The cloze test was identical to the prose passage except that the underlined words or phrases were replaced by underlined spaces 25 letters in length.

The nature of the prose task was explained by the Experimenter by referring to a sample passage and test. The sample prose passage was 95 words long (one-half page typed) and had four underlined words or phrases. The sample passage dealt with an eighteenth century clarvoyant. The sample test was identical to the sample passage, with the exception that the four underlined words were replaced with underlined spaces 25 letters in length. The sample passage and test are presented in Appendix H.

The vocabulary subscale and the forward digit span subscale of the Weschler Adult Intelligence Scale-Revised (WAIS-R) (Weschler, 1981) was also used. These are presented in Appendices I and J, respectively. Each of the 35 vocabulary words was printed on a 7.5 by 12.5 cm flashcard. Three sample vocabulary words were also printed on flashcards. These 3 words, 'slice', 'cavern' and 'impale' are three of the words used on the vocabulary subtest of the WAIS, but not of the WAIS-R. All words were printed in lowercase letters approximately 5 cm high. (Note that the items from the WAIS-R were presented in random order by shuffling these cards rather than in the order prescribed by the test publisher. This modification was necessary since the WAIS-R was given for two purposes in Experiment 2: to assess vocabulary skill and as one of the monitoring tasks). Each of the 14 number series from the digit span task was likewise printed on a flashcard in numerals approximately 5 cm high.

A questionaire (presented in Appendix K) which pertained to memory in general and to performance on the prose recall task in particular was also used, as was a subject data questionaire which differed little from the one used in Experiment 1. This questionaire is presented in Appendix L.

Procedure. Table 22 presents a summary of the procedure followed in Experiment 2. Participants were tested individually. Students were seen in one of two similar offices at the University of Western Ontario, London. Non-student younger adults and older adults were seen in their homes. The sample passage and test were used to instruct all subjects in the procedure of the prose recall task, as follows:

> I have a number of tasks for you to do today. I'll describe each as we go along. In the first task, you are going to read a short story. Some of the words in the story will be underlined, as in this example. You will pay special attention to these underlined words. After you have read the story once, you will give it back to me and I will give you a

Table 22

Summary of Procedure for Experiment 2

Condition		
Before Study	After Study	After Test
Prose instruction	Prose instruction	Prose instruction
Prose monitoring		
Prose study	Prose study	Prose study
Distraction	Prose monitoring	Distraction
Prose test	Prose test	Prose test
		Prose monitoring
Digit Span instruction[a]		Digit Span instruction
Digit Span monitoring		
Digit Span test		Digit Span test
		Digit Span monitoring
Vocabulary instruction[b]		Vocabulary instruction
Vocabulary monitoring		
Vocabulary test		Vocabulary test
		Vocabulary monitoring
Questionaire		Questionaire
Subject data		Subject data
Debriefing		Debriefing

[a]Digit Span presented as third task to half subjects in each condition.
[b]Vocabulary presented as second task to half subjects in each condition.

fill-in-the-blank test. The test will look like this: You see, the underlined words are missing. Your job is to try and fill in each of these blanks. You do not have to remember the exact words. That is important, so I want you to be sure you understand it: You can fill in the blanks with any words which express the thought contained in the original underlined phrase. For example, the words in this sample story are the French Revolution. On the test, your answer would still be right if you wrote down the civil war in France.

The story which you are going to read is set up just like this example, except that it is three and a half pages long and there are 40 underlined phrases and words. You are going to read the story once only. What that means is that when you finish a paragraph or turn a page, you will not go back. You can, however, take as long as you want to go through the story the one time. When you finish the story, I will give you the test. You can take as much time as you want to do the test and you can go through the test as many times as you want. You must put down an answer for every blank, even if you have to guess.

Subjects in the before-study condition were asked to estimate how well they would do on the fill-in-the-blank test, as follows:

> I want you to think about the story and the test as I've described them to you and tell me how many of the 40 blanks you think you will be able to fill in correctly on the test. What do you think your score will be out of 40? What is the very best score you think you might get? What is the very worst score you think you might get? (This format was followed for all subsequent requests for performance estimates.)

After subjects in the before-study condition had made their performance estimates, and immediately after instruction on the prose task for subjects in the other two conditions, all subjects were instructed to:

> Take this story and read it through once only, at your own speed. Remember, do not backtrack and pay special attention to the ideas contained in the underlined words and phrases.

When the subjects indicated they had finished reading the story, those in the after-study condition were asked to estimate how well they would do on the upcoming test. Subjects in the before-study and after-test conditions were engaged for approximately one minute in a distraction task. Specifically, they were told:

> Just before we go on, I should fill in this spot on my record sheet to make sure I have you in the right category. What year were you born? So that makes you how old now?

All subjects then completed the cloze test at their own speed. Following the test, subjects in the after-test condition were asked to estimate their performance.

The prose task was presented first to all subjects. Presentation of the vocabulary and digit span tasks was counterbalanced in second and third position. Instruction in the vocabulary-digit span order of presentation was as follows:

> Now we will move on to the second task. This time, I am going to give you a list of 35 words. Your task will be to give me a definition for each word. About a third of the words are fairly easy - an example is, 'slice', which can be defined as a thin piece of something. Another third are more difficult, such as the word 'cavern', which means a large underground hole, and the other

third of the words are quite difficult. An example is 'impale' which means to pierce through with a sharp object.

Subjects in the before-study condition were asked to estimate how many words they would be able to define correctly, and to predict the very best and very worst scores they might get. The vocabulary test was administered to all subjects. The flashcards were shuffled for each subject and the experimenter both pronounced the word and showed the flashcard with the word printed on it. Responding was self-paced, and responses were written down verbatim. Following the test, subjects in the after-test condition were asked to estimate their performance.

The digit span task was introduced as follows: In the third task, I am going to read strings of numbers to you. For example, I will read 9-1-5. You will repeat what I said back to me. So, for this example, you would say to me "9-1-5". I will read 14 strings of numbers to you, one at a time. There will be between 3 and 9 numbers in each string and I will read the strings in random order.

Subjects in the before-study condition were asked to estimate how many numbers there would be in the longest string that they would be able to remember in the correct order, and to predict their very best and very worst possible scores. The digit span test was administered to all subjects. The flashcards were shuffled for each subject, and the experimenter read the numbers at a rate of one digit per second. Subjects were not allowed to see the stimuli. Responding was self-paced, and the stimuli and responses were written down verbatim. Following testing, subjects in the after-test condition were asked to estimate their performance.

Next, the memory questionaire was administered. Subjects were told:

> What I would like you to do now is to use this scale to answer some questions about the first test you took, the fill-in-the-blank test, and about your memory in general. When I ask you a question, I want you to answer by picking one of the five response choices from this scale. On some of the questions, I will ask you to go on and explain your choice.

When the memory questionaire had been completed, the subject data questionaire was administered, and subjects were debriefed.

Results

Subjects in Experiment 2 were asked for the same personal data as were subjects in Experiment 1. Subjects reported their sex, age, educational history, living arrangements, health, memory, use of medication, and any difficulty they had seeing or hearing the test materials. Vocabulary skill was assessed by administering the vocabulary subscale of the WAIS-R. Table 23 provides data on subject characteristics by condition for each age group. There were three experimental conditions in Experiment 2: before study, after study and after test. Subjects were assigned to conditions in an approximately random fashion.

All comparisons were conducted between subjects. For each subject in Experiment 2, there were four aspects to performance on each task: 1) actual recall, 2) predicted recall, 3) range of predicted recall (i.e., 'best' minus 'worst' possible score) and 4) recall prediction accuracy. Recall prediction accuracy was assessed by constructing a derived variable, the absolute value of the actual recall score minus the predicted recall score. Scores on this variable were unsigned since the goal was to assess absolute monitoring accuracy, not direction of error. Specific sets of planned pairwise comparisons were conducted for each

Table 23

Subject Data as a Function of Age and Point of Prediction

Variable		Before Study O	Y	S	After Study O	Y	S	After Test O	Y	S
N		20	20	20	20	20	20	20	20	20
Age	Mean	69.5	31.8	19.8	72.2	30.4	19.6	70.7	30.3	20.1
	S.D.	7.7	5.0	3.1	6.0	5.1	2.3	6.9	6.0	3.4
WAIS Vocabulary										
	Mean	11.8	10.7	10.7	9.0	11.2	11.8	11.5	10.6	10.7
	S.D.	1.7	3.0	2.2	1.6	2.3	2.2	2.0	2.7	2.2
Years of Education										
	Mean	12.2	14.7	13.1	10.5	14.5	13.0	12.0	14.1	12.9
	S.D.	3.6	2.0	1.5	3.9	2.3	0.5	2.6	3.0	0.3
Level of Education										
elementary		5	0	0	8	0	0	4	0	0
secondary		5	5	0	6	6	0	8	9	0
post-secondary		10	15	20	6	14	20	8	11	20
Sex	Males	4	8	6	9	9	4	2	7	4
	Females	16	12	14	11	11	16	18	13	16
Residence										
alone		7	0	1	9	1	1	8	1	1
with Spouse		11	8	1	10	7	0	9	5	0
Spouse,child		1	10	1	1	8	1	2	9	0
with child		1	0	0	0	1	0	1	1	0
with parents		0	1	5	0	3	2	0	2	1
with roomate		0	1	11	0	0	15	0	2	14
room and board		0	0	1	0	0	1	0	0	4
Health										
below average		2	1	1	2	1	1	0	1	1
average		3	7	9	14	8	9	10	6	8
above average		15	12	10	4	11	10	10	13	11
Memory										
below average		3	1	2	2	0	1	0	1	4
average		12	12	11	14	13	10	17	13	9
above average		5	7	7	4	7	9	3	6	7
Medication		0	0	0	0	0	1	1	0	0
Trouble seeing		0	0	0	0	0	0	1	0	0
Trouble hearing		0	0	0	1	0	0	0	0	0

dependent variable to assess, first, age differences within each condition, and second, changes in performance across the study-test cycle.

The prose task. Table 24 presents means and standard deviations for each dependent variable, by age level and condition. The per contrast Type I error rate was .0167. Thus, the overall Type I error rate for each variable was less than or equal to .15, which is the overall Type I error rate that would be obtained in the corresponding 3x3 analysis of variance with a Type I error rate of .05 per effect. The critical value for the comparisons involving each dependent variable can be expressed as $t(171)=2.43$.

Age differences were assessed within each condition by comparing scores on each dependent variable between age levels. Table 25 presents a summary of the t statistics associated with each of these pairwise comparisons. An age difference in favour of the younger adults was obtained for actual performance on the prose task. There were, however, no age differences in recall predictions. The only age differences in accuracy were that in the after-study condition, older adults monitored their performance less accurately than non-student younger adults and students, and in the before-study condition older adults provided a smaller prediction range than did either group of younger adults. In

Table 24

Mean Performance on Prose Task as a Function of Age and Point of Prediction

Age/Point of Prediction	Recall	Prediction	Accuracy	Range	Correct
Older/ Before Study	18.05 (5.99)	23.90 (6.51)	7.75 (3.88)	11.05 (5.29)	55
Older/ After Study	10.30 (4.61)	21.85 (6.45)	11.55 (5.70)	12.70 (6.95)	40
Older/ After Test	16.90 (7.06)	21.00 (7.97)	6.50 (4.99)	10.10 (7.22)	65
Younger/ Before Study	24.15 (6.88)	23.90 (5.52)	5.95 (3.65)	17.40 (6.44)	70
Younger/ After Study	23.70 (6.11)	24.50 (6.93)	4.00 (2.32)	15.70 (6.72)	90
Younger/ After Test	23.35 (8.06)	21.90 (7.14)	4.15 (3.94)	11.55 (3.93)	75
Student/ Before Study	27.45 (6.48)	23.20 (6.93)	6.25 (6.19)	16.95 (6.24)	75
Student/ After Study	26.00 (5.60)	23.25 (9.61)	6.60 (5.33)	15.70 (4.93)	70
Student/ After Test	23.15 (5.60)	19.35 (9.61)	5.90 (5.33)	8.10 (4.93)	40

Notes: 1. Standard deviations in brackets.
2. Maximum score possible: 40.
3. Accuracy measure: (\|Recall - Prediction\|).
4. Range measure: (Best Possible - Worst Possible Score)
5. Correct measure: Percent of subjects for whom Range encompassed Recall Score.

Table 25

Summary of t-Statistics for Between Age Level Comparisons of Performance on Prose Task for Each Condition

Comparison	Recall[a]	Prediction[b]	Accuracy[c]	Range[d]
Before Study				
Older vs Younger	3.10*	0.00	1.23	3.38*
Older vs Student	4.78*	0.31	1.01	3.13*
Younger vs Student	1.68	0.31	0.20	0.24
After Study				
Older vs Younger	6.81*	1.17	5.11*	1.60
Older vs Student	7.98*	0.62	3.34*	1.60
Younger vs Student	1.68	0.31	0.20	0.24
After Test				
Older vs Younger	3.28*	0.40	1.35	0.77
Older vs Student	3.18*	0.73	0.41	1.06
Younger vs Student	0.10	1.12	1.18	1.84

*critical $t(171)=2.43$ at $p<.0167$ per comparison
[a]MSE=38.75
[b]MSE=51.44
[c]MSE=21.85
[d]MSE=35.33

general, the prediction range of older adults was less likely to encompass the actual score.

To assess the impact of the study-test cycle, performance was compared between conditions within each age level. Table 26 provides a summary of the *t* statistics associated with each of these pairwise comparisons. Older adults in the after-study condition both performed more poorly, and predicted their performance more poorly than did older adults in either of the other conditions. For subjects at all age levels, the tendency (non-significant) was towards an improvement in monitoring from before study to after testing. For all age levels the size of the prediction range was smaller after testing than it had been before study. The likelihood that the prediction range encompassed the actual score showed a corresponding decrease only for students.

Age differences in the direction of monitoring error were assessed by comparing the number of subjects at each age level who overpredicted their performance to the number who underpredicted, within each condition. (Subjects who predicted their performance accurately were omitted from these analyses). These comparisons were conducted using the chi-square test of significance, based on a per comparison Type I error rate of .05. The critical value for each comparison can be expressed as

Table 26

Summary of t-Statistics for Between Condition Comparisons of Performance on Prose Task for Each Age Level

Comparison	Recall[a]	Prediction[b]	Accuracy[c]	Range[d]
Older				
Before vs After Study	3.94*	0.90	2.57	0.88
After Study vs After Test	3.35*	0.37	3.42*	1.38
Before Study vs After Test	0.58	1.28	0.85	0.51
Younger				
Before Study vs After Study	0.23	0.26	1.32	0.53
After Study vs After Test	0.18	1.15	0.10	2.21
Before Study vs After Test	0.41	0.88	1.22	3.11*
Student				
Before Study vs After Study	0.74	0.02	0.24	0.66
After Study vs After Test	1.45	1.72	0.47	4.04*
Before Study vs After Test	2.18	1.70	0.24	4.71*

*critical $t(171)=2.43$ at $p<.0167$ per comparison
[a] MSE=38.75
[b] MSE=51.44
[c] MSE=21.85
[d] MSE=35.33

$\chi^2(2)=5.99$. Table 27 presents obtained and expected values and a summary of the chi-square statistics. In all conditions older adults were more likely to overpredict their scores, while younger adults were more likely to underpredict.

The vocabulary task. All subjects were requested to predict how many words they would be able to/had defined correctly. Responses were scored either 0, 1 or 2 according to the WAIS-R criteria. For the purposes of determining a score for the number of words defined correctly, responses scored either 1 or 2 were considered correctly defined. Table 28 presents means and standard deviations for each dependent variable. The per contrast Type 1 error rate was .0167. Thus, the overall Type 1 error rate for each variable was less than or equal to .15, which is the overall error rate that would be obtained in the corresponding 3x2 analysis of variance with a Type 1 error rate of .05 per effect. The critical value for the comparisons involving each dependent variable can be expressed as $t(174)=2.43$.

Age differences within each condition were assessed by comparing scores on each dependent variable between age levels. A summary of the t statistics associated with these pairwise comparisons is presented in Table 29. There were no age differences in vocabulary scores or estimates. Only one comparison revealed an age

Table 27

Summary of Chi-Square Statistics for Between Age Level Comparisons of the Number of Subjects Who Over- and Underpredicted Their Performance on the Prose Task for Each Condition

Comparison	Overpredicted	Underpredicted	Chi-Square
Before Study			
Older	17 (10.18)	3 (9.82)	
Younger	9 (10.18)	11 (9.82)	17.08*
Student	3 (8.65)	14 (8.35)	
After Study			
Older	19 (12.33)	0 (6.67)	
Younger	10 (12.98)	10 (7.02)	15.53*
Student	8 (11.68)	10 (6.32)	
After Test			
Older	15 (10.00)	5 (10.00)	
Younger	5 (8.00)	11 (8.00)	8.14*
Student	7 (9.00)	11 (9.00)	

*critical value $X^2(2)=5.99$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

Table 28

Mean Performance on Vocabulary Defining Task as a Function of Age and Point of Prediction

Age/Point of Prediction	Recall	Prediction	Accuracy	Range	Correct
Older/ Before Study	27.93 (3.76)	20.47 (7.14)	8.67 (5.26)	9.20 (4.31)	30
Older/ After Test	25.67 (5.59)	27.80 (6.04)	4.10 (3.28)	7.83 (5.61)	53
Younger/ Before Study	26.23 (6.54)	22.57 (5.48)	5.03 (3.60)	11.47 (5.58)	53
Younger/ After Test	27.93 (5.71)	27.93 (5.21)	3.00 (2.42)	6.60 (3.60)	67
Student/ Before Study	26.57 (4.78)	21.17 (5.50)	7.13 (5.21)	11.07 (5.54)	40
Student/ After Test	25.33 (4.31)	25.03 (6.74)	3.97 (2.97)	6.70 (4.46)	50

Notes:
1. Standard deviations in brackets.
2. Maximum score possible: 40.
3. Accuracy measure: (|Recall - Prediction|).
4. Range measure: (Best Possible - Worst Possible Score)
5. Correct measure: Percent of subjects for whom Range encompassed Recall Score.

Table 29

Summary of t-Statistics for Between Age Level Comparisons of Performance on Vocabulary Defining Task for Each Condition

Comparison	Recall[a]	Prediction[b]	Accuracy[c]	Range[d]
Before Study				
Older vs Younger	0.94	1.35	3.50*	1.79
Older vs Student	0.76	0.45	1.48	1.47
Younger vs Student	0.19	0.90	2.02	0.31
After Test				
Older vs Younger	1.26	0.08	1.06	0.97
Older vs Student	0.19	1.78	0.13	0.89
Younger vs Student	1.44	1.86	0.93	0.08

*critical $t(174)=2.43$ at $p<.0167$ per comparison
[a]MSE=27.04
[b]MSE=36.69
[c]MSE=15.53
[d]MSE=24.13

difference in monitoring accuracy: non-student young adults were more accurate than older adults, in the before-study condition. There were no age differences in the size of the prediction ranges. In general, the prediction ranges of older adults were the least likely to contain the actual score, while the prediction ranges of non-student younger adults were the most likely to contain the actual score.

To assess the role of test-taking, scores on each dependent varible were compared within each age level between conditions. A summary of the _t_ statistics associated with these pairwise comparisons is presented in Table 30. For all age levels, before-study predictions were lower than after-test predictions. After-test monitoring was significantly more accurate than before-study monitoring for older adults and students, and this trend was apparent, although not significant, for non-student younger adults. All age levels reduced their prediction ranges after testing. This reduction was significant for the two groups of young adults. For all age levels the likelihood that the prediction range encompassed the actual score increased from before study to after test.

Table 30

Summary of t-Statistics for Between Condition Comparisons of Performance on Vocabulary Defining Task for Each Age Level

Comparison	Recall[a]	Prediction[b]	Accuracy[c]	Range[d]
Older				
Before Study vs After Test	1.26	4.70*	4.39*	1.10
Younger				
Before Study vs After Test	0.94	3.44*	1.95	3.83*
Student				
Before Study vs After Test	0.69	2.74	3.04*	3.44*

*critical $t(174)=2.43$ at $p<.0167$ per comparison
[a]MSE=27.04
[b]MSE=36.69
[c]MSE=15.53
[d]MSE=24.13

Age differences in the direction of monitoring error were assessed between age levels within each condition by comparing the number of subjects who overpredicted their recall to the number who underpredicted. These comparisons were conducted using the chi-square test of significance, based on a per contrast Type I error rate of .05. (Accurate estimates were omitted from these analyses). The critical value for each comparison can be expressed as $X^2(2)=5.99$. Obtained and expected values and a summary of the chi-square statistics are presented in Table 31. Subjects at all age levels underpredicted their performance before study. After testing, older adults tended to overpredict their scores, while younger adults tended to underpredict their scores.

The digit span task. Table 32 present means and standard deviations for each dependent variable, by age level and condition. The per contrast Type I error rate was .0167. Thus, the overall Type I error rate for each variable was less than or equal to .15, which is the overall error rate that would be obtained in the corresponding 3x2 analysis of variance with a Type I error rate of .05 per effect. The critical value for the comparisons involving each dependent variable can be expressed as $t(174)= 2.43$.

Table 31

Summary of Chi-Square Statistics for Between Age Level Comparisons of the Number of Subjects Who Over- and Underpredicted Their Performance on the Vocabulary Defining Task for Each Condition

Comparison	Overpredicted	Underpredicted	Chi-Square
<u>Before Study</u>			
Older	4 (4.09)	26 (25.91)	
Younger	2 (3.82)	26 (24.18)	1.98
Student	6 (4.09)	24 (25.91)	
<u>After Test</u>			
Older	19 (12.78)	5 (11.22)	
Younger	9 (13.31)	16 (11.69)	9.98*
Student	13 (14.91)	15 (13.09)	

*critical value $X^2(2)=5.99$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

Table 32

Mean Performance on Digit Span Task as a Function of Age and Point of Prediction

Age/Point of Prediction	Recall	Prediction	Accuracy	Range	Correct
Older/ Before Study	6.67 (0.99)	5.50 (0.97)	1.50 (0.97)	2.00 (1.02)	60
Older/ After Test	7.00 (1.29)	6.83 (1.70)	0.97 (1.07)	1.83 (1.42)	73
Younger/ Before Study	6.73 (1.05)	6.57 (1.22)	1.03 (0.93)	3.37 (1.30)	83
Younger/ After Test	7.20 (1.16)	6.73 (1.14)	1.07 (0.98)	1.67 (1.40)	60
Student/ Before Study	7.43 (0.82)	6.67 (1.32)	1.37 (0.96)	3.07 (1.46)	70
Student/ After Test	7.47 (1.04)	7.07 (1.44)	1.13 (0.86)	1.27 (1.11)	70

Notes: 1. Standard deviations in brackets.
2. Maximum score possible: 9.
3. Accuracy measure: (|Recall - Prediction|).
4. Range measure: (Best Possible - Worst Possible Score).
5. Correct measure: Percent of subjects for whom Range encompassed Recall Score.

To assess age differences within each condition, scores on each dependent variable were compared between age groups. A summary of the t statistics associated with these pairwise comparisons is presented in Table 33. In the before-study condition, students had larger digit spans than either older adults or non-student younger adults. Older adults provided lower predictions before-study than did either students or non-student younger adults. There were no age differences in monitoring accuracy. Before study older adults provided smaller prediction ranges than did either group of younger adults, and the range they provided was the least likely to encompass the actual score. After testing, there were no age differences in the size of the prediction ranges, and the ranges provided by older adults were the most likely to encompass the actual score.

To assess the role of test-taking, scores on each dependent variable were compared within each age level, between conditions. A summary of the t statistics associated with these comparisons is presented in Table 34.. Older adults provided higher predictions in the after-test condition than in the before-study condition. Older adults and students showed a non-significant tendency towards increased monitoring accuracy across the study-test cycle, while non-student younger adults showed

Table 33

Summary of t-Statistics for Between Age Level Comparisons of Performance on Digit Span Task for Each Condition.

Comparison	Recall[a]	Prediction[b]	Accuracy[c]	Range[d]
Before Study				
Older vs Younger	0.21	3.15*	1.88	4.15*
Older vs Student	2.71*	3.44*	0.52	3.24*
Younger vs Student	2.50*	0.29	1.36	0.91
After Test				
Older vs Younger	0.71	0.29	0.40	0.48
Older vs Student	1.68	0.70	0.64	1.70
Younger vs Student	0.96	0.88	0.24	1.21

*critical $t(174)=2.43$ at $p<.0167$ per comparison
[a]MSE=1.14
[b]MSE=1.74
[c]MSE=0.93
[d]MSE=1.68

Table 34

Summary of t-Statistics for Between Condition Comparisons of Performance on Digit Span Task for Each Age Level

Comparison	Recall[a]	Prediction[b]	Accuracy[c]	Range[d]
Older				
Before Study vs After Test	1.18	3.91*	2.12	0.52
Younger				
Before Study vs After Test	1.68	0.47	0.16	5.15*
Student				
Before Study vs After Test	0.14	1.18	0.96	5.54*

*critical $t(174)=2.43$ at $p<.0167$ per comparison
[a]MSE=1.14
[b]MSE=1.74
[c]MSE=0.93
[d]MSE=1.68

stable accuracy. A reduction in the size of the prediction range (significant for the two groups of younger adults) was seen for all age levels. For older adults the likelihood that the prediction range encompassed the actual score increased, while for non-student younger adults it decreased and for students it remained stable.

The direction of monitoring error was compared between age levels within each condition. These comparisons were conducted using the chi-square test of significance, based on a per contrast type I error rate of .05. The critical value for each comparison can be expressed as $\underline{X}^2(2)=5.99$. (Accurate estimates were omitted from these comparisons). Obtained and expected values and a summary of the chi-square statistics are presented in Table 35. Subjects at all age levels underpredicted their scores both before study and after test.

Memory Questionaire. When they had completed the three tasks, subjects were asked to respond to a number of questions about their performance on the prose task, and about their memory performance in general. Responses were made using a five-point scale. Means and standard deviations for responses to the questionaire items are presented in Table 36. Since only age differences were of interest here, condition was not included as a factor

Table 35

Summary of Chi-Square Statistics for Between Age Level Comparisons of the Number of Subjects Who Over- and Underpredicted Their Performance on the Digit Span Task for Each Condition

Comparison	Overpredicted	Underpredicted	Chi-Square
Before Study			
Older	3 (5.62)	23 (20.38)	
Younger	9 (4.32)	11 (15.68)	8.90*
Student	4 (6.05)	24 (21.95)	
After Test			
Older	7 (5.09)	9 (10.91)	
Younger	6 (7.00)	16 (15.00)	1.35
Student	8 (8.91)	20 (19.01)	

*critical value $X^2(2)=5.99$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

Table 36

Mean Responses to Questionaire as a Function of Age

Question	Older	Younger	Student
Interested in story	2.28 (1.25)	2.53 (1.71)	2.70 (1.11)
Memory tricks used on prose	4.77 (0.65)	4.20 (1.15)	3.67 (1.48)
Memory tricks normally used	4.12 (1.30)	3.82 (1.44)	2.55 (1.52)
Good at remembering	3.12 (1.24)	2.82 (1.32)	2.88 (1.22)
Important to have good memory	2.33 (1.42)	2.12 (1.17)	1.50 (0.81)
Important to do well on prose	2.08 (1.48)	2.45 (1.33)	2.47 (1.20)
Nervous taking test	4.17 (1.18)	4.00 (1.31)	3.80 (1.61)
Nervousness affected performance	3.58 (1.48)	3.00 (1.48)	3.50 (1.31)
Could train for another test	3.33 (1.61)	2.38 (1.42)	2.79 (1.48)
Try to keep memory sharp	2.42 (1.48)	2.78 (2.78)	3.30 (1.63)

Notes: 1. Standard deviations in brackets.
2. Responses based on 1 (positive) to 5 (negative) rating scale.

in these analyses. Age differences were assessed by comparing mean responses to each question pairwise between age levels. The per contrast Type I error rate was .0167. Thus, with three contrasts performed for each variable, the overall Type I error rate was less than or equal to .05, which corresponds to the overall error rate that would be obtained in the corresponding one-way analysis of variance. The critical value for each comparison can be expressed as $\underline{t}(177)=2.43$. Table 37 presents a summary of the $\underline{t}$ statistics associated with each pairwise comparison.

Age differences were obtained in the extent to which subjects reported using "tricks" to facilitate performance; older adults reported using tricks less than non-student younger adults, who in turn reported using tricks less than students. This pattern was obtained both when questioned as to whether memory tricks were used on the prose task and when questioned as to whether memory tasks were normally used. Age differences were also obtained in that older adults thought it less likely that they could prepare for a future test and thereby perform better than did non-student younger adults. Older adults thought it was less important to have a good memory than did non-student younger adults, who in turn thought it was less important to have a good memory than did students. However, older adults reported a greater

Table 37

Summary of t-Statistics for Between Age Level Comparisons of Responses to Questionaire

Question	Older vs Younger	Comparison Older vs Student	Younger vs Student
[a]Interested in story	1.16	1.95	0.79
[b]Memory tricks used on prose	2.73*	5.26*	2.54*
[c]Memory tricks normally used	1.15	6.04*	4.88*
[d]Good at remembering	1.30	1.04	0.26
[e]Important to have good memory	0.99	3.87*	2.92*
[f]Important to do well on prose	1.51	1.60	0.08
[g]Nervous taking test	0.76	1.66	0.90
[h]Nervousness affected performance	2.32	0.31	1.92
[i]Could train for another test	3.44*	1.95	1.48
[j]Try to keep memory sharp	1.25	3.06*	1.81

*critical value $t(177)=2.43$, $p<.0167$ per comparison
[a]MSE=1.39
[b]MSE=1.31
[c]MSE=2.03
[d]MSE=1.59
[e]MSE=1.35
[f]MSE=1.80
[g]MSE=1.49
[h]MSE=2.03
[i]MSE=2.29
[j]MSE=2.48

Note: It is acknowledged that the risk of making a Type I error increases when multiple dependent variables are used; however, given the exploratory nature of this questionaire, the risk was considered acceptable in this instance.

tendency than students to deliberately do things to "exercise" their memories.

Four factors, ability, éffort, task difficulty and luck were presented to the subjects in pairs, and the subjects were asked to select the member of each pair which had been more influential in determining performance on the prose task. The number of subjects at each age level who selected each attribute out of a given pair was compared for each pair, using the chi-square test of significance with a per comparison Type I error rate of .05. The critical value for each comparison can be expressed as $\underline{X}^2(2)=5.99$. Obtained and expected values and a summary of the chi-square statistics are presented in Table 38.

Subjects at all age levels viewed ability as more important in determining performance than was effort. Subjects at all age levels viewed task difficulty as more influential than luck. Age differences were found in that older adults considered effort more influential than task difficulty, whereas students and non-student younger adults were more evenly split in their selection of these two attributes. Age differences were also found in that older adults discounted the role of luck in preference for effort less than did non-student younger adults and students.

Table 38

Summary of Chi-Square Statistics for Between Age Level Comparisons of the Number of Subjects Who Made Attributions to Each Factor for Performance on the Prose Task

Comparison	Older	Younger	Student	Chi-Square
Effort	25 (24.00)	25 (24.00)	22 (24.00)	
Ability	35 (36.00)	35 (36.00)	38 (36.00)	0.42
Task	46 (48.33)	52 (48.33	47 (48.33)	
Luck	14 (11.67)	8 (11.67)	13 (11.67)	2.20
Effort	39 (29.67)	27 (29.67)	23 (29.67)	
Task	21 (30.33)	33 (30.33)	37 (30.33)	10.25*
Effort	37 (44.33)	49 (44.33)	47 (44.33)	
Luck	23 (15.67)	11 (15.67)	13 (15.67)	7.13*
Ability	34 (31.00)	27 (31.00)	32 (31.00)	
Task	26 (29.00)	33 (29.00)	28 (29.00)	1.73
Ability	44 (47.33)	52 (47.33)	46 (47.33)	
Luck	16 (12.67)	8 (12.67)	14 (12.67)	3.47

*critical $\underline{X}^2(2)=5.99$, $\underline{p}<.05$ per comparison

Notes: 1. Expected values in brackets.

The mean ages for the three groups in Experiment 1 were 19, 33 and 72, for students, non-student younger adults and older adults, respectively. Subjects in Experiment 2 were asked to speculate on whether they performed better or worse at their present age than they would have if they were the mean age of the two categories to which they did not belong. Thus, for example, older adults were asked whether their performance would have been better or worse at ages 19 and 33 than at their present age.

The number of subjects at each age level who rated their performance as likely to have been better, the same, or worse, at each suggested age was compared using a chi-square test of significance with a Type I error rate of .05 per comparison. The critical value for each comparison can be expressed as $\underline{X}^2(2)=5.99$. Obtained and expected values and a summary of the chi-square statistics are presented in Table 39.

In general, older adults thought that they performed more poorly now than they would have at either age 19 or age 33. Non-student younger adults were fairly evenly divided as to whether their performance would be better, worse or the same if they were 19 or 72 years of age. Students thought that their performance would be worse if they were 33, and even more likely to have deteriorated if they were 72.

Table 39

Summary of Chi-Square Statistics for Within Age Level Comparisons of the Number of Subjects Who Expected Memory to Get Better, Worse or Stay the Same Across the Adult Life Span

Comparison	Better	Same	Worse	Chi-Square
Older				
at 19	45 (45.50)	3 (3.50)	12 (11.00)	
at 33	46 (45.500	4 (3.50)	10 (11.00)	0.48
Younger				
at 19	24 (21.50)	16 (17.00)	20 (21.50)	
at 72	19 (21.50)	18 (17.00)	23 (21.50)	0.90
Student				
at 33	11 (7.50)	11 (7.50)	38 (45.00)	
at 72	4 (7.50)	4 (7.50)	52 (45.00)	8.70*

*critical value $X^2(2)=5.88$, $p<.05$ per comparison

Notes: 1. Expected values in brackets.

Discussion

The results of Experiment 2 lend further support to the hypothesis that metacognitive monitoring skills remain relatively stable across the adult lifespan, despite the occurence of an age-related decrement in recall abilities. On the prose task, subjects at all age levels exhibited a consistent, albeit non-significant shift towards improved prediction accuracy from before-study to after-test. This finding attests to the importance of actual test experience in the monitoring of ongoing cognitive activity.

The after-study results were not consistent with a pattern of progressive improvement for any age level. The greatest accuracy was observed in the after-study condition for the non-student younger adults, and the least accuracy was observed in the after-study condition for students and older adults. These findings are in contrast to the findings of previous research with prose materials (Pressley, Snyder, Levin, Murray and Ghatala, submitted). Previous research indicated that study resulted in slight enhancement of predictive accuracy, and study and testing resulted in even greater enhancement. The reasons for the unexpected pattern of after-study results in the present study were not readily apparent. Analysis of the data with age and education level included as covariates yielded no change in results. It may be that requesting performance predictions between study and testing in the after-study condition introduced interference not present in the before-study and after-test conditions for the students and older adults and that the non-student younger adults were not similarily affected because their performance was somewhat better overall.

Additional evidence for the stability of monitoring skill across adulthood was revealed by the prose prediction range data. The pattern obtained with the before-study and after-test prediction range data generally paralleled the pattern obtained with the before-study and after-test accuracy data. Subjects at all age levels showed a reduction in the size of their prediction ranges (i.e., highest possible score minus lowest possible score) from before study to after test. For older adults and non-student younger adults, this reduction occurred without a corresponding reduction in the likelihood that the prediction range encompassed the actual score: Students exhibited a reduction from 75% before-study to only 40% after test in the likelihood that their prediction range encompassed their actual score. However, when their prediction range plus or minus one was calculated, the percentage obtained after testing returned to the before-study level. It is important to take this observation into consideration, given that the prediction range provided by the students after testing was considerably smaller than that provided by either of the other age groups.

In general, then, the prediction range data indicated that subjects were able to "zero-in" (Pressley, Snyder, Levin, Murray and Ghatala, submitted) on their actual scores better after testing than before study. Not surprisingly, given the accuracy data, the after-study prediction range results were inconsistent with a clear pattern of progressive improvement.

The results obtained with the two discrete item tasks, the vocabulary defining task and the digit span task, generally support the view that subjects at different age levels do not differ in their ability to monitor their ongoing cognitive activities. On the vocabulary-defining task accuracy improved and prediction range decreased (but was as likely to encompass the actual score) from before-study to after-test for all age levels. On the digit span task, older adults and students showed the above pattern of improvement. Non-student younger adults showed evidence of improvement only in their prediction range data. This finding should be interpreted in context, however: due to their greater accuracy initially, they had less room for improvement across the study-test cycle.

Thus, the results of Experiment 2 support an optimistic view of aging: memory recall ability may decrease with age, but the ability to monitor ongoing cognitive activity remains stable. This allows for the possibility that older adults know when and how to adjust their cognitive activities to compensate for the minor age-related memory deficits they experience (e.g., Kausler, 1982).

The monitoring of ongoing cognitive activity is a procedural aspect of metacognitive knowledge. The present study assessed age differences in declarative aspects as well. That is, age differences in the ability to predict future performance without the benefit of current experience with the task were assessed. Some researchers have found age differences in the direction and accuracy of performance predictions made before study (e.g., Bruce et al., 1982), while others have not (e.g., Experiment 1, reported here). Lovelace and Marsh (1984) suggest that the tendency of older adults to overpredict their performance stems from a calibration problem: that is, they overpredict not because their awareness of their own abilities is flawed, but because their lack of recent experience with memory tasks leads them to underestimate task difficulty.

The prose task was administered first to all subjects. On this task, the pattern of overprediction by older adults and underprediction by younger adults was observed. However, on the vocabulary definition and digit span tasks subjects at all age levels underpredicted their performance before study. On the vocabulary task, older adults switched to overprediction after testing, while younger adults continued to be slightly more likely to underpredict. On the digit span task, subjects at all age levels were more likely to underpredict after testing. Thus, a consistent, age-related pattern of over- and underpredictions was not observed.

It is possible that task variables, rather than age <u>per se</u> make the major contribution to whether an individual over- or underpredicts his or her performance. This suggestion is in keeping with the observation made by previous researchers such as Lovelace and Marsh (1984) and Craik and Rabinowitz (1984) that age differences in direction of monitoring error are ameliorated by allowing time for task familiarization and practice.

When all of the tasks used in Experiment 2 had been administered to a subject, the subject was questioned about his or her performance. The most notable data to emerge from this questioning pertained to the issue of strategy use. Older adults reported less use of memory "tricks" on the prose task and less use of memory tricks in general than younger adults. Furthermore, they were less likely to believe that they could do anything to prepare for a future prose task like the one they had just completed.

These age differences appeared attributable to two sources: older adults were less likely to use memory tricks to facilitate remembering because they were less aware of what those tricks might be, and they were less inclined to use any memory tricks of which they were aware. Their lack of knowledge might stem from differences in the educational practices to which younger and older adults have been exposed. In recent years an emphasis on understanding and imagery-based learning methods has supplanted the previous emphasis on rote memorization (e.g., Schaie, 1980). Their disinclination may result from an awareness that memory capacity decreases with age. Thus, older adults may be unwilling to undertake mnemonic activities to enhance remembering because, without being very well versed in those activities, the amount of effort and processing resources required to perform the activities would jeopardize successful completion of the task itself. Both of these hypotheses allow for the possibility that the memory performance of older adults may be enhanced through the provision of sufficient training in appropriate mnemonic techniques.

General Discussion

The most important finding of the present research was that both younger and older adults were able to monitor their ongoing cognitive activity. This finding, discussed below, has implications for an optimistic model of aging. While the primary concern of this research was with procedural aspects of metacognitive knowledge (i.e., performance monitoring across the study-test cycle) age differences in declarative aspects of metacognitive knowledge (i.e., in the ability to predict performance before study on the basis of stored metacognitive knowledge concerning the likely interaction of self, task and stimuli (Flavell and Wellman, 1977)) were assessed as well. This assessment is important, given that almost all of the limited amount of previous research available has addressed age differences in this aspect of metacognitive performance.

The main measure of monitoring skill was a within age level comparison of the accuracy of performance predictions made before study to the accuracy of those made after testing. Previous research using both list stimuli and prose materials has indicated that younger adults are capable of using the metacognitive experience of test-taking to enhance the accuracy of their after-test performance judgements (e.g., Pressley, Ross, Levin and Ghatala, 1984; Pressley, Snyder, Levin, Murray and Ghatala, submitted). Furthermore, previous age difference research using the before-study/after-test design has revealed that both younger and older adults predicted their performance more accurately after test experience (Devolder and Pressley, in preparation; Lachman and Jelilian, 1984). The present results supported the previous research: a consistent tendency in the direction of greater accuracy after testing was observed for all age levels. While this tendency was consistent, however, it was not significant for all age levels on all tasks. This observation motivated a miniature "meta-analysis" (Glass, 1977; Hedges and Olkin, 1985).

The effect size was based on the pooled within-group standard deviation for the before-study and after-test cells for each experimental task, calculated separately for each age level. For older adults the improvement in accuracy across the study-test cycle is represented by the following change in standard deviation units between before-study and after-test prediction accuracy: keyword recall task, .39; context recall task, 1.06; prose recall task, .28; vocabulary defining task, 1.07; and digit span task .52. The mean effect size was .66. For non-student younger adults the changes, in standard deviation units, were as follows: keyword, .47; context, .73; prose, .47; vocabulary defining, .67; and digit span, -.04 (note that this was the one instance in which improved accuracy was not found). The mean effect size was .46. For students, the changes, in standard deviation units, were: keyword, .83; context, 1.51; prose, .06; vocabulary defining, .77; and digit span, .26. The mean effect size was .69.

These mean effect sizes (i.e., .66, .46 and .69 for older adults, non-student younger adults and students, respectively) are "moderate" in size (Cohen, 1977). That is, in many cases the effects were too small to be picked up as statistically significant on the individual tasks used in Experiments 1 and 2. Individual tasks were sensitive only to large effect sizes of one standard deviation or more. However, the meta-analysis revealed that the tendency towards improved accuracy was consistent, albeit undramatic. Most important to the concerns of this study, this tendency was observed regardless of age level.

These monitoring accuracy data are supported by other monitoring measures. On the prose, vocabulary defining and digit span tasks, subjects were requested to provide a prediction range for their scores in addition to estimating the actual score. On all tasks for all age levels the prediction range provided after testing was smaller than that provided before study. On the prose task older adults reduced their prediction ranges by 9%, non-student younger adults reduced their prediction ranges by 34% and students reduced their prediction ranges by 52%. On the vocabulary defining task older adults provided a 15% reduction, non-student younger adults provided a 42% reduction and students provided a 40% reduction. On the digit span task the reductions after testing were 9%, 50% and 59% for older adults, non-student younger adults and students, respectively. In most cases, however, the actual score was as likely or more likely to be encompassed by the prediction range.

Students and non-student younger adults showed a greater decrease in the size of their prediction ranges than did older adults. This finding must be considered in the context of the actual scores: no age differences were found in the size of prediction ranges provided after testing. That is, older adults showed less of a decrease because they started out providing smaller ranges than their younger counterparts, not because their after-test ranges were wider. After testing younger and older adults tended to "zero-in" on their actual scores to a comparable degree.

Additional, although less straightforward, evidence for stability in monitoring skills across the adult lifespan was found in the strategy choice and strategy choice rationale data of Experiment 1. It was expected that if subjects detected the superiority of the keyword learning method this recognition would be reflected in their after-test strategy choices: subjects would prefer the more potent learning method. This pattern was observed for younger adults. While they did not show a pronounced preference for the keyword strategy before study, after test experience they almost unanimously selected the keyword strategy. Furthermore, they justified this preference in terms of potency: the keyword method was preferred because it resulted in greater recall.

Older adults tended to recognize the superiority of the keyword method after testing. However, unlike their younger counterparts they often did not act on this knowledge in the expected fashion. Instead, many older adults chose to stay with the less effective context strategy. They indicated that this was because they perceived the context strategy to be easier, more "sensible" and more similar to their accustomed methods of learning new words than the keyword method.

This may represent adaptive cognitive functioning and is an area of investigation worthy of future research attention. Specifically, the efficient cognizer is able to sequence and plan his or her cognitive activities so as to achieve a given cognitive goal while consuming as few intellectual resources as necessary (Pressley, Borkowski and Schneider, in press). This efficiency is crucial to all cognizers, given the limited capacity of working memory (e.g., Kahneman, 1973). However, it may be especially crucial to older adults. It is widely believed that memory deterioration with age is related to factors such as the slowing of mental processes (e.g., Birren, Woods and Williams, 1980) and the reduction in processing capacity (e.g., Craik and Simon, 1980). These factors place limitations on the processing resources of the older individual (Craik and Rabinowitz, 1984). However, if the individual is aware of this (i.e., his or her general metacognitive knowledge is accurate), then he or she will be able to engineer cognitive activities so as to make maximal use of available resources. Important considerations for future research include the extent to which the consequences of age-related memory deterioration can be ameliorated by training monitoring skills. Training could pertain to specific monitoring skills, with the aim being the recognition of and compensation for specific memory deficits, and to higher order monitoring skills, with the aim being to streamline cognitive functioning so as to maximize use of the available (and perhaps diminishing) resources.

The interview data were pertinent to this issue. Older adults considered it less important to have a good memory than did younger adults. They also made less use of "tricks" to enhance memory performance and thought that there was less that they could do to improve future performance on a similar task. These beliefs could be interpreted as self-defeating and likely to interfere with efficient cognitive functioning (e.g., Borkowski et al., submitted). This interpretation is appropriate from a child-oriented developmental perspective: an important antecedent of the general tendency to deploy task-appropriate learning strategies is the motivation to exert the necessary effort. This motivation derives from the belief that these efforts will be successful. From an aging perspective, however, it may be more adaptive to minimize than to exert effort. Thus, the metacognitive activities of older adults may be directed by accurate awareness of their cognitive limitations.

Thus, the finding that the ability to monitor ongoing cognitive activity is stable across the adult lifespan allows for an optimistic theory of aging. Regulation of one's cognitive activities hinges on having accurate knowledge concerning the progress of those activities: errors cannot be corrected if they are not detected (e.g., Baker and Brown, 1982). If monitoring processes remain intact, it may be possible for older adults who experience minor age-related decrements in other aspects of cognitive functioning to recognize and compensate for these deficits. An older adult with deteriorating recall abilities but accurate metacognitive knowledge would realize the need to supplement his reliance on internal memory aids (which may have been sufficient in the past) with task-appropriate external memory aids. For example, older adults who realize that they have begun to have difficulty keeping track of their financial affairs would be able to delegate the more problematic aspects of their affairs to appropriate sources of help, and thus, would be able to maintain their independence longer than might otherwise be expected.

The primary concern of this research was with changes in monitoring accuracy across the study-test cycle. However, much of the previous research has assessed performance within, rather than across conditions. In general, previous research has provided little evidence for the hypothesis that older adults predict their performance less accurately than do their younger counterparts. In keeping with these findings, there were few significant within-condition age differences in prediction accuracy. However, as was done in the between-condition data, a miniature meta-analysis (Glass, 1977; Hedges and Olkins, 1985) was conducted across tasks within each condition to determine whether any consistencies were present. The effect size was based on the pooled (pairwise) within group standard deviation for age level, calculated separately for each condition. For the before-study condition, age differences in prediction accuracy are represented by the following differences in standard deviation units between older adults and non-student younger adults: keyword, .45; context, .74; prose, .48; vocabulary defining, .82; and digit span, .49. The mean effect size was .60. The differences between older adults and students were as follows: keyword, .17; context, -.04; prose, .30; vocabulary defining, .30; and digit span, .13. The mean effect size was .19. The differences between students and non-student younger adults were: keyword, .24; context, .78; prose, .06; vocabulary defining, .48; and digit span, .36. The mean effect size was .38.

Thus, the meta-analysis revealed a small to moderate tendency for the before-study performance predictions of non-student younger adults to be better than those of both older adults and students, and a very small tendency for the before-study predictions of students to be better than those of older adults.

The meta-analysis procedure was repeated for the after-test condition. Age differences in after-test accuracy are represented by the following differences in standard deviation units between older adults and non-student younger adults: keyword, .63; context, .39; prose, .53; vocabulary defining, .39; and digit span, -.10. The mean effect size was .37. The differences between older adults and students were as follows: keyword, .81;, context, .42; prose, .11, vocabulary defining, .04; and digit span, -.16. The mean effect size was .24. The differences between students and non-student younger adults were: keyword, -.23; context, 0; prose, .38; vocabulary defining, .36; and digit span, .07. The mean effect size was .17.

Thus, consistent with the pattern observed before study, there was a slight tendency for non-student younger adults to be the most accurate and for older adults to be the least accurate. However, the mean effect sizes for these age differences were very small (Cohen, 1977). Moreover, it can be seen that the least ecologically-valid task, the keyword vocabulary learning task, made the most substantial contribution to age differences between younger and older adults.

In previous research direction of error has also been used as a measure of within condition age differences. The general pattern revealed in previous research was that older adults tended to overpredict their performance when assessed before study, while younger adults tended to underpredict. It has been suggested that these tendencies represent a calibration problem: older adults overpredict because they expect tasks to be easier than they are, while younger adults underpredict because of the reverse expectations. The present results offered some support for this hypothesis, not as an explanation for age differences _per se_, but as an explanation for performance differences in direction of monitoring error in general. For example, in Experiment 1, subjects at all age levels indicated that they found the context vocabulary learning task more difficult than expected, and consequently subjects at all ages levels tended to overpredict their performance. Similarily, on the prose task in Experiment 2, older adults indicated that they found the task more difficult than expected, and they again overpredicted. Research is needed to assess the extent to which these mistaken beliefs concerning the nature of the task, rather than mistaken beliefs concerning abilities, contribute to the pattern of over- and under-predictions. It may be that to consider this pattern as indicative of an age difference, rather than as an artifact of the experimental situation is erroneous. Lovelace and Marsh (1984) suggest that considerably more attention should be paid to the issue of pre-experimental task familiarity in age difference research. The present results support this view.

The possibility that older and younger adults may be differentially familiar with the memory and learning tasks used in typical laboratory experiments was addressed in the present research. Previous research had indicated that the ecological validity of the materials and task play a role in determining whether or not age differences are found (e.g., Lachman and Jelilian 1984;

Perlmutter, 1978). The present results supported this view. The most notable differences between younger and older adults occurred with the keyword recall task. Older adults both predicted and monitored their performance with the keyword strategy less accurately than did younger adults. Furthermore, by far the majority of the older adults rejected the keyword method of learning, despite its superior efficacy, on the grounds that it did not seem a sensible or easy way to learn vocabulary words, especially when compared to the context strategy. Younger adults, on the other hand, typically recognized the keyword strategy as a method of "word association", and as such, as a method likely to assist memory.

Given that they were less familiar with the keyword strategy, older adults may have had to focus more of their attention on the technique, and less on the recall task, than younger adults. Since, in the efficient learner, the components of strategy use have become automatic, leaving the learner free to focus exclusively on the cognitive task (Pressley, Forrest-Pressley, Elliot-Faust, 1985 ; Borkowski et al., submitted), any differences between age levels in the efficiency of strategy use would handicap the less-familiar age group. From this perspective, it is not surprising to find that older adults predicted and monitored their keyword performance less accurately than younger adults: the store of general and specific strategy knowledge on which they based their predictions was less complete than that of younger adults, and the cognitive demands of using an unfamiliar strategy would leave few processing resources available to undertake effective monitoring activities.

The ecological-validity of the tasks used to assess age differences in metacognitive skills is of paramount importance. An important consideration for future research is to determine the extent to which various memory tasks relate to the practical experience of adults, both older and younger. One very practical way to approach this problem is by simply asking study participants to rate the experimental tasks in terms of their applicability to real-life situations. As Lachman et al. (1979) state, the maintenance of memory skills may extend over the life span. in healthy adults, when those memory skills are measured by "tasks of high salience to everyone's daily life (p.550)". The image of competent older adults in their own environments is certainly more consistent with reality than is the picture of severe decline obtained in studies using esoteric tasks (Lachman et al., 1979).

Thus, the practical significance of finding age differences on memory tasks with little relevance to the real-life activities of older adults must be considered.

Moreover, the present results suggest that it is not enough to identify age-related deficits in cognitive function; the impact of those deficits must be determined by assessing the relationship between those deficits and the individual's capacity to function effectively. Where the individual is able to compensate for his or her deficits because of accurate metacognitive knowledge, the importance of the deficits is questionable.

Thus, an important concern for future research on age differences in metacognition is the underlying theoretical model on which the research is premised. Current conceptualizations of metacognition pertain largely to the cognitive development of children. Consequently, they emphasize the acquisition of cognitive knowledge and skills. Thus, a deficiency model may be an appropriate framework for viewing the metacognitive development of children (e.g., children progress from being unable to deploy strategies efficiently, to being able, but unlikely, to deploy strategies efficiently, to being efficient strategy users (Borkowski et al., submitted). Adulthood, on the other hand "seems to be a period in which a large arsenal of symbolic ability and learned skills are already available and many changes take place in how adults apply these skills and ability" (Yussen, 1985, p. 271).

Dixon and Hultsch (1983a, 1983b) found evidence to suggest that different dimensions of metacognition were differentially related to monitoring performance in younger and older adults. The present results similarily revealed differences between younger and older adults in several aspects of memory beliefs. Thus, it may be that the priorities and activities of older adults warrant a different approach to cognitive tasks than that of younger adults. Labouvie-Vief (1985) has argued that the tasks of later adulthood require the ability to generalize and to see alternatives, while the tasks of early adulthood are more logic- and goal-oriented. In a similar vein, Yussen (1985) suggests, that Flavell's taxonomy of self, task and strategies be replaced, for an understanding of metacognition in adulthood, by a taxonomy which involves knowledge about human purposes, integating group desires and values, and 'time-lining', or long-term planning. These are cognitive concerns of adulthood.

Thus, while the procedural components of metacognitive knowledge appear to have relevance and stability across the adult lifespan (e.g., the present experiments), it may be that the overall context in which cognitive performance is interpreted follows a developmental course. In tune with the suggestions of Labouvie-Vief (1985) and Yussen (1985), a primary focus of future research on age differences in metacognition should be on identifying and defining the processes and states of metacognitive knowledge that are important to different stages of adulthood.

REFERENCES

Baker, L., & Brown, A. L. (1982). Metacognitive skills in reading. In D. P. Pearson (Ed.), Handbook of Reading Research (pp.353-394). New York: Academic Press.

Birren, J. E., Woods, A. M. & Williams, M. V. (1980). Behavioural slowing with age: Causes, organization and consequences. In L. W. Poon (Ed.), Aging in the 1980s: Psychological Issues (pp. 293-308). Washington, D. C.: American Psychological Association.

Borkowski, J., Carr, M. & Pressley, M. (submitted). "Spontaneous" strategy use: Perspectives from metacognitive theory.

Brown, A., Bransford, J. D. Ferrara, R. A. & Campione, J. C. (1983). Learning, remembering and understanding. In J. H. Flavell & E. M. Markman (Eds.), Handbook of Child Psychology: Vol. III: Cognitive Development (pp. 177-266). New York: John Wiley and Sons.

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

Cohen, J. (1977). Statistical Power for the Behavioural Sciences (rev. ed.). New York: Academic Press.

Craik, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), _Handbook of the Psychology of Aging_ (pp. 384-420). New York: Van Nostrand Reinhold.

Craik, F. I. M. & Rabinowitz, J. C. (1984). Age differences in the acquisition and use of verbal information. In H. Bouma & D. G. Bouwhuis (Eds.) _Attention and Performance: Control of Language Processes_ (pp. 471-499). Hillsdale, N. J.: Lawrence Erlbaum Associates.

Devolder, P. & Pressley, M. (in preparation). University of Western Ontario, London, Ontario.

Dixon, R. A. & Hultsch, D. F. (1983a). Structure and development of metamemory in adulthood. _Journal of Gerontology_, _38_, 682-688.

Dixon, R. A. & Hultsch, D. F. (1983b). Metamemory and memory for text relationships in adulthood: A cross-validation study. _Journal of Gerontology_, _38_, 689-694.

Dixon, R. A. & Hultsch, D. F. (1984). The Metamemory in Adulthood (MIA) Instrument. _Psychological Documents_, _14_, 3.

Epstein, W., Glenberg, A. M. & Bradley, M. M. (1984). Coactivation and comprehension: Contribution of text variables to the illusion of knowing. _Memory and Cognition_, _12_, 355-360.

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

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. American Psychologist, 34, 906-911.

Flavell, J. H. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), Children's Oral Communication Skills (pp. 35-60). New York: Academic Press.

Flavell, J. H. (1985). Cognitive Development. Englewood Cliffs, N. J.: Prentice Hall.

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

Glass, G. V. (1977). Integrating findings: The meta-analysis of research. In L. S. Shulman (Ed.), Review of Research in Education, Vol. 5 (pp.351-379). Itasca, Ill.: Peacock.

Glenberg, A. M., Wilkinson, A. C. & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. Memory and Cognition, 10, 547-602.

Hedges, L. V. & Olkin, I. (1985). Statistical Methods for Meta-analysis. Orlando, Florida: Academic Press.

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

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, N. J.: Prentice Hall.

Kausler, D. H. (1982). Experimental Psychology and Human Aging. New York: John Wiley and Sons.

Labouvie-Vief, G. (1985). Intelligence and cognition. In J. E. Birren and K. W. Schaie (Eds.), Handbook of the Psychology of Aging, 2nd ed. (pp.500-530). New York: Van Nostrand Reinhold.

Lachman, M. E. & Jelilian, E. (1984). Self-efficacy and attributions for intellectual performance in young and elderly adults. Journal of Gerontology, 39, 577-582.

Lachman, J. L., Lachman, R. & Thronesbery, 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.

Maki, R. H. & Berry, S. L. (1984). Metacomprehension of text material. Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 663-679.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. Child Development, 48, 986-992.

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

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

Perlmutter, M. & Mitchell, D. B. (1982). The appearance and disappearance of age differences in adult memory. In F. I. M. Craik & S. Trehub (Eds.), Aging and Cognitive Processes (pp.127-143). New York: Plenum Press.

Poon, L. W. (1985). Differences in human memory with aging: Nature, causes and clinical implications. In J. E. Birren & K. W. Schaie (Eds.), Handbook of the Psychology of Aging, 2nd ed. (pp.427-462). New York: Van Nostrand Reinhold.

Poon, L. W., Krauss, I. K. & Bowles, N. L. (1984). On subject selection in cognitive aging research. Experimental Aging Research, 10, 43-49.

Pressley, M., Borkowski, J. G. & O'Sullivan, J. T. (1983). Children's metamemory and the teaching of memory strategies. In D. L. Forrest-Pressley, G. E. MacKinnon & T. G. Waller (Eds.), Metacognition, Cognition and Human Performance, Vol. 1 (pp. 111-153). New York: Academic Press.

Pressley, M., Borkowski, J. and Schneider, W. (in press). Cognitive strategies: Good strategy users co-ordinate metacognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), Annals of Child Development, Vol.4. Greenwich, Conn.: JAI Press.

Pressley, M., Forrest-Pressley, D. L., Elliot-Faust, D. & Miller, G. (1985). Children's use of cognitive strategies, and what to do if they can't be taught. In M. Pressley & C. J. Brainerd (Eds.), Cognitive Processes in Memory Development (pp. 1-46). New York: Springer-Verlag.

Pressley, M., Levin, J. R. & Delaney, H. D. (1982). The mnemonic keyword method. Review of Educational Research, 52, 61-91.

Pressley, M., Levin, J. R., & Ghatala, E. S. (1984). Memory strategy monitoring in adults and children. Journal of Verbal Learning and Verbal Behaviour, 32, 270-288.

Pressley, M., Ross, K. A., Levin, J. R. & Ghatala, E. S. (1984). The role of strategy utility knowledge in children's stratetgy decision making. Journal of Experimental Child Psychology, 38, 491-504.

Pressley, M, Snyder, B., Levin, J. R., Murray, H. G. & Ghatala, E. S. (submitted). Perceived readiness for examination performance (PREP) produced by initial reading of text and text containing adjunct questions.

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.

Schmitt, F. A., Murphy, M. D. & Sanders, R. E. (1981). Training older adult free recall rehearsal strategies. Journal of Gerontology, 36, 329, 337.

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

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

Schaie, K. W. (1980). Intelligence and problem-solving. In J. E. Birren & R. B. Sloane (Eds.), Handbook of Mental Health and Aging (pp.262-284). Englewood Cliffs, N. J.: Prentice Hall Inc.

Shaughnessey, J. (1981). Memory monitoring accuracy and modification of rehearsal strategies. Journal of Verbal Learning and Verbal Behaviour, 20, 216-230.

Wallechinsky, D., Wallace, A. & Wallace, I. (1981) The People's Allmanac Presents the Book of Predictions. New York: Bantam Books.

Weschler, D. (1955). Weschler Adult Intelligence Scale. New York: The Psychological Corporation.

Weschler, D. (1981). Weschler Adult Intelligence Scale - Revised. New York: The Psychological Corporation.

Yesavage, J. A. (1983). Imagery pretraining and memory training in the elderly. Gerontology, 29, 271-275.

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

Yussen, S. R. (1985). The role of metacognition in contemporary theories of cognitive development. In D. L. Forrest-Pressley, G. E. MacKinnon & T. G. Waller (Eds.), Metacognition, Cognition and Human Performance, Vol. 1 (pp. 253-283, New York: Academic Press.

APPENDICES

APPENDIX A

Experiment 1 Word Lists 1 and 2

Old English Word	Meaning	Keyword
WORD LIST 1		
dogger	boat	dog
carlin	old woman	car
corniche	road	corn
hamate	bone	ham
gunnel	fish	gun
anitar	poison	ant
bustard	bird	bus
claymore	sword	clay
fanfold	tablet	fan
girdler	beetle	girdle
bullace	plum	bull
henbane	plant	hen
windling	straw	wind
manchet	bread	man
ratine	fabric	rat
flagon	pitcher	flag
cowry	seashell	cow
tarn	lake	tar
flamen	priest	flame
catkin	flowers	cat
dottle	tobacco	dot
cordite	explosive	cord
WORD LIST 2		
necroplis	cemetary	neck
postern	gate	post
legist	lawyer	leg
nestor	old man	nest
bolter	machine	bolt
inkle	tape	ink
lapidist	jeweller	lap
melonite	stone	melon
pantheon	temple	pan

APPENDIX B

Experiment I Sample Stimuli

Old English Word	Meaning	Keyword
handsel	payment	hand
jarvey	driver	jar
casern	barracks	case
piggin	bucket	pig

APPENDIX C

Experiment I Rating Scales

1	2	3	4	5
I just took a guess. I'm completely uncertain that I would learn more with the strategy that I said.	Very uncertain that I would learn more with the strategy that I said.	Not really sure but I think that I would learn more with the strategy that I said	Pretty sure that I would learn more with the strategy that I said.	Absolutely certain that I would learn more with the strategy that I said.

1	2	3	4	5
I just took a guess. I'm completely uncertain.	Very uncertain that I learned more with the strategy I I said.	Not really sure, but I think I learned more with the strategy I said.	Pretty sure that I learned more with the strategy I said.	Absolutely certain that I learned more with the strategy I said.

1	2	3	4	5
I just took a guess. I'm completely uncertain.	Very uncertain that I picked the right one.	Not really sure, but I think I picked the right one.	Pretty sure that I picked the right one.	Absolutely certain that I picked the right one.

APPENDIX D

Experiment I WAIS Vocabulary Test

bed
ship
penny
winter
repair
breakfast
fabric
slice
assemble
conceal
enormous
hasten
sentence
regulate
commence
ponder
cavern
designate
domestic
consume
terminate
obst[illegible]
remo[illegible]e
sanctuary
matchless
reluctant
calamity
fortitude
tranquil
edifice
compassion
tangible
perimeter
audacious
ominous
tirade
encumber
plagarize
impale
travesty

APPENDIX E

Experiment 1 Subject Data Questionaire

Subject Data

Name:______________________ Age:______ Sex:_____

Education: grade school _____ years
high school _____ years
college/university _____ years

Residence: ____ live alone
____ live with family
____ live in Home for elderly
____ other: please specify ______________

Relative to other people your own age, how would you rate
1) your general health (circle one)

1	2	3	4	5
very poor	worse than average	average	better than average	very good

2) your memory for everyday memory tasks:

1	2	3	4	5
very poor	worse than average	average	better than average	very good

Are you currently taking any medication that is known to affect memory functioning: ____ yes _____no

Did you have any difficulty ____ seeing or _____ hearing the materials used in this task?

APPENDIX F

Experiment 2 Prose Passage

The future cannot be known in advance. To try to predict the future, therefore, is [illegible] even very interesting unless one is a fiction writer seeking to entertain. Thus, it is less important to ask what the future will be than it is to ask what we can make the future become. Both good and bad outcomes can be predicted for the future. The bad outcomes can be considered a challenge, warning us to examine our way of life. The good outcomes can be considered a message of hope, an indication of what might be accomplished with the resources at our disposal.

You may wonder if all, or any, of this forecast is believable. To convince yourself that these predictions are possible, take a look over your shoulder. Had someone told you, only a few years ago that we would soon have portable computers, videotape machines, digital watches, microwave ovens and test-tube babies, you might have had doubts. But it all happened. In light of the recent past, there is no reason to disbelieve the possibility of advances and changes in the future. With all this in mind, then, and if current trends continue, our future might look like the picture described below.

Politically, we are headed towards a time of unrest.

Many cities and countries will come under military rule. Maintaining international trade will be a challenge as diplomatic ties between countries are disrupted. Production of war weapons will increase. As time goes on, however, and the political boundaries between nations are eroded due to advances in technology and fear of nuclear war, this time of unrest will pass. We will move away from nationalism. This does not necessarily mean world government, but rather an awareness of how interdependent all parts of the world have become.

The period of political unrest will be fueled by economic concerns, including the high cost of conventional energy, the unwillingness of energy producers to move to renewable resources, and unemployment. Unusual weather patterns will result in the occurence of expensive natural disasters. These economic concerns will gradually be alleviated by scientific discoveries, thereby paving the way for a new world order. We can expect a general conversion to solar energy. This will eliminate pollution from our environment, as well as having economic benefits. Hybrid plants which are rich in nutrients and capable of being grown in the oceans and deserts will be developed. This, along with improved food distribution systems, will make starvation a thing of the past. Our industries and institutions will become more concerned with the wellbeing of their employees, and most will adopt a flexible workweek designed to fit the employee's lifestyle.

The future will see a massive reorganization of the education system. These changes will be most pronounced in the industrialized nations. Initially, a large number of schools, colleges and universities will close, or they will be taken over for community use. Strikes and vandalism will be common in the schools that remain open. Many young people will be forced into the army. The value of a degree will decline. With the restructuring of society, however, a new type of school will emerge. These schools will emphasize practical skills and will be attended by people of all ages for varying lengths of time. Much of the educational process will be devoted to scientific research aimed at improving the quality of life.

Many new things will be invented or discovered. For example, we will have computerized maps built into the dashboards of our cars. Punching in your present location and your destination will cause these maps to light up with directions and information on weather conditions. Everyone will have a telephone which they carry with them at all times. The television will be used for at-home shopping.

Among the medical advances which we can expect will be the discovery of new pain killers with no side-effects. Vitamin therapy will be found to be highly effective in combating disease. Hypnosis will be used extensively to treat mental disorders. Lab-grown body parts will be used to replace unhealthy organs.

Thus, while a period of strife and dissension may occur, the future is far from bleak. We can look forward to major increases in our knowledge and understanding.

APPENDIX G

Experiment 2 Prose Test

The future cannot be known in advance. To try to predict the future, therefore, is not even very interesting unless one is a ____________________ seeking to entertain. Thus, it is less important to ask what the future will be than it is to ask what we can make the future become. Both good and bad outcomes can be predicted for the future. The bad outcomes can be considered a ____________________, warning us to examine our ____________________. The good outcomes can be considered a message of hope, an indication of what might be accomplished with the ____________________.

You may wonder if all, or any, of this forecast is believable. To convince yourself that these predictions are possible, take a look over your shoulder. Had someone told you, only a few years ago that we would soon have ____________________, videotape machines, digital watches, microwave ovens and ____________________, you might have had doubts. But it all happened. In light of the recent past, there is no reason to disbelieve the possibility of advances and changes in the future. With all this in mind, then, and if ____________________ continue, our future might look like the picture described below.

Politically, we are headed towards a time of unrest.

Many cities and countries will come under ____________________. Maintaining ____________________ will be a challenge as diplomatic ties between countries are disrupted. Production of ____________________ will increase. As time goes on, however, and the political boundaries between nations are eroded due to advances in ____________________ and fear of ____________________, this time of unrest will pass. We will move away from ____________________. This does not necessarily mean world government, but rather an awareness of how ____________________ all parts of the world have become.

The period of political unrest will be fueled by economic concerns, including the ____________________ of conventional energy, the unwillingness of energy producers to move to renewable resources, and ____________________. Unusual ____________________ will result in the occurence of expensive natural disasters. These economic concerns will gradually be alleviated by ____________________, thereby paving the way for a new world order. We can expect a general conversion to ____________________. This will eliminate ____________________ from our environment, as well as having economic benefits. Hybrid plants which are ____________________ and capable of being grown in the oceans and deserts will be developed. This, along with improved ____________________, will make starvation a thing of the past. Our industries and institutions will become more concerned with the ____________________ of their employees, and most will adopt a flexible workweek designed to ____________________.

The future will see a massive reorganization of the education system. These changes will be most pronounced in the ____________________. Initially, a large number of schools, colleges and universities will close, or they will be taken over for ____________________. Strikes and ____________________ will be common in the schools that remain open. Many young people will be forced into ____________________. The value of a degree will ____________________. With the restructuring of society, however, a new type of school will emerge. These schools will emphasize ____________________ and will be attended by ____________________ for varying lengths of time. Much of the educational process will be devoted to scientific research aimed at improving the ____________________.

Many new things will be invented or discovered. For example, we will have ____________________ built into the dashboards of our cars. Punching in your present location and your destination will cause these maps to light up with directions and information on ____________________. Everyone will have a ____________________ which they carry with them at all times. The television will be used for ____________________.

Among the medical advances which we can expect will be the discovery of new pain killers with ____________________. ____________________ will be found to be highly effective in combating disease. Hypnosis will be used extensively to treat ____________________. Lab-grown body parts will be used to ____________________.

Thus, while a period of strife and dissension may occur, the future is far from bleak. We can look forward to major increases in our knowledge and understanding.

APPENDIX H

Experiment 2 Sample Prose Passage and Test

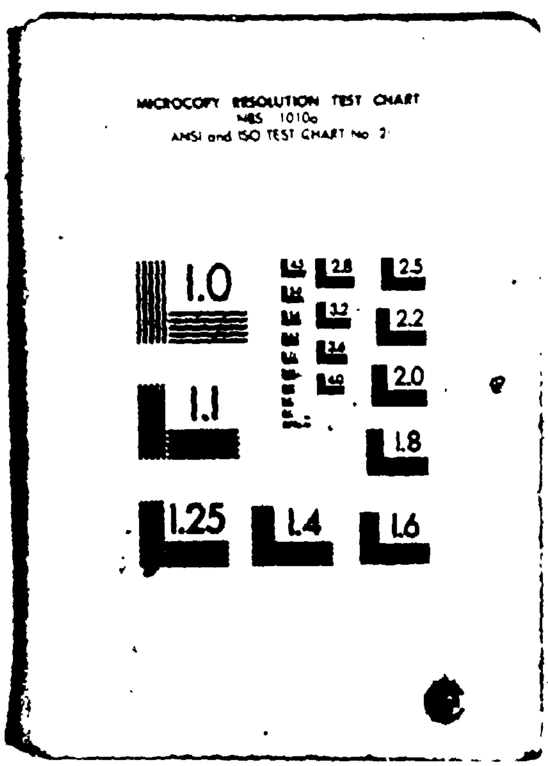
Many people would like to be able to foresee the future. However, history records only a handful of people who seem to have had this gift. Among these is Jacques Cazotte, a French writer of the <u>18th century</u>. Following a <u>dinner party</u> in Paris one evening, Cazotte was asked to predict the fates of each of the guests. His predictions were chilling - <u>six</u> of the guests and Cazotte himself would die violently in the next six years. Soon after this, <u>the French Revolution</u> began and by its end, all of Cazotte's prediction had come true.

Many people would like to be able to foresee the future. However, history records only a handful of people who seem to have had this gift. Among these is Jacques Cazotte, a French writer of the ____________________. Following a ____________________ in Paris one evening, Cazotte was asked to predict the fates of each of the guests. His predictions were chilling - ____________________ of the guests and Cazotte himself would die violently in the next six years. Soon after this, ____________________ began and by its end, all of Cazotte's prediction had come true.

APPENDIX I

Experiment 2 WAIS-R Vocabulary Test

3 OF/DE 3


MICROCOPY RESOLUTION TEST CHART
NBS 1010a
ANSI and ISO TEST CHART No. 2
1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

bed
ship
penny
winter
breakfast
repair
fabric
assemble
enormous
conceal
sentence
consume
regulate
terminate
commence
domestic
tranquil
ponder
designate
reluctant
obstruct
sanctuary
compassion
evasive
remorse
perimeter
generate
matchless
fortitude
tangible
plagarize
ominous
encumber
audacious
tirade

APPENDIX J

Experiment 2 WAIS-R Digit Span Test

5-8-2
6-4-3-9
4-2-7-3-1
6-1-9-4-7-3
5-9-1-7-4-2-8
5-8-1-9-2-6-4-7
2-7-5-8-6-2-5-8-4

6-9-4
7-2-8-6
7-5-8-3-6
3-9-2-4-8-7
4-1-7-9-3-8-6
3-8-2-9-5-1-7-4
7-1-3-9-4-2-5-6-8

APPENDIX K

Experiment 2 Memory Performance Questionaire

(Subjects provided a rating for each question using a five-point scale where 1 = totally positive response and 5 = totally negative response, then went on to verbally clarify their responses.)

Did you find the story interesting? Was your performance on the test influenced by how interesting you found the story to be?

Did you use any memory techniques or tricks while you were reading the story to help you remember the underlined parts? What were they?

Are there any techniques or tricks you would normally use when trying to remember the details of things that you've read that you did not use here? What are they?

Are you usually good at remembering the details of things that you've read?

Is it important to you to be good at remembering the details of things that you've read?

Was it important to you to do well on this test?

Did taking this test make you nervous?

Did how you felt while doing the test affect your performance? How?

Which do you think had the most control over your performance on the test: effort, ability, luck, how easy/difficult the task was? (Presented pairwise in randomized order.)

Could you train your memory so that you would do better if you took another test like this one? How?

Do you try to do things to keep your memory sharp? What do you do? Would your performance on this test have been different if you were 19 (33, 72) years old? (Only the categories to which the subject did not belong were presented.) Would it have been better or worse?

APPENDIX L

Experiment 2 Subject Data Questionaire

Subject Data Questionaire

Subject number: ___
Age: ____
Sex: ____

Education: grade school ____ years
high school ____ years
college/university ____ years

Residence: please record your living arrangements (e.g., with spouse, alone, boarder, etc.)

___

Relative to OTHER PEOPLE YOUR OWN AGE how would you rate

1) your health (circle one)

1	2	3	4	5
much worse than average	worse than average	average	better than average	much better than average

2) your memory (circle one)

1	2	3	4	5
much worse than average	worse than average	average	better than average	much better than average

Are you currently taking any medication that is known to affect memory functioning? ____ yes ____ no

Did you have any difficulty seeing the materials used in these tasks?
____ yes ____ no

Did you have any difficulty hearing the materials used in these tasks?
____ yes ____ no

Santi, A., Grossi, V. & Gibson, M. (1982). Differences in matching-to-sample performance with element and compound sample stimuli in pigeons. Learning and Motivation, 13, 240-256.

PRESENTATIONS

Brigham, M., Devolder, P. & Pressley, M. Cognitive monitoring across the adult years. Presented at the annual meeting of the American Psychological Association, Washington, D.C., August, 1986.

Weisman, R., Gibson, M. & Rochford, J. Advances in delayed sequence discrimination. Presented at the annual meeting of the Psychonomic Society, Philadelphia, November, 1981.

Subject Data Questionaire

Subject number: ___
Age: ___
Sex: ___

Education: grade school ____ years
high school ____ years
college/university ____ years

Residence: please record your living arrangements (e.g., with spouse, alone, boarder, etc.)

____

Relative to OTHER PEOPLE YOUR OWN AGE how would you rate

1) your health (circle one)

1	2	3	4	5
much worse than average	worse than average	average	better than average	much better than average

2) your memory (circle one)

1	2	3	4	5
much worse than average	worse than average	average	better than average	much better than average

Are you currently taking any medication that is known to affect memory functioning? ____ yes ____ no

Did you have any difficulty seeing the materials used in these tasks?
____ yes ____ no

Did you have any difficulty hearing the materials used in these tasks?
____ yes ____ no