Psychometric Properties of a New Metamemory Questionnaire for Older Adults

Angela K. Troyer¹ and Jill B. Rich¹,²

¹Baycrest Centre for Geriatric Care, Toronto, Canada. ²York University, Toronto, Canada.

Subjective memory ratings provide information that is distinct from objective memory performance, and there is a need for reliable and valid metamemory measures. The Multifactorial Memory Questionnaire (MMQ), developed to assess separate dimensions of memory ratings that are applicable to clinical assessment and intervention, includes scales of Contentment (i.e., affect regarding one’s memory), Ability (i.e., self-appraisal of one’s memory capabilities), and Strategy (i.e., reported frequency of memory strategy use). Among a group of 115 older adults, analyses revealed excellent content validity, factorial validity, test-retest and intratest reliability, convergent and discriminant construct validity, and independence from demographic variables. The psychometric strengths of the MMQ, together with descriptive statistics provided for healthy older adults, make this questionnaire useful in both clinical and research settings.

In a national survey of nearly 15,000 adults aged 55 and older, 15% reported having had trouble remembering things frequently during the past year (Cutler & Grams, 1988). Examination of a wide range of variables from that survey indicated that potential correlates, such as sex, age, educational attainment, and health factors, accounted for very little variance in the frequency of self-reported memory problems. As this large-scale survey suggests, self-report measures of memory function provide important information about aspects of subjective memory abilities in aging. Self-appraisals of memory add a unique dimension to the assessment of memory performance that is not captured by objective memory testing alone. Self-report tools allow the measurement of everyday memory problems rather than more artificial, laboratory-based memory problems. Consequently, such tools allow estimation of the clinical significance of a documented memory problem and can guide the development of intervention programs aimed at improving everyday memory function.

In addition, self-reports provide information about metamemory, or insight and awareness of an individual’s own memory functioning. Such insight is known to differ among memory-disordered patients of differing etiologies (Schacter, 1991). Moreover, even among healthy individuals, self-report measures often reveal inaccurate beliefs about memory and aging influenced by culturally based negative stereotypes (Cutler & Grams, 1988; Hertzog & Hultsch, 2000). This aspect of metamemory may be associated with a poor appraisal of one’s own memory efficacy, which in turn may lead to a self-propagating cycle of lowered expectations, reduced effort on everyday tasks involving memory, and, ultimately, impaired performance (i.e., a self-fulfilling prophecy of age-related memory reductions). Given the potential significance of metamemory, there is a need for well-developed self-report memory measures that are reliable, valid, and applicable to clinical and research settings.

Several self-report memory questionnaires have been developed for use with research participants and brain-injured clinical patients. The Metamemory in Adulthood Questionnaire (MIA; Dixon, Hultsch, & Hertzog, 1988) contains 108 items that describe memory issues and changes relevant to healthy aging. The seven subscales provide information about an individual’s knowledge of general memory processes and tasks, frequency of memory strategy use, self-rated memory ability, perceptions of memory stability over time, anxiety regarding memory, memory and achievement motivation, and locus of control in memory abilities. Good psychometric data have been presented for the factor structure, reliability, and validity of this test (Dixon et al., 1988; Dixon & Hultsch, 1983a, 1983b). Another metamemory scale, the Memory Functioning Questionnaire (MFQ; Gilewski, Zelinski, & Schaie, 1990), contains 64 items distributed across four subscales that reflect subjective appraisals of frequency of forgetting in different situations, the seriousness of the consequences of forgetting in these situations, comparison of present and past memory functioning, and frequency of memory strategy use. Psychometric evidence supports the factor structure, reliability, and validity of this test (Gilewski et al., 1990; Zelinski, Gilewski, & Anthony-Bergstone, 1990). Additional self-report instruments provide information primarily or exclusively about the frequency with which various memory mistakes occur. These include the Memory Assessment Clinics Self-Rating Scale (Crook & Larrabee, 1990, 1992), the Everyday Memory Questionnaire (Martin, 1986), and the Cognitive Failures Questionnaire (Broadbent, Cooper, FitzGerald, & Parkes, 1982). Recently, Lachman, Bandura, Weaver, and Elliott (1995) developed a metamemory questionnaire, the Memory Controllability Inventory, focusing on perceptions of control over one’s memory.

Despite the considerable number of self-report memory questionnaires described in the literature, there are several drawbacks associated with their use, especially in clini-
cal settings. First, because some of these questionnaires were developed primarily for research use, the items included do not necessarily reflect aspects of memory that are amenable to clinical memory intervention. For example, items concerning self-evaluation of past memory ability and those involving memory for very remote events would not reflect typical targets of clinical interventions. Rather, clinical interventions tend to focus on problems with recent memory (e.g., remembering new names) and on strategies relevant to everyday life (e.g., repetition, written aids). Additional items pertaining to contentment with one’s memory, which is relevant to clinical intervention, are not included in current measures. Second, several of the questionnaires include items that are not applicable to some individuals, such as public speaking, reading novels, driving, or working. It can be difficult to interpret the results of a questionnaire when not all items are answered. Third, some questionnaires, such as the MIA, contain a very large number of items. The considerable amount of time required to complete such a questionnaire could adversely affect compliance, especially in older or cognitively impaired individuals. Finally, although the MIA and MFQ address the multidimensional aspects of metamemory, several existing questionnaires do not. For example, unidimensional questionnaires, such as those that address only the frequency with which memory mistakes occur, do not provide as much information about self-reported memory as does a multifactorial questionnaire. Some aspects of metamemory are not addressed at all in the existing questionnaires; for example, although there are subscales assessing anxiety about memory (e.g., Dixon et al., 1988), there are no scales assessing other common emotional reactions to memory, such as frustration, irritation, unhappiness, or pleasure.

In this article, we present psychometric data on a new questionnaire, the Multifactorial Memory Questionnaire (MMQ), that assesses three dimensions of self-reported memory, including overall contentment or satisfaction with one’s own memory ability (MMQ-Contentment), perception of everyday memory ability (MMQ-Ability), and use of everyday memory strategies and aids (MMQ-Strategy). To increase clinical utility, we did not include items unlikely to be amenable to clinical interventions (e.g., comparison of present memory ability to performance when younger) or of little relevance to everyday functioning (e.g., use of laboratory-based mnemonics such as the peg-word system) in the questionnaire. To increase compliance, we made the scales as short as possible (i.e., requiring about 10 min to complete all three) without compromising either completeness of content or statistical reliability and validity. Thus, the MMQ is an improvement over other memory self-report instruments because it encompasses a number of features in combination (i.e., multidimensionality, clinical relevance, brevity, and ease of administration) that are found only partially in any single existing questionnaire. In addition, the Contentment scale of the MMQ includes items pertaining to aspects of memory-related affect, which are not typically included in existing questionnaires.

**METHODS**

**Participants**

Participants were 130 community-dwelling middle-aged and older adults recruited via newspaper advertisements, community lectures, and a database of research volunteers. Data were collected either at the beginning of a five-session, weekly educational program about memory and aging (n = 84), in the first of two research sessions conducted over the same time interval without the educational program (n = 24), or in a single research session (n = 22). To identify participants with possible memory impairment, we conducted cognitive screening using the modified version (Welsh, Breitner, & Magruder-Habib, 1993) of the Telephone Interview for Cognitive Status (TICS; Brandt, Spencer, & Folstein, 1988; n = 40) and/or a 10-item word-list memory task (n = 101). Participants were required to earn at least 30 out of 50 possible points on the modified TICS, a cutoff score recommended by Welsh and colleagues. In addition, participants were required to recall at least two items following a 30-s distraction interval on the word-list memory task. This cutoff score was derived from scores (M = 4.8, SD = 1.6) obtained by an independent sample of 24 older adults (Troyer & Craik, 1995) who were screened both for cognitive decline with the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) and for medical problems that could affect cognition. Fifteen participants were excluded from the sample because of low scores on the memory task alone (n = 11), low scores on both the modified TICS and the memory task (n = 1), or invalid data on the memory task because of failure to follow instructions (n = 3). This resulted in a total of 115 participants who met the inclusionary criteria. (Although excluded from the main analyses, the 12 participants with low memory were retained in a separate sample for one of the validity analyses, described in the Results section.) In the main sample, the mean age was 71.7 years (SD = 9.9, range = 40–91), and the mean level of education was 13.8 years (SD = 2.8, range = 3–21). A majority (i.e., 79%) of the participants were female.

**Memory Questionnaire Scales**

The original version of MMQ-Contentment contained 21 items addressing a variety of emotions and perceptions that participants may have about their current memory ability. Statements address positive emotions (e.g., confidence, satisfaction), negative emotions (e.g., embarrassment, irritation), and subjective ability ratings (e.g., comparison to peers, belief that one has a serious memory problem). Respondents rated their level of agreement with each statement on a 5-point scale (i.e., strongly agree, agree, undecided, disagree, and strongly disagree) according to how they felt over the past 2 weeks. For each item, 0 to 4 points were given on the basis of level of agreement such that higher scores indicated greater contentment.

MMQ-Ability contains 20 everyday memory situations, such as remembering appointments, names, and telephone numbers. Items are phrased as memory failures (e.g., forgetting an appointment) and were based in part on previously published questionnaires (Broadbent et al., 1982; Crook &
Larrabee, 1992; Dixon et al., 1988; Gilewski et al., 1990; Jennings & Hay, 1994). Respondents indicated the frequency with which each mistake occurred over the last 2 weeks on a 5-point scale (all the time, often, sometimes, rarely, never). For each item, 0 to 4 points were given on the basis of self-reported frequency, with higher scores indicating better subjective memory ability (i.e., fewer memory mistakes).

The original version of MMQ-Strategy contained 20 items describing different memory aids and strategies applicable to everyday memory tasks, such as writing appointments on a calendar and repeating information to oneself. Items were based in part on previously published questionnaires and interviews (Dixon et al., 1988; Harris, 1980; Intons-Peterson & Fournier, 1986; Lovelace & Twohig, 1990; Park, Smith, & Cavanaugh, 1990). Respondents indicated the frequency with which each strategy was used over the last 2 weeks using a 5-point scale (never, rarely, sometimes, often, all the time). For each item, 0 to 4 points were given on the basis of frequency of use, with higher scores indicating more frequent use of memory aids and strategies.

A few participants omitted some questionnaire items. For these scales, scores were prorated on the basis of the responses to completed items, and this prorated score was used in the analyses when possible. This method was used only if very few items (i.e., four or fewer) were omitted, which occurred on 26 out of 344 scales. There were no scales with more than four items omitted.

Other Questionnaires and Tests

In addition to the MMQ, some participants also completed previously validated self-report measures of memory (n = 52) or mood (n = 44). Memory questionnaires included the MIA (Dixon et al., 1988) and the MFQ (Gilewski et al., 1990). The entire questionnaires were administered, although we selected only certain subscales to compare to the MMQ. The MIA Anxiety and Change subscales and the MFQ Seriousness of Forgetting subscale were selected as measures of memory-related affect and perceived change in memory functioning expected to be related to MMQ-Contetntment. The MIA Capacity subscale and the MFQ General Frequency of Forgetting subscale were selected as measures of subjective memory ability expected to be similar to MMQ-Ability. The MIA Strategy subscale and the MFQ Mnemonics subscale were selected as measures of memory strategy use expected to be similar to MMQ-Strategy. Measures of general mood included the Geriatric Depression Scale (Yesavage et al., 1983) and the Anxiety scale from the Symptom Checklist-90-Revised (Derogatis, 1994).

Participants completed three objective memory tasks. Because of the group setting, it was not possible to use clinical memory tasks; thus, memory tasks that allowed simultaneous visual and oral presentation of information as well as a self-paced written test were created. The word-list memory test, as described previously, was used as a laboratory-based measure of recent memory. The score was the number of words correctly recalled after a 30-s delay. A name-memory task was used as an everyday measure of recent memory. For this test, two three-part name strings (i.e., first, middle, and last) were presented via slide projection for 10 s and were simultaneously read aloud. Free recall was tested following a 5-min delay filled with lecture. Points were awarded for both accuracy (i.e., correct name and spelling) and placement (i.e., correct sequencing of names within a string and no confounding of names between the two strings). As an everyday measure of prospective memory, participants were asked to make two telephone calls at specified times and to provide specific information (i.e., name and identification number). Points were awarded both for the promptness of the calls and for the information provided.

Participants also completed two tests of attention. As a speeded test of attention, a symbol-digit matching task was used. Each participant was given a sheet of paper with six symbol-digit pairs above rows of symbols alone. Participants were instructed to write the appropriate digit beneath each symbol as quickly as possible for 60 s. Scores were calculated as the number of correct symbol-digit matches. A digit span task was also created in which participants listened to strings of three to eight digits as they were read aloud and wrote down the digits immediately after presentation. The score was the longest span of digits written in correct sequence.

Procedure

Questionnaires and tests were administered in a fixed order to groups of 5 to 15 participants, as follows: MMQ Contetntment, Ability, and Strategy scales; then the objective memory tests (word-list memory, name memory, and prospective memory telephone task); then the tests of attention (symbol-digit followed by digit span). Participants who were tested on two occasions, only scores from the first session were used unless otherwise indicated. Although all 115 participants completed the three memory questionnaire scales at least once, some did not complete all of the other tests and questionnaires. The MIA, MFQ, and/or the mood measures were administered at the end of the session to the subsample of participants who completed them.

For the purpose of content validation, the 61 MMQ items were randomly ordered and were worded with similar terminology (i.e., all items began with the phrase “Do you . . .”). We recruited 12 memory experts from the psychology department and the research division at Baycrest Centre for Geriatric Care to rate questionnaire items. Each expert had a PhD in psychology; 6 were licensed clinicians who work with memory-disordered patients, and 6 were professionals who primarily conduct memory research. We asked raters to classify each item as belonging to one of three metamemory domains: emotions and perceptions of one’s own memory, everyday memory mistakes or problems, or memory strategies or aids.

Results

Content Validity

The criterion for content validity on the MMQ was 70% agreement among the 12 memory-expert raters. Agreement on the 61 items tended to be high, with 100% agreement obtained for 53 items, 92% agreement for 6 items, and 83% agreement for 1 item. Only 1 item failed to meet the 70%
agreement criterion. This item (i.e., “Do you feel that, if you try hard enough, you can find ways to help you remember better?”), which had been assigned a priori to MMQ-Contentment, was subsequently eliminated.

Factor Structure
We performed a principal components analysis with varimax rotation using the responses on the remaining 60 items from all 115 participants. We forced the solution to three components and, with few exceptions, each item loaded most highly onto the expected component. Eigenvalues of the three components ranged from 3.6 to 15.3. Eighteen of the 20 items from MMQ-Contentment loaded most highly onto a single component. The loadings from these 18 items ranged from .35 to .79, with a mean loading of .66. All 20 items from MMQ-Ability loaded most highly onto another component, with loadings ranging from .45 to .74 and a mean loading of .59. Nineteen of the 20 items from MMQ-Strategy loaded most highly onto a third component. The loadings from these items ranged from .27 to .70 with a mean loading of .50. The 3 items that did not load onto the expected components were eliminated from the subsequent analyses, for a total of 57 remaining items. These items and their component loadings are listed according to their respective scales in the Appendix. All loadings are positive because each item was scored such that higher scores indicated higher levels of contentment. In all, the three factors accounted for 40.4% of total score variance.

We retabulated scores for each participant using the retained items. Pearson correlations calculated between the three questionnaires indicated a strong relation only between MMQ-Contentment and MMQ-Ability, \( r(115) = .62, p < .001 \). MMQ-Strategy was not highly correlated with either MMQ-Contentment, \( r(114) = -.17, p = .078 \), or MMQ-Ability, \( r(114) = -.15, p = .104 \).

Reliability
We examined 4-week test-retest reliability using data from 24 participants who were tested on two occasions and did not receive any memory interventions. Correlations indicated highly reliable scores on MMQ-Contentment, \( r(24) = .93, p < .001 \); MMQ-Ability, \( r(24) = .86, p < .001 \); and MMQ-Strategy, \( r(24) = .88, p < .001 \). Unlike other correlation coefficients (which need to be squared), reliability coefficients are interpreted directly as the proportion of variance accounted for (Anastasi & Urbina, 1997; McNemar, 1969). Thus, 86–93% of the total test-retest score variance was due to true variance rather than to the effects of random daily changes in participant or environmental conditions.

We examined internal consistency using Cronbach’s alpha on data from all participants. These analyses indicated highly reliable scores on MMQ-Contentment (\( \alpha = .95 \)), MMQ-Ability (\( \alpha = .93 \)), and MMQ-Strategy (\( \alpha = .83 \)). Thus, at least 83–95% of the total within-test score variance was due to true score variance rather than to item content heterogeneity or poor item quality.

Construct Validity
We tested convergent validity by examining effect sizes (Cohen, 1988) based on correlations between scores obtained on the present memory questionnaires, other self-report memory and mood questionnaires, and objective memory tests. For MMQ-Contentment, correlations showed a large effect size with both the Anxiety and Change subscales of the MIA, \( r(50) = -.57 \) and .61, respectively, \( p < .001 \); medium effect sizes with the Seriousness of Forgetting subscale of the MFQ, \( r(51) = .45, p = .001 \), and with the Geriatric Depression Scale, \( r(43) = -.41, p = .007 \); and a small effect size with the Anxiety subscale of the Symptom Checklist-90-Revised, \( r(43) = -.27, p = .079 \). For MMQ-Ability, correlations showed large effect sizes with the Capacity subscale of the MIA, \( r(50) = .60, p < .001 \), and the General Frequency of Forgetting subscale of the MFQ, \( r(51) = .70, p < .001 \). For MMQ-Strategy, correlations showed large effect sizes with the Strategy subscale of the MIA, \( r(50) = .64, p < .001 \), and the Mnemonics subscale of the MFQ, \( r(51) = .66, p < .001 \).

As expected, the relation between self-reported memory function and performance on objective memory tests was weaker. Small to negligible effect sizes were obtained from correlations between MMQ-Ability and performance on the word-list memory test, \( r(101) = .10, p = .309 \); name memory test, \( r(101) = .09, p = .398 \); and the telephone task, \( r(94) = .14, p = .166 \).

To further examine the relation between self-reported memory and objective memory performance, we examined MMQ-Contentment and MMQ-Ability obtained by a group of participants with documented low memory performance (i.e., more than 2 standard deviations below the mean on the word-list memory task) who were excluded from the main sample, as described previously. Each individual from this low-memory sample was matched with 1 or 2 (whenever possible) individuals from the main sample according to sex, age (within 3 years), and education (within 2 years). The low-memory (\( n = 12 \)) and normal-memory (\( n = 20 \)) groups did not differ in age (\( M_s = 74.9 \) and 74.5 years), level of education (\( M_s = 12.6 \) and 12.8 years), or sex (67% and 65% female), \( p > .80 \). Scores on MMQ-Contentment were lower in the low-memory group (\( M = 30.7 \)) than the normal-memory group (\( M = 43.1 \)), \( t(30) = 3.10, p = .004 \). Similarly, scores on MMQ-Ability were lower in the low-memory group (\( M = 37.3 \)) than the normal-memory group (\( M = 46.2 \)), \( t(30) = 3.19, p = .003 \).

We examined discriminant validity by calculating correlation coefficients between the memory questionnaires and tests of attention. Scores on the memory questionnaires were unrelated (i.e., showed negligible effect sizes) to the symbol-digit task, \( r = -.02 \) to .00, \( p > .85 \), and the digit span task, \( r = .01 \) to .07, \( p > .50 \).

Relation to Demographic Variables and Mental Status
Scores on MMQ-Contentment, MMQ-Ability, and MMQ-Strategy were not related to age, \( r = -.13 \) to .07; education, \( r = .02 \) to .07; or sex, \( r = -.13 \) to .13, all \( p > .170 \). Moreover, MMQ scores were not significantly related to performance on the modified TICS, \( r = .03 \) to .20, \( p > .220 \).

Descriptive Statistics
Means, standard deviations, and percentiles based on performance by all 115 participants are provided in Table 1.
To examine the shape of the score distributions, we calculated skewness (S) and kurtosis (K) coefficients and tested them for significance using z scores. Scores from the three questionnaire scales did not show significant skewness, S = −0.43 to 0.21, z = −1.89 to 0.94, ps > .05. Scores from MMQ-Contentment and MMQ-Strategy did not show significant kurtosis, Ks = −0.80 to −0.03, z = −1.78 to −0.07, ps > .05, but scores from MMQ-Ability were significantly peaked, K = 1.69, z = 3.79, p < .001.

**DISCUSSION**

In this article we present a new self-report metamemory questionnaire developed specifically for eventual application in clinical settings. Examination of the psychometric properties of the three scales finds them to be psychometrically sound, which is a necessary precondition to demonstrations of their clinical utility. Overall, analyses reveal excellent content validity, factorial validity, reliability (both test-retest and intratreat), construct validity (both convergent and discriminant), and independence from demographic variables. Together, these findings provide a basis for use of the final version of the questionnaire in subsequent studies aimed at demonstrating its relevance to clinical applications.

On the basis of previous reports indicating the multidimensional nature of metamemory (e.g., Dixon & Hultsch, 1983a; Gilewski et al., 1990), the present scales were developed so that they could be administered separately for the assessment of separate dimensions of memory appraisal (i.e., Contentment, or feelings about one’s memory; Ability, or the subjective impression of one’s memory capabilities; and Strategy, or one’s reported frequency of use of various memory aids). The final assignment of items to each questionnaire scale is supported in several ways. First, there is excellent interrater agreement among 12 memory experts from both research and clinical settings, which demonstrates content validity for our initial distribution of items across the various scales.

Second, principal components analysis yields three meaningful factors, each of which is associated with high loadings by the items from one scale but not the other two. This pattern of loadings provides a measure of factorial validity that confirms our a priori division of items into scales representing different dimensions of metamemory. Correlational analysis reveals a strong, positive relation between MMQ-Contentment and MMQ-Ability. As one might expect intuitively, self-ratings of high competence regarding one’s memory abilities in performance-demanding situations are associated with high ratings of satisfaction with one’s memory. Conversely, individuals with feelings of low competence naturally are less satisfied with their memory. The lack of a relation between self-reported strategy use and ratings of either contentment or ability is also not surprising. Individuals with perceived memory deficiencies might use strategies as an intentional compensatory technique, in which case frequency of strategy use would correlate negatively with self-reported memory abilities and contentment. However, those who rate their memory ability positively and feel satisfied with it may achieve such memory success precisely because they use strategies frequently, in which case reported frequency of strategy use would correlate positively with ratings on the other two scales.

Reliability of the MMQ scales is demonstrated in two ways. Test-retest reliability in a subgroup of participants who completed the questionnaires in two sessions 4 weeks apart (with no intervention) is excellent for all three scales. Moreover, the intratreat reliability of each of the questionnaires is acceptable, as indicated by measures of internal consistency.

Construct validity is also assessed in several ways. As expected, convergent validity is demonstrated by medium-to-large effect sizes of correlations between the present instruments and their corresponding scales on other multidimensional metamemory questionnaires and subjective mood scales. Specifically, feelings about one’s memory (contentment vs dissatisfaction) are related to ratings of anxiety (on the MIA and Symptom Checklist-90-Revised), perceived change in memory functioning (on the MIA), feelings of seriousness of forgetting (on the MFQ), and endorsement of depressive symptoms (on the Geriatric Depression Scale).

The fact that a much larger effect size is observed for the correlation between MMQ-Contentment and the Anxiety scale of the MIA than for its correlation with depressive symptomatology or generalized feelings of anxiety further validates the construct of feelings about one’s memory as a specific dimension in a multidimensional model of metacognition. Similarly, scores on MMQ-Ability and MMQ-Strategy, which reflect one’s perceived memory capabilities and reported frequency of use of memory aids, respectively, are related to scores on corresponding scales of the MIA (Capacity and Strategy) and MFQ (General Frequency of Forgetting and Mnemonics).

Both the MIA and MFQ are well-established multidimensional questionnaires with good psychometric properties (Dixon et al., 1988; Gilewski et al., 1990), and both have been used frequently in studies of cognitive aging. Thus, it is a testament to the construct validity of the three scales developed and reported in this study that scores on them are significantly related to scores on the corresponding scales of these other instruments. At the same time, it is gratifying to note that the largest correlation observed (i.e., between
MMQ-Capacity and MFQ General Frequency of Forgetting) was .70, which accounts for only half the variance in scores on the two scales. That scores on the present scales are significantly but not excessively related to scores on other instrument subscales indicates that the new questionnaire scales are similar to existing questionnaires yet provide additional information about unique aspects of memory self-appraisals.

Among the main sample of healthy, elderly individuals who participated in this study, questionnaire ratings show very low correlations with objective memory performance across both traditional experimental tasks, such as word-list memory, and everyday memory tasks, such as name memory and a prospective memory telephone task. Other studies have also found self-reports of memory ability to be highly reliable but not very accurate in predicting actual memory performance (e.g., Little, Williams, & Long, 1986; Rabbitt & Abson, 1990; see Herrmann, 1982, 1990, for reviews). Others (Hertzog, Saylor, Fleece, & Dixon, 1994) have found that the relationship between metamemory and objective memory performance is mediated by a number of variables, such as age, type of memory task, and experience with the task. Differences between memory self-appraisals and objective memory performance may arise from several sources. For social reasons, memory failures are not necessarily pointed out to people by their friends; on the contrary, reports by others may be deliberately misrepresented in order to flatter—or to intimidate.

If people cannot rely on their friends to give them truthful feedback about their performance, they likely do no better to rely on themselves either. When reporting all the words one can recall on a memory test, for example, one may not be aware of those that are not recalled. These factors may contribute to inaccuracies in evaluating one’s own memory performance. In addition, to the extent that variation in self-reported memory failures is driven by general beliefs about memory and aging rather than monitoring of specific incidents of faulty memory, self-appraisal of memory ability would be uncorrelated with objective memory performance. Posing questionnaire items as failures (as opposed to an “agree-disagree” format) might elicit implicit theories or stereotypes about age-related memory declines in general. If so, the format of the MMQ may have contributed to the lack of relation between reported and actual memory abilities.

Another reason for the poor relationship between memory appraisals and performance is that most memory tests are not predictive of everyday memory performance. People may use compensatory strategies (such as written aids) in everyday life that are not used in laboratory testing. They may also have more practice with the types of everyday memory situations presented on questionnaires than with those in laboratory tasks. The name-memory and telephone tasks used in the present study appear to be more relevant to everyday memory than word-list learning in the sense that learning names and making telephone calls are everyday tasks; nevertheless, they were contrived with standardized administration and scoring procedures and likely do not tap the same underlying construct as learning the name of someone to whom one has been introduced recently or remembering to make a phone call that is meaningful in one’s everyday life. Moreover, self-report questionnaires are not intended to be used as screening instruments for actual memory problems; instead, researchers should use the two types of tools in a thorough assessment of memory to provide information about both subjective memory beliefs and objective memory performance.

The lack of relation between the MMQ scales and the telephone task, which is a prospective memory task, is consistent with the results of several other studies indicating that subjective memory assessment is unrelated (or is negatively related) to performance on prospective memory tasks (Dobbs & Rule, 1987; Sunderland, Watts, Baddeley, & Harris, 1986; Zelinski et al., 1990). Some studies have found modest correlations between self-reported memory and objective memory performance. For example, Dixon and Hultsch (1983b) found that metamemory scores on the Task and Strategy subscales of the MIA were the best overall predictors of memory for text in adulthood. However, consistent with the present results, Dixon and Hultsch found no relation among older adults between memory performance and any of the three MIA scales (i.e., Anxiety, Capacity, and Strategy) that correspond to the MMQ scales. Conversely, there does seem to be a good correspondence between self-reports and memory performance when the self-reported aspects of metacognition or metamemory are directly related to the outcome measure. For example, self-reported strategy use on a recall test is positively related to recall performance on that test (Hertzog, McGuire, & Lineweaver, 1998). Similarly, self-reported memory efficacy in both younger and older adults, as measured by predicted memory performance on a specific task, is related to recall performance on that task, thereby demonstrating age-invariant accuracy in metacognitive predictions (Connor, Dunlosky, & Hertzog, 1997). Thus, it appears that the relation between metamemory and objective memory performance differs depending on whether the self-reported memory reflections refer to general aspects of memory (i.e., abilities or strategy use) or, instead, are directed to the specific test used as the outcome measure of memory performance.

Although memory self-appraisals are unrelated to memory performance in our main sample of healthy, elderly participants, scores on two of the three scales do differ as a function of objective memory performance when participants with documented poor memory abilities (who were excluded from the main analyses) are included. Specifically, participants with poorer memory have lower scores on both the MMQ-Contentment and MMQ-Ability scales relative to the main sample. The fact that the scales discriminate between older adults with low versus normal objective memory performance suggests that the failure to find a relation between questionnaire ratings and memory performance in this and other studies may be due in part to the standard practice of excluding individuals with memory impairments, which results in a restricted range of memory ability.

We assessed discriminant validity by comparing scores on the memory questionnaires to performance on attention measures (i.e., symbol-digit and digit span). As expected, memory appraisals and attention are unrelated, thereby demonstrating another aspect of construct validity. In addi-
tion, MMQ scores are unrelated to demographic variables and global mental status among the older adults participating in this study. It is possible, however, that a significant correlation between age and MMQ would have emerged if younger adults were included, thereby increasing the age range.

In sum, evidence for the reliability and validity of the MMQ is obtained among a sample of 115 healthy older adults. The relatively short length of the MMQ (i.e., 57 items covering three distinct scales) and the intended applicability to clinical intervention should enhance the usefulness of the questionnaires in clinical and research settings. Future studies should be conducted to determine whether the factor structure observed here remains constant across different samples (e.g., young individuals, cognitively impaired elderly persons) and longer time frames.

Acknowledgments

We thank Kimberly Kane, Katy Kamkar, Cara Mirabelli, and Samantha Dworsky for assistance in data collection and management and Sabine Doebel for assistance in data scoring.

Address correspondence to Angela Troyer, PhD, Department of Psychology, Baycrest Centre for Geriatric Care, 3560 Bathurst Street, Toronto, Ontario, Canada M6A 2E1. E-mail: atroyer@baycrest.org

References


Appendix

The items included in the final versions of the questionnaires are as follows. Component loadings are presented in parentheses.

**Metamemory Questionnaire-Contentment**

1. I am generally pleased with my memory ability. (.79)
2. There is something seriously wrong with my memory. (.68)
3. If something is important, I will probably remember it. (.58)
4. When I forget something, I fear that I may have a serious memory problem, like Alzheimer's disease. (.35)
5. My memory is worse than most other people my age. (.59)
6. I have confidence in my ability to remember things. (.74)
7. I feel unhappy when I think about my memory ability. (.70)
8. I worry that others will notice that my memory is not very good. (.59)
9. When I have trouble remembering something, I'm not too hard on myself. (.52)
10. I am concerned about my memory. (.76)
11. My memory is really going downhill lately. (.71)
12. I am generally satisfied with my memory ability. (.79)
13. I don't get upset when I have trouble remembering something. (.58)
14. I worry that I will forget something important. (.64)
15. I am embarrassed about my memory ability. (.67)
16. I get annoyed or irritated with myself when I am forgetful. (.68)
17. My memory is good for my age. (.78)
18. I worry about my memory ability. (.78)

**Metamemory Questionnaire-Ability**

1. How often do you forget to pay a bill on time? (.66)
2. How often do you misplace something you use daily, like your keys or glasses? (.52)
3. How often do you have trouble remembering a telephone number you just looked up? (.55)
4. How often do you not recall the name of someone you just met? (.54)
5. How often do you leave something behind when you meant to bring it with you? (.60)
6. How often do you forget an appointment? (.53)
7. How often do you forget what you were just about to do; for example, walk into a room and forget what you went there to do? (.67)
8. How often do you forget to run an errand? (.64)
9. How often do you have difficulty coming up with a specific word that you want? (.67)
10. How often do you have trouble remembering details from a newspaper or magazine article you read earlier that day? (.65)
11. How often do you forget to take medication? (.54)
12. How often do you not recall the name of someone you have known for some time? (.64)
13. How often do you forget to pass on a message? (.63)
14. How often do you forget what you were going to say in conversation? (.74)
15. How often do you forget a birthday or anniversary that you used to know well? (.52)
16. How often do you forget a telephone number you use frequently? (.56)
17. How often do you retell a story or joke to the same person because you forgot that you had already told him or her? (.45)
18. How often do you misplace something that you put away a few days ago? (.62)
19. How often do you forget to buy something you intended to buy? (.60)
20. How often do you forget details about a recent conversation? (.54)

**Metamemory Questionnaire-Strategies**

1. How often do you use a timer or alarm to remind you when to do something? (.27)
2. How often do you ask someone to help you remember something or to remind you to do something? (.30)
3. How often do you create a rhyme out of what you want to remember? (.45)
4. How often do you create a visual image of something you want to remember, like a name and a face? (.55)
5. How often do you write things on a calendar, such as appointments or things you need to do? (.40)
6. How often do you go through the alphabet one letter at a time to see if it sparks a memory for a name or word? (.47)
7. How often do you organize information you want to remember; for example, organize your grocery list according to food groups? (.58)
8. How often do you say something out loud in order to remember it, such as a telephone number you just looked up? (.64)
9. How often do you use a routine to remember important things, like checking that you have your wallet and keys when you leave home? (.38)
10. How often do you make a list, such as a grocery list or a list of things to do? (.41)
11. How often do you mentally elaborate on something you want to remember; for example, focus on a lot of the details? (.52)
12. How often do you put something in a prominent place to remind you to do something, like putting your umbrella by the front door so that you will remember to take it with you? (.53)
13. How often do you repeat something to yourself at increasingly longer and longer intervals so that you will remember it? (.70)
14. How often do you create a story to link together information you want to remember? (.63)
15. How often do you write down in a notebook things that you want to remember? (.53)
16. How often do you create an acronym out of the first letters in a list of things to remember, such as carrots, apples, and bread (cab)? (.39)
17. How often do you intentionally concentrate hard on something so that you will remember it? (.70)
18. How often do you write a note or reminder for yourself (other than on a calendar or in a notebook)? (.39)
19. How often do you mentally retrace your steps in order to remember something, such as the location of a misplaced item? (.64)