DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF BENEFICIARY KNOWLEDGE INDICES FROM THE MEDICARE CURRENT BENEFICIARY SURVEY

DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF BENEFICIARY KNOWLEDGE INDICES FROM THE MEDICARE CURRENT BENEFICIARY SURVEY

by Carla Bann, Ph.D., and
Nancy Berkman, Ph.D.

Scientific Reviewer: Arthur Bonito, Ph.D.

Federal Project Officer: Sherry Terrell, Ph.D.

RTI

CMS Contract No. 500-00-0024/003

October 2002

This project was funded by the Centers for Medicare & Medicaid Services under contract no. 500-00-0024. The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of the Centers for Medicare & Medicaid Services. RTI assumes responsibility for the accuracy and completeness of the information contained in this report.

CONTENTS

Symbols	
*	$p < 0.05$
**	$p < 0.01$
***	$p < 0.001$

EXECUTIVE SUMMARY

The Balanced Budget Act of 1997 significantly increased the number and range of health plan options potentially available to Medicare beneficiaries. To inform Medicare beneficiaries about these choices, the law provided for specific information dissemination activities that the Centers for Medicare & Medicaid Services implemented as a part of its National Medicare Education Program (NMEP). The NMEP is the largest and most comprehensive information program ever undertaken by the Medicare program. One important feature of the NMEP is the *Medicare & You* handbook which was first nationally distributed in 1999.

A goal of RTI's project, *Analysis of Medicare Beneficiary Knowledge Data Using the Medicare Current Beneficiary Survey: Phase Three*, is to evaluate the impact of the NMEP on beneficiaries' knowledge of the Medicare program. As a part of this project, RTI is comparing knowledge before and after the first national distribution of the *Medicare & You* handbook using data from the 1998 to 2000 administrations of the Medicare Current Beneficiary Survey (MCBS). To assess beneficiary knowledge, we developed two potential measures of knowledge using data from the Beneficiary Knowledge supplemental rounds 23, 26, and 29 and Beneficiary Needs supplemental rounds 24, 27, and 30 of the MCBS Access to Care files. The first measure, the perceived knowledge index, includes five questions that ask beneficiaries to subjectively rate how much they know about a particular topic related to Medicare. The other measure is a nine-item quiz that requires participants to respond to a series of true/false questions.

We evaluated the psychometric properties of the knowledge measures by conducting item-level analyses and scale-level descriptive statistics, assessing internal consistency reliability, and establishing construct validity. The item-level analyses of the perceived knowledge items suggest that these items appropriately do not exhibit floor or ceiling effects. Comparisons of the nine-item quiz scores by beneficiaries' educational achievement suggest that the difficulty level of this quiz is equivalent to a high school level. The reliability analyses indicate that both the perceived knowledge index and the nine-item quiz reached acceptable levels of internal consistency reliability.

We used three different approaches to assess construct validity of the knowledge indices. First, we evaluated the relationship between the knowledge indices and a global knowledge item that asks beneficiaries to rate how much they know about the Medicare program. A strong relationship between a knowledge index and the global knowledge item would provide support for the construct validity of the index. Second, we compared the mean knowledge index scores for groups previously shown to differ in knowledge. Finally, we compared the mean knowledge scores for groups that we hypothesized would have different levels of knowledge.

Based solely on the quantitative results, the five-item perceived knowledge index seems to have the best psychometric properties. This index had the highest Cronbach's alpha, demonstrated a strong, monotonic relationship with the global knowledge question, and showed significant differences in index scores between respondents who were expected to have differing knowledge levels. However, other criteria, such as the content of the items, must be considered when selecting the most appropriate knowledge measure. The perceived knowledge index relies on beneficiaries to be the sole judge of their knowledge. Individuals' subjective ratings may be influenced by factors other than their knowledge, such as confidence in decision making or

satisfaction with information received. Therefore, the perceived knowledge index may not provide the most accurate assessment of true knowledge level. A more accurate measure of knowledge would require respondents to actually demonstrate their knowledge.

On the basis of both content considerations and the psychometric analysis results, the nine-item quiz appears to be the most useful measure of beneficiary knowledge overall. The quiz requires beneficiaries to demonstrate their knowledge by correctly answering true/false questions rather than merely stating that they know everything they need to know. In addition, the quiz had good variability in scores, reached an acceptable level of internal consistency reliability, and performed well on the validity analyses.

Possible modifications to the nine-item quiz depend on the intended use of the measure. If the goal of the quiz is to obtain the most precise estimate of beneficiaries' knowledge possible, then the quiz should contain questions that cover all difficulty levels. The education-level analyses suggest that the items on the nine-item quiz do in fact cover all education levels. However, the quiz contains only one item at the high school graduate level whereas it has four items at the college level. Because more respondents report being a high school graduate than any other educational level, it may be helpful to add more items targeted at this group.

In contrast, if the purpose of the quiz is to determine whether Medicare beneficiaries reach a certain proficiency level, then the primary emphasis should be on the content of the quiz items. The items should cover the entire range of information that beneficiaries need to know in order to make informed decisions. Less emphasis should be placed on the difficulty of the items. In this case, a limitation of the nine-item quiz is that four of its nine questions deal with managed care plans which restricts the range of knowledge the quiz can measure. Other questions could be added to improve the comprehensiveness of the quiz. In fact, as a part of RTI's *Questionnaire Development and Cognitive Testing Using Item Response Theory* project, several new knowledge questions were developed that address a variety of issues, including beneficiary rights and health plan decision making (Uhrig et al., 2001), and could be used to expand the quiz.

SECTION 1
INTRODUCTION

The Balanced Budget Act (BBA) of 1997 significantly increased the number and range of health plan options potentially available to Medicare beneficiaries. To inform Medicare beneficiaries about these choices, the law provided for specific information dissemination activities that the Centers for Medicare & Medicaid Services implemented as a part of its National Medicare Education Program (NMEP). The NMEP is the largest and most comprehensive information program ever undertaken by the Medicare program. Some important features of the NMEP include the *Medicare & You* handbook, the Medicare website (www.medicare.gov), and the helpline (1-800-MEDICARE).

A goal of RTI's project, *Analysis of Medicare Beneficiary Knowledge Data Using the Medicare Current Beneficiary Survey: Phase Three*, is to evaluate the impact of the NMEP on beneficiaries' knowledge of the Medicare program and available health plan options. The Medicare Current Beneficiary Survey (MCBS) provides a useful source of data with which to assess levels of Medicare beneficiary program knowledge. In the MCBS, a large national probability sample of 14,000 or more Medicare beneficiaries are interviewed in a rotating panel design every 4 months for up to 4 years. Each year, approximately one-quarter of the sample is rotated out of the survey and replaced with new sample members. Therefore, 25 percent of each annual MCBS data set represents a cross section of the Medicare population enrolled in the program continuously since January 1st of that year and 75 percent represents a longitudinal beneficiary panel.

As a part of the NMEP, the *Medicare & You* (HCFA, 2000) handbook was first distributed nationally in the fall of 1999. To evaluate the impact of the handbook, this project will assess changes in beneficiary knowledge from immediately before to immediately after national distribution of the handbook. A first step in determining the level of Medicare program knowledge among Medicare beneficiaries is to develop an index to measure knowledge of the Medicare program. This report describes the development of knowledge measures from questions included in the Beneficiary Knowledge (BK) supplemental rounds 23, 26, and 29 and Beneficiary Needs (BN) supplemental rounds 24, 27, and 30 of the MCBS Access to Care files respectively, which were administered before and after national distribution of the first *Medicare & You* (HCFA, 2000) handbook. ***Table A.1*** in the appendix outlines the timelines for the administration of the supplemental rounds of the MCBS and the national distributions of the *Medicare & You* handbook. In addition, this report evaluates the psychometric properties of each knowledge measure, including internal consistency reliability and construct validity.

SECTION 2
DEVELOPMENT OF KNOWLEDGE MEASURES

We developed two knowledge measures using questions from the 1998 to 2000 Access to Care files and related supplemental rounds of the MCBS. The knowledge measures and the items comprising them are listed in ***Table A.2*** in ***Appendix A***. The first knowledge measure, the Perceived Knowledge Index, includes five items that were administered during BN Rounds 24, 27, and 30.[1] These questions ask beneficiaries how much they feel they know about five topics: (a) what services Medicare covers, (b) how much they have to pay for medical services covered by Medicare, (c) supplemental or Medigap insurance, (d) the availability and benefits of Medicare HMOs, and (e) choosing or finding a doctor or other health care provider.

The second measure is a nine-item quiz including items from BK Rounds 23, 26, and 29 and BN Rounds 24 and 27.[2] The quiz is comprised of true/false questions concerning Medicare benefits and managed care plans. In some years, additional quiz items were also included, however, to allow for comparisons from 1998 to 2000, only the nine items included in all three years were used to compute the quiz scores in this report.

2.1 Scoring Algorithms

2.1.1 Perceived Knowledge Index

The perceived knowledge index was created by reverse scoring each of the five response categories across the items. For example, in the original coding of the variables, knowing "just about everything you need to know" was coded as "1," and knowing "almost none of what you need to know" was coded as "5." The former response was recoded as "5," while the latter was recoded as "1." Responses for "2" and "4" were also switched. "Don't know" and "Refused" responses were recoded as missing. Then the recoded responses for all five questions were summed to compute the overall perceived knowledge index score.

For example, if a respondent answered "3," "4," "2," "1," and "5," respectively, to the original questions, the questions would be reverse coded and the knowledge score would be calculated as follows:

$$3 + 2 + 4 + 5 + 1 = 15.$$

Higher scores on this index reflect beneficiaries reporting that they knew more of what they needed to know about the five different topics. Scores produced by this system may range from 5 to 25, thus providing greater variability in scores and more power to discriminate among beneficiaries than would be obtained with only true/false (two response) questions.

---

[1]The perceived knowledge index was referred to as the "know-all-need-to-know index" in Bann et al. (2000).

[2]During the 2000 MCBS, the items were all administered in Round BK 29. In the other two years, seven of the questions were administered in the BK round and two were included in the BN round.

We excluded some potential survey questions from the index. One item regarding the Medicaid program was not included because it was asked of Medicaid recipients only. Another item that asks whether respondents know what they need to know to stay healthy was excluded from the calculation of the scale because the question does not address an insurance benefit or option and therefore appeared conceptually unrelated to the other five items.

We calculated the perceived knowledge index by imputing values for missing data rather than eliminating a respondent's answers. The mean of the remaining items was substituted for the missing item values (Chapman, 1976). Imputation was used only for respondents who answered at least half of the items (three of the five items). Individuals missing responses to more than half of the items were assigned a value of missing for the index.

2.1.2 Nine-Item Quiz

Correct responses to each of the nine quiz questions were coded as "1," while incorrect or "don't know" responses were coded as "0." The recoded responses were then summed to create the quiz scores. For example, a respondent who had three correct responses and six incorrect responses received a score of "3," as shown:

$$0 + 1 + 1 + 0 + 1 + 0 + 0 + 0 + 0 = 3.$$

An advantage of this scoring method is that it yields scores with meaningful interpretations, specifically, the number of questions to which the respondent knew the correct answer. Scores range from "0" to "9," with higher scores indicating greater knowledge of the Medicare program.

For the quiz, respondents who refused or were missing data on all of the quiz items were coded as missing for the overall quiz score. In this sample, all of the remaining respondents provided answers to all of the quiz questions or gave a response of "don't know." As mentioned earlier, "don't know" responses were coded as incorrect. Therefore, in contrast to the perceived knowledge index, we did not need to impute missing values in order to calculate scores for the nine-item quiz.

SECTION 3
PSYCHOMETRIC ANALYSIS METHODS

To evaluate the psychometric properties of the knowledge measures, we conducted item-level and scale-level analyses. The item-level analyses examined the psychometric properties of each individual item included in the scale, while the scale-level analyses evaluated the scale as a whole.

3.1 Item-Level Analyses

For both knowledge indices, we computed the correlations between each item on the index and the total index score. Item–total score correlations provide information about the contribution of the item to the reliability of the scale. Ideally, items should have item–total correlations of at least 0.30. Because the item itself is included in the calculation of the index score, the item–total score correlations may be inflated. To avoid possible inflation, the item–total score correlations in this report were corrected for overlap (Howard and Forehand, 1962).

For the perceived knowledge items, we examined the distribution of responses for possible floor or ceiling effects (i.e., scale compression). Ideally, the item should have variability in responses, indicating that participants are utilizing all five response options. For the true/false quiz questions, we examined the percentage of correct responses which provides an indicator of the difficulty of the questions. In addition, we utilized information on beneficiaries' highest school grade completed to try to match the difficulty level of an item to an education level. If possible, a knowledge index should contain items with a wide range of difficulty levels to enable it to discriminate among respondents with a variety of knowledge levels.

3.2 Scale-Level Analyses

This section describes the analyses used to assess the psychometric properties of the knowledge indices. For both scales, three sets of analyses were computed: (1) descriptive statistics, (2) reliability analyses, and (3) validity analyses.

3.2.1 Descriptive Statistics

For each scale, we calculated descriptive statistics to determine the most representative scale scores and to examine the distribution of scores. For each scale, we computed three measures of central tendency: mean, median, and mode. We also calculated standard deviations to examine variability in scores. A lack of variability can compromise the validity of the scale scores.

3.2.2 Reliability

The internal consistency reliability of the scales was estimated using Cronbach's alpha coefficient (Cronbach, 1951). Internal consistency measures the degree to which items on a scale are related to each other and therefore appear to be measuring the same construct. One common rule of thumb is to require alpha coefficients to be 0.70 and above in order for the index score to be considered reliable for use in group-level statistical analyses (Guilford, 1956; Nunnally, 1978). The coefficient alphas for each index were also calculated separately for

several subgroups defined by insurance and service utilization variables. These analyses helped determine whether the reliability of the index remained consistent across different groups.

Test–retest reliability was not assessed because this type of reliability is used to measure the stability of a scale over time and is usually assessed over a short period. The time between the administration of each wave of the knowledge supplement MCBS is relatively long (i.e., 1 year), during which time several factors (e.g., experience with the program, use of services) could affect a respondent's level of knowledge. Calculating the test–retest reliability using assessments administered so far apart would greatly underestimate the reliability of the scales. Therefore, test–retest reliability is not an appropriate type of reliability assessment for these knowledge measures.

3.2.3 Validity

Several different approaches may be used to establish the validity of a new scale. Ideally, the new scale would be shown to be highly related to a well-established and validated scale measuring the same construct, often called a "gold standard." Unfortunately, a gold standard for measuring Medicare beneficiary knowledge does not exist. Therefore, we assessed validity using three alternative approaches. These methods may not individually provide evidence as strong as that of a comparison with a gold standard, but together they may provide at least preliminary evidence to support the validity of the knowledge indices.

For the first set of validity analyses, we examined the relationship between the knowledge indices and another measure of the same construct. If the indices are highly related to the other knowledge measure, they will have demonstrated evidence of construct validity. The MCBS includes a knowledge question that asks respondents how much they feel they know about the Medicare program; the question is followed by a 5-point rating scale ranging from "almost none of what you need to know" to "just about everything you need to know." This question, which we refer to as the global perceived knowledge question, obtains the respondents' own perceptions of their Medicare-related knowledge and would be expected to be related to the level of knowledge indicated by their scores on the knowledge indices. A strong monotonic relationship between this question and the knowledge indices would support the validity of the indices.

Next we conducted analyses to determine if the knowledge scale scores discriminated among groups of Medicare beneficiaries who have previously been shown to differ in their knowledge of Medicare. This approach is sometimes referred to as known-groups comparisons. We based our expectations of differing levels of knowledge among beneficiaries on previous research. For example, factors related to socioeconomic status are often predictive of levels of insurance knowledge. Several studies report that respondents with more education have higher levels of insurance knowledge (Lambert, 1980; Marquis, 1983; McCall, Rice, and Sangl, 1986; Hibbard et al., 1998; McCormack et al., 2002). Higher knowledge levels have also been associated with higher incomes (Lambert, 1980; Marquis, 1983; McCall, Rice, and Sangl, 1986; Hibbard et al., 1998) and having a supplemental insurance plan (Cafferata, 1984). Other researchers have found that, among older adults, those who are younger have more insurance-related knowledge (Lambert, 1980; Cafferata, 1984).

Based on this research, we expected that the following groups of beneficiaries would have higher levels of knowledge about the Medicare program: (1) beneficiaries with more education, (2) beneficiaries with higher incomes, and (3) beneficiaries with supplemental insurance. We also expected that among beneficiaries who are eligible for Medicare because of their age, those who are younger would have more knowledge. In this report, evidence for construct validity of a particular scale is provided if the results of the analyses on that scale showed these expected patterns.

Finally, we compared groups of respondents that we hypothesized would differ in knowledge. For this report, we called these exploratory comparisons. We expected that beneficiaries enrolled in managed care during the past year would have more knowledge than those not enrolled in managed care during the past year because they are required to make more choices regarding their insurance arrangements than beneficiaries using fee-for-service. Because four of the items on the nine-item quiz address issues involving managed care, we expected the greatest effect to be present for this measure.

In addition, we hypothesized that beneficiaries who have more experience with the Medicare program would have higher levels of knowledge. For these validity analyses, we used level of service utilization as an approximation of experience with the Medicare program. Two types of service utilization during the past year were included: (1) institutional utilization and (2) Part B utilization. Institutional utilization includes hospice, home health agency, or skilled nursing facilities and inpatient or outpatient hospital visits. Part B utilization includes professional services and durable medical equipment. The amounts of allowable and reimbursed charges were also used as indicators of experience with the Medicare system.

Some prior research supports this hypothesis. Cafferata (1984) found that among a subsample of older adults with private insurance, service utilization was positively associated with knowledge. McCormack and colleagues (2002) found that hospitalization and number of doctor visits were positively related to beneficiary knowledge of the Medicare system. Results from the national *Medicare & You* evaluation also suggest a positive relationship between beneficiary knowledge and number of doctor visits (McCormack et al., 2001).

The following background and experience variables were used for both the known-groups comparisons and the exploratory comparisons:

- income (under $25,000 or $25,000 or more),
- age (65 to 75 years old, or over 75 years old),
- educational achievement (8th grade or less, more than 8th grade but no college, or college),
- enrollment in managed care during the past year (enrolled or not enrolled),
- private supplemental insurance (have supplemental insurance or do not have supplemental insurance),

- institutional utilization (some utilization or no utilization),
- Part B utilization (some utilization or no utilization),
- total reimbursed dollars ($0; $1 to $499; $500 to $4,999; or $5,000 or more), and
- covered institutional charges ($0; $1 to $499; $500 to $4,999; or $5,000 or more).

Because complete information on service utilization was available only for respondents who were not enrolled in an HMO, our analyses of these variables included only individuals who were not enrolled in managed care during the year before the survey data were collected.

SECTION 4
RESULTS

This section describes the results of the psychometric analyses of the perceived knowledge index and the nine-item quiz. For both indices item-level and scale-level analyses were computed. As a part of the scale-level analyses, the reliability and construct validity of the indices were evaluated.

Data from the 1998 to 2000 Access to Care BK supplemental rounds 23, 26, and 29 and BN supplemental rounds 24, 27, and 30 of the MCBS were used for the psychometric analyses. During 1998 and 2000, the knowledge questions were administered to all respondents; however, in 1999 the questions comprising the nine-item quiz were only administered to respondents who were in their first year of participation in the MCBS. Therefore, a much smaller number of respondents was available for the 1999 analyses of these indices.

Only participants who were living in the community, rather than an institution, were included in the psychometric analyses. In addition, because this was an elderly population whose members were likely to experience disabilities, the use of a proxy was sometimes necessary to obtain information on a respondent.[3] Therefore, for completeness, proxy information was included in these analyses. However, for each of these comparisons, data for sample members and proxy respondents were analyzed separately. It was expected that proxy and sample member participants would respond to the knowledge indices differently.

4.1 Perceived Knowledge Index

4.1.1 Item-Level Analyses

Each item in the perceived knowledge index contains these five response options: (1) "Almost none of what you need to know," (2) "A little of what you need to know," (3) "Some of what you need to know," (4) "Most of what you need to know," and (5) "Just about everything you need to know." Because these items had more than two response options, we examined the distribution of responses for any possible floor or ceiling effects. A floor effect would be present if respondents tended to select only the lowest response options, while a ceiling effect would occur if respondents selected only the highest response options. The presence of either of these effects would restrict the possible range of scores and thereby limit the ability of the index scores to discriminate among respondents with different levels of knowledge.

Table 1 presents the response distributions of the perceived knowledge questions for the sample members in each year while ***Table 2*** presents the response distributions for proxy respondents. Overall, the perceived knowledge items demonstrated good variability across responses, suggesting that there are no floor or ceiling effects for any of the items.

[3]In all three years, proxy interviews comprised approximately 10 percent of the interviews conducted.

Table 1
Response distribution of the perceived knowledge questions among sample members

Question	Almost none	A little	Some	Most	Just about everything
Sample Members—1998					
Services Medicare covers	13%	17%	28%	29%	13%
Paying for medical services	13%	16%	24%	29%	18%
Supplemental insurance	26%	16%	20%	23%	15%
Medicare HMOs	41%	17%	16%	15%	11%
Choosing a doctor	8%	11%	20%	34%	28%
Sample Members—1999					
Services Medicare covers	11%	16%	27%	31%	16%
Paying for medical services	11%	15%	24%	31%	20%
Supplemental insurance	25%	15%	19%	25%	16%
Medicare HMOs	41%	15%	15%	15%	13%
Choosing a doctor	7%	9%	19%	34%	30%
Sample Members—2000					
Services Medicare covers	10%	15%	26%	33%	15%
Paying for medical services	10%	15%	24%	32%	20%
Supplemental insurance	22%	15%	20%	27%	16%
Medicare HMOs	42%	16%	15%	16%	12%
Choosing a doctor	7%	10%	20%	36%	27%

SOURCE: Centers for Medicare & Medicaid Services, MCBS 1998, 1999, and 2000 Access to Care and Supplemental Files BN 24, 27, and 30.

Table 2
Response distribution of the perceived knowledge questions among proxy respondents

Question	Almost none	A little	Some	Most	Just about everything
Proxy Respondents—1998					
Services Medicare covers	15%	18%	31%	24%	12%
Paying for medical services	14%	15%	24%	27%	20%
Supplemental insurance	30%	19%	18%	20%	13%
Medicare HMOs	42%	20%	16%	13%	9%
Choosing a doctor	8%	12%	20%	34%	27%
Proxy Respondents—1999					
Services Medicare covers	16%	15%	28%	27%	14%
Paying for medical services	16%	11%	22%	29%	22%
Supplemental insurance	31%	15%	17%	21%	16%
Medicare HMOs	45%	16%	15%	14%	11%
Choosing a doctor	8%	10%	19%	35%	28%
Proxy Respondents—2000					
Services Medicare covers	12%	17%	27%	30%	14%
Paying for medical services	12%	13%	22%	32%	21%
Supplemental insurance	28%	16%	18%	24%	14%
Medicare HMOs	43%	17%	15%	15%	9%
Choosing a doctor	8%	11%	20%	35%	25%

SOURCE: Centers for Medicare & Medicaid Services, MCBS 1998, 1999, and 2000 Access to Care and Supplemental Files BN 24, 27, and 30.

Tables 3 and ***4*** display the item–total score correlations for the perceived knowledge items, separately for sample members and proxies. The correlations are very similar across all 3 years. All of the correlations are 0.5 or greater, suggesting that these items are highly related and appear to be measuring the same construct. The first two items (services Medicare covers and paying for medical services) are the most highly related to the underlying construct with item–total score correlations around 0.7. The item about Medicare HMOs is the least related to the construct with correlations of approximately 0.5; however, it still contributes to the reliability of the index.

Table 3
Item–total score correlations for the perceived knowledge questions for sample member respondents

	Sample Member Interviews		
Question	1998	1999	2000
Services Medicare covers	0.71	0.71	0.72
Paying for medical services	0.69	0.70	0.71
Supplemental insurance	0.62	0.63	0.64
Medicare HMOs	0.50	0.50	0.49
Choosing a doctor	0.56	0.56	0.56

SOURCE: Centers for Medicare & Medicaid Services, MCBS 1998 to 2000 Access to Care and Supplemental Files BN 24, 27, and 30.

Table 4
Item–total score correlations for the perceived knowledge questions for proxy respondents

	Proxy Interviews		
Question	1998	1999	2000
Services Medicare covers	0.72	0.74	0.76
Paying for medical services	0.74	0.72	0.73
Supplemental insurance	0.70	0.70	0.67
Medicare HMOs	0.54	0.54	0.56
Choosing a doctor	0.61	0.62	0.63

SOURCE: Centers for Medicare & Medicaid Services, MCBS 1998 to 2000 Access to Care and Supplemental Files BN 24, 27, and 30.

4.1.2 Scale-Level Analyses

This section describes the scale-level descriptive statistics and the reliability and validity analyses for the perceived knowledge index.

Descriptive statistics—The descriptive statistics for the perceived knowledge index are presented in ***Table 5***. The modes for sample members in 1999 and 2000 are higher than the means and medians, suggesting that the distributions of scores are skewed to the right. In other words, the scores tend to be clustered on the higher end of the scale, suggesting that many respondents received scores higher than 15 (the middle possible score). The distributions of scores for proxy respondents in 1998 and 2000 are also slightly skewed to the right while the distributions for sample members in 1998 and proxies in 1999 appear to be fairly normally distributed with similar values for the means, medians, and modes.

Table 5
Descriptive statistics for the perceived knowledge index

	N	Mean	S.D.	Median	Mode
Sample member interviews (survey year)					
1998 (Round 24)	12,524	15.2	5.0	15	15
1999 (Round 27)	12,606	15.6	5.0	16	20
2000 (Round 30)	12,446	15.7	5.0	16	20
Proxy interviews (survey year)					
1998 (Round 24)	1,330	14.9	5.6	15	17
1999 (Round 27)	1,382	15.1	5.3	15	15
2000 (Round 30)	1,304	15.2	5.2	15	20

Note: The possible range of scores for the perceived knowledge index was from 5 to 25.

Reliability—The value of Cronbach's alpha was 0.82 for sample members in all three years. Similar values for coefficient alpha were obtained for proxy respondents in 1998 ($\alpha = 0.84$), 1999 ($\alpha = 0.85$), and 2000 ($\alpha = 0.85$). These values indicate that the perceived knowledge index demonstrated strong internal consistency reliability.

Coefficient alphas of the perceived knowledge index were also calculated separately for various subgroups classified according to enrollment in managed care, institutional and Part B utilization, and amounts of allowable and reimbursed charges. For details of these results, please refer to ***Tables A.3*** through ***A.7*** of ***Appendix A***. The coefficient alphas of these different groups were very similar. Among sample member respondents, values of alpha ranged from 0.81 to 0.85 while proxy respondents had alphas ranging from 0.81 to 0.90.

Validity—Relationship with global perceived knowledge question. We assessed the construct validity of the perceived knowledge index with three different approaches. First, we evaluated the relationship between the perceived knowledge index and a global knowledge item that asks beneficiaries how much they feel they know about the Medicare program. Analyses of variance (ANOVAs) indicated that perceived knowledge index scores varied significantly across levels of the global perceived knowledge question for sample members in all three years: 1998 ($F(4, 12336) = 762.71$, $p < 0.0001$), 1999 ($F(4, 3617) = 250.33$, $p < 0.0001$), and 2000 ($F(4,12248) = 1011.51$, $p < 0.0001$). Similar results were found for proxy respondents in 1998 ($F(4, 1323) = 67.82$, $p < 0.0001$), 1999 ($F(4, 369) = 17.16$, $p < 0.0001$), and 2000 ($F(4,1274) = 86.18$, $p < 0.0001$). These results indicate that the perceived knowledge index scores were related to the global perceived knowledge scores other than by chance.

To help interpret the ANOVA results, we computed the means and standard deviations of the perceived knowledge index scores for each of the five response categories included in the global perceived knowledge question. As shown in ***Table 6***, there is a clear, monotonic relationship between the perceived knowledge index scores and the global knowledge question across all three years and both types of respondents. Individuals who rated their knowledge higher on the global perceived knowledge question had higher scores on the perceived knowledge index.

Table 6
Means (and standard deviations) of perceived knowledge index scores by response to global perceived knowledge question

Survey year	Almost none	A little	Some	Most	Just about everything
Sample member interviews					
1998 (Round 24)	11.8 (5.0)	13.4 (4.5)	15.3 (4.3)	17.5 (4.2)	18.7 (4.7)
1999 (Round 27)	11.6 (4.8)	13.9 (4.7)	15.6 (4.4)	17.8 (4.0)	18.9 (4.4)
2000 (Round 30)	11.2 (4.7)	13.5 (4.5)	15.5 (4.1)	17.9 (4.0)	19.3 (4.6)
Proxy interviews					
1998 (Round 24)	11.7 (5.2)	13.3 (4.6)	15.0 (4.3)	17.2 (4.3)	18.4 (5.3)
1999 (Round 27)	11.6 (5.0)	14.3 (5.3)	15.2 (5.2)	16.4 (5.2)	20.0 (5.2)
2000 (Round 30)	11.7 (5.1)	13.1 (4.6)	15.2 (4.4)	17.6 (4.2)	19.5 (5.4)

Group comparisons. For the next set of validity analyses, we compared the perceived knowledge index scores of beneficiaries who have been shown in previous research to have different levels of knowledge; we referred to this set of analyses as known-groups comparisons. As another set of validity analyses, we compared beneficiaries who we hypothesized would have different levels of knowledge which we referred to as exploratory comparisons. In both sets of analyses, the index value validity measures were computed separately for sample member and

proxy respondents. T-tests were used to compare the perceived knowledge index scores for sample members and proxy respondents. As hypothesized, sample member respondents had significantly different scores on the perceived knowledge index than proxy respondents in all three years: 1998 ($t(13883)=2.97$, $p = 0.003$), 1999 ($t(1660)=3.35$, $p < 0.001$), and 2000 ($t(1562) = 3.00$, $p = 0.003$).

ANOVAs and t-tests were used to compare the perceived knowledge index scores of respondents according to the various background and experience variables. For the t-tests, the assumption of equal variance for the two groups was tested using the F´ (folded) statistic (Steel and Torrie, 1980). As implemented in SAS, if this test was significant (i.e., the variances for the two groups were not equal), an approximate t statistic was computed, using Satterthwaite's approximation to estimate the degrees of freedom (Satterthwaite, 1946).

The mean perceived knowledge index scores for each of the groups are presented separately according to interview type in ***Tables A.8*** and ***A.9***. Asterisks are used to denote the significance of the relevant ANOVA or t-test. As shown in ***Table A.8***, perceived knowledge index scores for sample members in all three years differed significantly on all of the variables. Examining the patterns of means indicates that the results are generally in the expected direction. Sample members with more education and higher incomes received higher knowledge index scores. Also, higher scores were found for sample members between 65 and 75 years of age, with some institutional utilization, some Part B utilization, private supplemental insurance, or enrollment in managed care. However, the means of the charges variables did not always follow an entirely monotonic pattern, possibly due to the number and range of cut-off points chosen. In some cases, the mean knowledge index scores for respondents with the most charges (e.g., $5,000 or more) were similar or smaller than the mean scores for respondents with fewer charges (e.g., $500-$4,999). However, overall, there does appear to be a general pattern with individuals who have no charges having lower knowledge scores than those with any charges.

Table A.9 displays the results for proxy respondents in each year. Generally, the patterns of means are very similar to those of sample members. Higher knowledge scores were found among respondents with higher education, higher incomes, some Part B utilization, or any charges. However, there were fewer significant differences for proxies than sample member respondents, possibly due to the smaller sample size. Another possibility is that perhaps proxies use some of their own knowledge in answering the questions. This could result in less of a relationship between the information proxies provide and the sample member's characteristics.

4.2 Nine-Item Quiz

This section outlines the item-level and scale-level analyses of the nine-item quiz. For the education-level analyses, only data from sample members in 1999 are used. As mentioned earlier, during 1999, only participants who were new to the MCBS were administered the quiz questions compared to the other two years in which all respondents received the questions. Because the participants in 1999 had not seen the quiz questions before, their responses should provide a truer measure of the difficulty level of the items. Proxy respondents were excluded from the grade-level analyses because it was unclear whether responses provided by a proxy would correspond directly to a sample member's education level.

4.2.1 Item-Level Analyses

While the perceived knowledge questions ask respondents to rate their own knowledge, the questions on the nine-item quiz require respondents to demonstrate their knowledge by determining whether a statement is true or false. Therefore, participants' responses to these questions may be classified as correct or incorrect. To examine the difficulty of the items, we computed the percentage of sample member respondents who answered each item correctly; these results are presented in ***Table A.10***.

The item about flu shots appeared to be the easiest item with 84 percent of respondents answering it correctly. In contrast, the item concerning whether HMOs cover more health services was the most difficult with only 37 percent of respondents answering it correctly.

To help make the difficulty level of the items more meaningful, we attempted to match an education level to the items, using information about beneficiaries' educational achievement. We began by classifying sample members in 1999 into the following five categories based on the highest school grade they completed: (1) no formal education, (2) 8th grade or less, (3) 9th to 12th grade without a high school diploma, (4) high school graduate, and (5) college. Then we calculated the percentage of sample members answering an item correctly, according to their educational achievement. For an item to be matched to a particular education level, we expected at least 50 percent of respondents at that level (as well as at all higher levels) to have answered the item correctly.

Table A.10 contains a list of the quiz items, their corresponding education level, and the number of respondents at each education level who answered the item correctly. The education level assigned to the item is shown in bold. Using this criterion, the nine-item quiz contains items covering all education levels. The item concerning whether Medicare covers a flu shot was very easy; 71 percent of beneficiaries with no formal education answered this question correctly. However, the question concerning whether HMOs cover more services was much more difficult; only 46 percent of respondents with a college education answered this question correctly.

We also computed an education level for the overall nine-item quiz. To assign an educational level to the quiz, we required that at least 50 percent of respondents at that level answer at least 50 percent of the quiz items correctly. In other words, respondents must have received an overall quiz score of 5. The results indicated that the nine-item quiz corresponds to a high school level. Fifty-one percent of beneficiaries reporting some high school education received a quiz score of at least 5.

The item–total score correlations for the questions in the nine-item quiz are presented in ***Tables 7 and 8*** for sample members and proxies, respectively. A common criterion requires that correlations be greater than 0.30 for the items to be considered as having contributed significantly to the scale. Only one of the items (Medicare covers an annual flu shot) failed to meet this criterion. It had low correlations during 1998 and 1999 possibly due to the fact that this item, along with the item regarding supplemental insurance, was administered in a different round than the other items. In the time between rounds, beneficiaries' knowledge may have changed, making these items appear slightly less related to the other items. In fact, the item-total correlations for these items increase during 2000 when all of the items were included in the same round.

Table 7
Item–total score correlations for the nine-item quiz questions for sample members

	Sample Member Interviews		
Question	1998	1999	2000
Supplemental insurance is the same as managed care	0.31	0.29	0.35
Medicare covers an annual flu shot	0.22	0.23	0.30
Can select different health plan options	0.43	0.39	0.47
Medicare alone pays for all health care expenses	0.37	0.40	0.49
Medicare offers more information	0.36	0.37	0.47
Can report complaints to Medicare about HMOs and supplemental insurance	0.44	0.41	0.47
Limited choices of doctors if on HMOs	0.52	0.48	0.55
Can drop HMO and still be covered by Medicare	0.54	0.51	0.59
HMOs cover more health services	0.48	0.45	0.49

Table 8
Item–total score correlations for the nine-item quiz questions for proxy respondents

	Proxy Interviews		
Question	1998	1999	2000
Supplemental insurance is the same as managed care	0.35	0.35	0.37
Medicare covers an annual flu shot	0.22	0.22	0.35
Can select different health plan options	0.44	0.43	0.52
Medicare alone pays for all health care expenses	0.41	0.40	0.49
Medicare offers more information	0.47	0.36	0.53
Can report complaints to Medicare about HMOs and supplemental insurance	0.52	0.46	0.52
Limited choices of doctors if on HMOs	0.58	0.54	0.58
Can drop HMO and still be covered by Medicare	0.59	0.55	0.62
HMOs cover more health services	0.51	0.49	0.52

4.2.2 Scale-Level Analyses

Following the item-level analyses, additional analyses were conducted to explore the psychometric properties of the nine-item quiz as a whole. This section describes the results of these scale-level analyses.

Descriptive statistics—*Table 9* displays the descriptive statistics for the nine-item quiz scores, displayed separately for sample member and proxy respondents. As shown in the table, the distribution of scores was slightly skewed to the right with the mode generally having a higher value than the mean and median. Most respondents received scores of "5" or higher, indicating that they correctly answered at least half of the questions.

Table 9
Descriptive statistics for the nine-item quiz

	N	Mean	S.D.	Median	Mode
Sample member interviews (survey year)					
1998 (Round 23)	13,175	4.94	2.45	5	6
1999 (Round 26)	3,976	4.86	2.42	5	6
2000 (Round 29)	12,927	5.34	2.55	6	7
Proxy interviews (survey year)					
1998 (Round 23)	1,495	4.62	2.57	5	6
1999 (Round 26)	422	4.56	2.49	5	5
2000 (Round 29)	1,390	5.02	2.70	5	8

Reliability—The internal consistency reliability of the nine-item quiz was measured using Cronbach's alpha. The alpha coefficients for the nine-item quiz among sample members was 0.73 in 1998, 0.72 in 1999, and 0.78 in 2000. Among proxy respondents, the alpha coefficient for the quiz was 0.76 in 1998, 0.75 in 1999, and 0.81 in 2000. These values indicate that the nine-item quiz reached an acceptable level of reliability.

The alpha coefficients for the nine-item quiz were also calculated separately for various subgroups to assess whether its reliability remained consistent. The results are presented in ***Tables A.11*** through ***A.15*** of ***Appendix A***. Generally, the values of alpha were fairly consistent across all of the groups, with alphas ranging from 0.68 to 0.83 for sample members and from 0.69 to 0.83 for proxy respondents.

Validity—To assess construct validity, particular groups who were expected to differ in knowledge were compared to determine if they did in fact receive different scores on the nine-item quiz. (See ***Tables A.16*** and ***A.17*** in ***Appendix A*** for details of the statistical results.) These analyses were computed separately for sample member and proxy respondents. As

hypothesized, sample members received significantly higher scores on the nine-item quiz than proxy respondents in all three years: 1998 ($t(1816) = 4.64$, $p < 0.0001$), 1999 ($t(4396) = 2.41$, $p=.02$), and 2000 ($t(1671) = 4.16$, $p < 0.0001$).[4]

Relationship with global perceived knowledge question. ANOVAs revealed that the nine-item quiz scores of sample members varied significantly according to responses to the global perceived knowledge question in all three years: 1998 ($F(4, 12973) = 413.54$, $p < 0.0001$), 1999 ($F(4, 3909) = 134.79$, $p < 0.0001$), and 2000 ($F(4,12884) = 487.39$, $p < 0.0001$). Similar results were found for proxy respondents in 1998 ($F(4, 1455) = 46.83$, $p < 0.0001$), 1999 ($F(4, 411) = 9.52$, $p < 0.0001$), and 2000 ($F(4,1381) = 51.50$, $p < 0.0001$).

Table 10 displays the means and standard deviations of the nine-item quiz scores by responses to the global perceived knowledge question. Across years and respondent types, there is a monotonic relationship between scores on the nine-item quiz and the first four response categories ("Almost none," "A little," "Some," and "Most") for the global perceived knowledge question. Respondents who rated their knowledge higher on the global knowledge question received higher scores on the nine-item quiz. This pattern seems to taper off for the last response category ("Just about everything"). Generally, respondents with ratings of "Most" or "Just about everything" have similar quiz scores. It is possible that respondents do not clearly distinguish between the "Most" and "Just about everything" response categories.

Table 10
Means (and standard deviations) of nine-item quiz scores by response to global perceived knowledge question

	Almost none	A little	Some	Most	Just about everything
Sample member interviews (survey year)					
1998 (Round 23)	3.5 (2.4)	4.5 (2.3)	5.2 (2.2)	5.9 (2.2)	5.8 (2.3)
1999 (Round 26)	3.4 (2.3)	4.5 (2.3)	5.2 (2.2)	5.8 (2.1)	5.7 (2.5)
2000 (Round 29)	3.5 (2.6)	4.6 (2.5)	5.4 (2.3)	6.3 (2.2)	6.2 (2.3)
Proxy interviews (survey year)					
1998 (Round 23)	3.2 (2.5)	4.2 (2.5)	5.0 (2.3)	5.6 (2.3)	5.7 (2.4)
1999 (Round 26)	3.4 (2.1)	4.2 (2.5)	5.0 (2.4)	5.5 (2.3)	5.0 (2.8)
2000 (Round 29)	3.4 (2.8)	4.4 (2.7)	5.0 (2.4)	6.2 (2.3)	6.0 (2.6)

[4]Please note that for these t-tests the degrees of freedom are smaller for the 1998 and 2000 samples than the 1999 sample even though they have larger sample sizes. The Folded F test for equality of variances indicated that the variances for the proxy and sample members in these two years could not be assumed to be equal. Therefore, Satterthwaite's (1946) approximation was used which usually results in a decrease in the degrees of freedom.

Group comparisons. ANOVAs and t-tests were conducted to determine whether various subgroups received different scores on the nine-item quiz. The mean nine-item quiz scores for each of the groups are presented separately according to interview type in ***Tables A.16*** and ***A.17***. Asterisks indicate the significance level of the corresponding ANOVA or t-test.

As shown in ***Table A.16***, nine-item quiz scores for sample members in 1999 differed significantly on all of the variables. The results from sample members in 1998 and 2000 were significant for all variables except institutional utilization. The patterns of mean scores for all three years indicate that the results are generally in the expected direction. Higher quiz scores were associated with sample members who had higher education, higher incomes, some charges, private supplemental insurance, or were between the ages of 65 and 75. In addition, as hypothesized, respondents who had been enrolled in managed care during the past year had higher nine-item quiz scores than those who had not been enrolled during the past year.

Table A.17 displays the results for proxy respondents in each year. Across all years, there were significant differences for all variables except age, institutional utilization, and covered institutional charges. Examining the means reveals that higher quiz scores were received by respondents with more education, higher incomes, private supplemental insurance, some Part B utilization, or who had been enrolled in managed care during the past year.

5. CONCLUSIONS AND RECOMMENDATIONS

We developed two potential knowledge measures using data from the BK and BN supplemental rounds of the 1998 to 2000 MCBS. The first measure, the perceived knowledge index, includes five questions that ask beneficiaries to rate how much they know about a particular topic related to Medicare. The other measure is a nine-item quiz that requires participants to respond to a series of true/false questions.

We evaluated the psychometric properties of the knowledge measures by calculating item-level and scale-level descriptive statistics, assessing internal consistency reliability, and conducting construct validity analyses. The item-level analyses of the perceived knowledge items suggest that these items appropriately do not exhibit floor or ceiling effects. Comparisons of the nine-item quiz scores according to beneficiaries' educational achievement suggest that the difficulty level of the quiz is equivalent to a high school level.

To assess internal consistency reliability, we computed Cronbach's alphas for both of the measures. A level of 0.70 is commonly required for an alpha to be considered acceptable for use in group-level analyses. As shown in ***Table 11***, both the perceived knowledge index and the nine-item quiz met this criterion. The nine-item quiz had coefficient alphas of 0.72 to 0.81, while the perceived knowledge index demonstrated strong internal consistency reliability with alphas ranging from 0.82 to 0.85.

Table 11
Coefficient alphas of the knowledge indices

	Perceived knowledge index	9-item quiz
Sample member interviews (survey year)		
1998	0.82	0.73
1999	0.82	0.72
2000	0.82	0.78
Proxy interviews (survey year)		
1998	0.84	0.76
1999	0.85	0.75
2000	0.85	0.81

To evaluate the construct validity of the knowledge indices, we examined the relationship between each index and a global knowledge item that asks respondents how much they feel they know about the Medicare program, referred to as the global perceived knowledge question. A strong relationship between a knowledge index and another measure of the same construct (e.g., the global perceived knowledge question) would provide support for the construct validity of the knowledge index. The results indicated that both knowledge indices are significantly related to

the global perceived knowledge question. Respondents who rated their knowledge higher on this question received higher scores on each knowledge index. In particular, there appeared to be a strong, monotonic relationship between the perceived knowledge index and the global perceived knowledge question, possibly because of the similarity in wording between the two. Nonetheless we can still infer that they closely measure the same construct.

As a further evaluation of the validity of the knowledge indices, we analyzed comparisons of the mean knowledge index scores for various groups. The first set of analyses, sometimes referred to as known-groups comparisons, involve comparisons between groups previously shown to have different knowledge levels. We also compared groups we hypothesized would differ in knowledge which we referred to as exploratory comparisons. Both indices performed well on the known group comparisons. In general, groups that were expected to differ in knowledge had significantly different knowledge index scores. Because of the small number of proxy respondents in 1999, however, the knowledge index scores for this group had much lower statistical power and the differences were therefore less likely to be significant.

One of the most consistent findings across all years and interview types was a strong relationship between higher educational achievement and higher knowledge scores. As mentioned in Chapter 3, this relationship has been found repeatedly in prior research. Possible explanations include those with more education having better comprehension of the program materials or a greater motivation to learn about the Medicare program. Another finding was that, as hypothesized, the relationship between knowledge scores and enrollment in managed care was strongest for the nine-item quiz. This quiz includes four questions concerning managed care plans, while the perceived knowledge index contains only one question on this topic.

Based solely on the quantitative results, the perceived knowledge index seems to have the best psychometric properties. This index performed well in both the reliability and validity analyses. However, other criteria, such as the content of the items, must be considered when selecting the most appropriate measure of knowledge. The perceived knowledge index relies on beneficiaries to be the sole judge of their knowledge. Individuals' subjective ratings of their own knowledge may be influenced by factors other than knowledge, such as confidence in decision making or satisfaction with information received. Therefore, the perceived knowledge index may not provide the most accurate assessment of actual knowledge level. A more accurate measure of knowledge would require respondents to actually demonstrate their knowledge.

On the basis of both content considerations and the psychometric analysis results, the nine-item quiz appears to be the most useful measure of beneficiary knowledge overall. The quiz requires beneficiaries to demonstrate their knowledge by correctly answering true/false questions rather than merely stating that they know everything they need to know. In addition, the quiz had good variability in scores, reached an acceptable level of internal consistency reliability, and performed well on the validity analyses.

Possible modifications to the nine-item quiz depend on the intended use of the measure. If the goal of the quiz is to obtain the most precise estimate of beneficiaries' knowledge possible, then the quiz should contain questions that cover all difficulty levels. Through the education-level analyses suggest that the items on the nine-item quiz does in fact cover all education levels, the quiz contains only one item at the high school graduate level whereas it has four items at the

college level. Because more respondents report being a high school graduate than any other educational level, it may be helpful to add more items targeted at this group.

In contrast, if the purpose of the quiz is to determine whether Medicare beneficiaries reach a certain proficiency level, then the primary emphasis should be on the content of the quiz items. The items should cover the entire range of information that beneficiaries need to know in order to make informed decisions. Less emphasis should be placed on the difficulty of the items. In this case, a limitation of the nine-item quiz is that four of its nine questions deal with managed care plans which restricts the range of knowledge that the quiz can measure. Other questions could be added to improve the comprehensiveness of the quiz. In fact, as a part of RTI's *Questionnaire Development and Cognitive Testing Using Item Response Theory* project, several new knowledge questions were developed that address a variety of issues, including beneficiary rights and health plan decision-making (Uhrig et al., 2001), and could be used to expand the quiz.

REFERENCES

Bann, C., Lissy, K.S., Keller, S., Garfinkel, S.A., and Bonito, A.J.: Analysis of Medicare Beneficiary Baseline Knowledge Data from the Medicare Current Beneficiary Survey: Knowledge Index Technical Note. Report prepared for the Health Care Financing Administration under Contract No. 500-95-0061/004. Research Triangle Park, NC. Research Triangle Institute, 2000.

Cafferata, C.: Knowledge of their health insurance coverage by the elderly. Med. Care 22(9):835-847, 1984.

Chapman, D.: A survey of nonresponse imputation procedures. In American Statistical Association Proceedings, Section on Survey Research Methods. Washington, DC. American Statistical Association, 1976.

Cronbach, L.J.: Coefficient alpha and the internal structure of tests. Psychometrika 16(3):297-334, 1951.

Guilford, J.P.: Psychometric Methods. New York. McGraw-Hill, 1956.

Health Care Financing Administration: Medicare & You 2000. HCFA Pub. No. 10050. Baltimore, MD. Health Care Financing Administration, 2000.

Hibbard, J.H., Jewett, J.J., Englemann, S., and Tusler, M.: Can Medicare beneficiaries make informed choices? Health Aff. 17(6):181-193, 1998.

Howard, K.I., and Forehand, G.C.: A method for correcting item-total correlations for the effect of relevant item inclusion. Educational and Psychological Measurement 22(4):731-735, 1962.

Lambert, Z.V.: Elderly consumers' knowledge related to Medigap protection needs. Journal of Consumer Affairs 14(2):434-451, 1980.

Marquis, M.S.: Consumer's knowledge about their health insurance coverage. Health Care Financing Review 5(1):65-80, 1983.

McCall, N., Rice, T., and Sangl, J.: Consumer knowledge of Medicare and supplemental health insurance benefits. Health Serv. Res. 20(6):633-657, 1986.

McCormack, L.A., Anderson, W., Daugherty, S., Ross, K., Kuo, M., Lynch, J., Scheffler, S., and Garfinkel, S.A.: National Evaluation of the Medicare & You 2000 Handbook Final Report. Report prepared for the Centers for Medicare & Medicaid Services under Contract No. 500-96-0010/004. Research Triangle Park, NC. Research Triangle Institute, 2001.

McCormack, L.A., Garfinkel, S.A., Hibbard, J.H., Keller, S., and Kilpatrick, K.E.: Health insurance knowledge among Medicare beneficiaries. Health Serv. Res. 37(1):43-63, 2002.

Nunnally, J.O.: Psychometric Theory. New York. McGraw-Hill, 1978.

Satterthwaite, F.W.: An approximate distribution of estimates of variance components. Biometrics Bulletin 2:110-114, 1946.

Steel, R.G.D., and Torrie, J.H.: Principles and Procedures of Statistics (2nd edition). New York. McGraw-Hill, 1980.

Uhrig, J.D., Squire, C., McCormack, L.A., Bann, C., Hall, P.K., An, C., and Bonito, A.J.: Questionnaire Development and Cognitive Testing Using Item Response Theory (IRT): Questionnaire Development Report. Report prepared for the Centers for Medicare & Medicaid Services under Contract No. 500-00-0024/002. Research Triangle Park, NC. Research Triangle Institute, 2001.

APPENDIX A:
TABLES OF RESULTS FROM PSYCHOMETRIC ANALYSES

LIST OF TABLES IN APPENDIX A

Table A.1
Timeline of administration of MCBS supplemental rounds

Calendar year	January to April 1999	May to August 1999	Fall 1999	January to April 2000	May to August 2000	Fall 2000	January to April 2001	May to August 2001
MCBS year	1998 MCBS			1999 MCBS			2000 MCBS	
MCBS round number	Round 23 beneficiary knowledge (BK)	Round 24 beneficiary needs (BN)		Round 26 beneficiary knowledge (BK)	Round 27 beneficiary needs (BN)		Round 29 beneficiary knowledge (BK)	Round 30 beneficiary needs (BN)
NMEP event			Nationwide distribution of Medicare & You 2000 handbook			Nationwide distribution of Medicare & You 2001 handbook		

Table A.2
MCBS knowledge questions

Question wording	Round	Year
Perceived Knowledge Index	24, 27, 30	1998, 1999, 2000
BN 1: How much do you feel you know about what medical services Medicare covers or does not cover? Do you know just about everything you need to know, most of what you need to know, some of what you need to know, a little of what you need to know, or almost none of what you need to know about what Medicare covers or doesn't cover?		
BN 2: How much do you feel you know about how much you have to pay for medical services? Do you know just about everything you need to know, most of what you need to know, some of what you need to know, a little of what you need to know, or almost none of what you need to know about how much you have to pay for medical services?		
BN 3: How much do you feel you know about supplemental or Medigap insurance, such as what it covers or how it works with Medicare to pay medical claims? Do you know just about everything you need to know, most of what you need to know, some of what you need to know, a little of what you need to know, or almost none of what you need to know about supplemental insurance?		
BN 4: How much do you feel you know about the availability and benefits of Medicare managed care plans? Do you know just about everything you need to know, most of what you need to know, some of what you need to know, a little of what you need to know, or almost none of what you need to know about the availability and benefits of managed care plans?		
BN 5: How much do you feel you know about choosing or finding a doctor or other health care provider? Do you know just about everything you need to know, most of what you need to know, some of what you need to know, a little of what you need to know, or almost none of what you need to know about finding a doctor or other health care provider?		

SOURCE: MCBS Supplemental Rounds BN 24, 27, and 30.

Table A.2
(continued)

Question wording	Round	Year
Nine-Item Quiz[1, 2]	23, 24, 26, 27, 29	1998, 1999, 2000
BN 18: Medigap or supplemental insurance is the same as a Medicare managed care plan. *(False)*		
BN 19: Medicare covers an annual flu shot. *(True)*		
BK 43: Most people covered by Medicare can select among different kinds of health plan options within Medicare. *(True)*		
BK 44: Medicare without a supplemental insurance policy pays for all of your health care expenses. *(False)*		
BK 46: The Medicare program has begun to offer more information and help in order to answer your Medicare questions. *(True)*		
BK 47: People can report complaints to Medicare about their Medicare managed care plans (HMOs) or supplemental plans if they are not satisfied with them. *(True)*		
BK 48: If someone joins a Medicare managed care plan (HMO) that covers people on Medicare, they have limited choices about which doctors they can see. *(True)*		
BK 49: If someone joins a Medicare managed care plan (HMO) that covers people on Medicare, they can change or drop the plan and still be covered by Medicare. *(True)*		
BK 50: Medicare managed care plans (HMOs) that cover people on Medicare often cover more health services, like prescribed medicines, than Medicare without a supplemental policy. *(True)*		

[1]The correct answers to the quiz questions are presented in italics following each question.

[2]During 1998 and 1999, the first two quiz questions were administered in the BN supplement (Rounds 24 and 27) while the remaining items were administered as a part of the BK supplement (Rounds 23 and 26). In 2000, all of the items were administered as a part of the BK supplement (Round 29).

SOURCE: MCBS Supplemental Rounds BN 24 and 27 and BK 23, 26, and 29

Table A.3
Coefficient alphas of the perceived knowledge index, by enrollment in managed care during the past year

Interview type/survey year	No enrollment	Some enrollment
Sample member interviews		
1998 (Round 24)	0.82	0.83
1999 (Round 27)	0.83	0.84
2000 (Round 30)	0.83	0.85
Proxy interviews		
1998 (Round 24)	0.84	0.85
1999 (Round 27)	0.85	0.87
2000 (Round 30)	0.86	0.85

SOURCE: Centers for Medicare & Medicaid Services, MCBS 1998, 1999, and 2000 Access to Care and Supplemental Files.

Table A.4
Coefficient alphas of the perceived knowledge index, by institutional utilization[1]

Interview type/survey year	No utilization	Some utilization
Sample member interviews		
1998 (Round 24)	0.83	0.82
1999 (Round 27)	0.83	0.82
2000 (Round 30)	0.83	0.82
Proxy interviews		
1998 (Round 24)	0.86	0.83
1999 (Round 27)	0.87	0.84
2000 (Round 30)	0.86	0.85

[1]Respondents who were enrolled in managed care during the past year were excluded from these analyses.

Table A.5
Coefficient alphas of the perceived knowledge index, by part b utilization[1]

Interview type/survey year	No utilization	Some utilization
Sample member interviews		
1998 (Round 24)	0.83	0.82
1999 (Round 27)	0.83	0.82
2000 (Round 30)	0.85	0.82
Proxy interviews		
1998 (Round 24)	0.87	0.84
1999 (Round 27)	0.87	0.85
2000 (Round 30)	0.90	0.85

[1]Respondents who were enrolled in managed care during the past year were excluded from these analyses.

Table A.6
Coefficient alphas of the perceived knowledge index, by total reimbursed dollars[1]

Interview type/survey year	$0	$1 – $499	$500 – $4,999	$5,000 or more
Sample member interviews				
1998 (Round 24)	0.83	0.82	0.82	0.83
1999 (Round 27)	0.83	0.82	0.82	0.83
2000 (Round 30)	0.85	0.81	0.82	0.84
Proxy interviews				
1998 (Round 24)	0.86	0.85	0.83	0.82
1999 (Round 27)	0.87	0.86	0.84	0.83
2000 (Round 30)	0.89	0.86	0.84	0.84

[1]Respondents who were enrolled in managed care during the past year were excluded from these analyses.

Table A.7
Coefficient alphas of the perceived knowledge index, by covered institutional charges[1]

Interview type/survey year	$0	$1 – $499	$500 – $4,999	$5,000 or more
Sample member interviews				
1998 (Round 24)	0.83	0.82	0.82	0.83
1999 (Round 27)	0.83	0.81	0.83	0.83
2000 (Round 30)	0.83	0.81	0.83	0.84
Proxy interviews				
1998 (Round 24)	0.86	0.83	0.84	0.81
1999 (Round 27)	0.87	0.85	0.84	0.82
2000 (Round 30)	0.86	0.87	0.86	0.83

[1]Respondents who were enrolled in managed care during the past year were excluded from these analyses.

Table A.8
Mean perceived knowledge index scores and statistical significance of ANOVAs and t-tests among sample members

Variable	1998	1999	2000
Education	F(2,12521) = 427.93***	F(2,12569) = 488.15***	F(2,12415) = 505.81***
8th grade or less	12.92	13.06	13.02
More than 8th grade, but no college	15.28	15.54	15.39
College	16.59	17.19	17.24
Age category	t(10671) = 8.09***	t(11240) = 13.24***	t(10544) = 7.66***
65-75	15.84	16.31	16.30
Over 75	15.07	15.54	15.57
Income category	t(8164) = -25.81***	t(8886) = -26.39***	t(9778) = -30.15***
$25,000 or less	14.50	14.84	14.73
More than $25,000	16.86	17.24	17.38
Managed care	t(3552) = -12.64***	t(12592) = 22.74***	t(12431) = -10.32***
Some enrollment	16.37	16.71	16.60
No enrollment	14.95	15.38	15.46
Private Supplemental Insurance	t(11715) = -15.92***	t(1911) = -6.50***	t(9817) = 22.87***
Supplemental insurance	16.03	16.25	16.52
No supplemental insurance	14.62	14.75	14.44
Any institutional utilization[1]	t(10195) = 3.39***	t(10126) = 6.73***	t(9965) = 4.84***
Some utilization	15.07	15.61	15.62
No utilization	14.72	14.90	15.10
Any Part B utilization[1]	t(10195) = -12.29***	t(10126) = -11.79***	t(1130) = -10.09***
Some utilization	15.17	15.58	15.63
No utilization	13.28	13.63	13.85
Total reimbursed dollars[1]	F(3,10174) = 48.70***	F(3,10108) = 51.10***	F(3,9946) = 43.17***
$0	11.95	13.84	14.01
$1-$499	13.52	15.34	15.41
$500-$4,999	14.97	15.86	15.88
$5,000 or more	15.43	15.38	15.37
Covered institutional charges[1]	F(3,10193) = 3.83***	F(3,10124) = 15.78***	F(3,9963) = 10.85***
$0	14.72	14.90	15.10
$1-$499	15.05	15.72	15.81
$500-$4,999	15.11	15.61	15.65
$5,000 or more	15.04	15.50	15.36

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[1]Respondents who were enrolled in managed care during the past year were excluded from these analyses.

Table A.9
Mean perceived knowledge index scores and statistical significance of ANOVAs and t-tests among proxy respondents

Variable	1998	1999	2000
Education	F(2,1358) = 18.79***	F(2,1339) = 10.49***	F(2,1262) = 17.98***
8th grade or less	13.88	14.36	14.31
More than 8th grade, but no college	15.22	15.36	15.33
College	16.17	16.37	16.93
Age category	t(903) = 0.86	t(921) = 1.15	t(849) = 0.26
65-75	15.38	15.90	15.79
Over 75	15.09	15.48	15.70
Income category	t(1332) = –7.46***	t(1340) = –6.91***	t(440) = –9.16***
$25,000 or less	14.28	14.63	14.70
More than $25,000	16.85	17.13	17.68
Managed care	t(1359) = –1.71	t(1380) = –1.02	t(1302) = –1.70
Some enrollment	15.42	15.52	15.88
No enrollment	14.70	15.09	15.14
Private Supplemental Insurance	t(1359) = –5.81***	t(179) = –1.51	t(1295) = 7.56***
Supplemental insurance	15.99	15.64	16.55
No supplemental insurance	14.27	14.23	14.36
Any institutional utilization[1]	t(1192) = 2.45*	t(1197) = 1.88	t(1139) = 1.57
Some utilization	14.98	15.30	15.32
No utilization	14.24	14.70	14.81
Any Part B utilization[1]	t(1192) = –4.25***	t(1197) = –3.70***	t(166) = –4.72***
Some utilization	14.97	15.32	15.43
No utilization	13.25	13.72	12.99
Total reimbursed dollars[1]	F(3,1182) = 5.23**	F(3,1191) = 5.20**	F(3,1129) = 9.18***
$0	13.55	13.93	13.28
$1-$499	14.55	14.88	15.02
$500-$4,999	15.24	15.35	15.72
$5,000 or more	14.91	15.82	15.59
Covered institutional charges[1]	F(3,1190) = 2.95*	F(3,1195) = 3.22*	F(3,1137) = 2.25
$0	14.23	14.70	14.81
$1-$499	14.90	14.90	14.70
$500-$4,999	15.36	14.96	15.65
$5,000 or more	14.67	15.91	15.48

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[1]Respondents who were enrolled in managed care during the past year were excluded from these analyses.

Table A.10
Percentage of correct responses by educational achievement for nine-item quiz questions in 1999

Question/education level	Percent Correct (%)
Supplemental insurance is the same as managed care	43%
No formal education	4%
8th grade or less	27%
9th to 12th grade, no high school diploma	35%
High school graduate	46%
College	56%
Medicare covers an annual flu shot	84%
No formal education	71%
8th grade or less	80%
9th to 12th grade, no high school diploma	82%
High school graduate	85%
College	86%
Can select different health plan options	44%
No formal education	13%
8th grade or less	32%
9th to 12th grade, no high school diploma	40%
High school graduate	45%
College	51%
Medicare alone pays for all health care expenses	78%
No formal education	44%
8th grade or less	65%
9th to 12th grade, no high school diploma	74%
High school graduate	81%
College	85%
Medicare offers more information	52%
No formal education	26%
8th grade or less	40%
9th to 12th grade, no high school diploma	45%
High school graduate	55%
College	58%

(continued)

Table A.10
(continued)

Question/education level	Percent Correct (%)
Can report complaints to Medicare about HMOs and supplemental insurance	56%
No formal education	24%
8th grade or less	46%
9th to 12th grade, no high school diploma	52%
High school graduate	60%
College	60%
Limited choices of doctors if on HMOs	61%
No formal education	24%
8th grade or less	44%
9th to 12th grade, no high school diploma	59%
High school graduate	63%
College	69%
Can drop HMO and still be covered by Medicare	46%
No formal education	7%
8th grade or less	28%
9th to 12th grade, no high school diploma	43%
High school graduate	48%
College	56%
HMOs cover more health services	37%
No formal education	6%
8th grade or less	23%
9th to 12th grade, no high school diploma	30%
High school graduate	39%
College	46%

SOURCE: Centers for Medicare & Medicaid Services, MCBS Supplemental Rounds BN 27 and BK 26.

Table A.11
Coefficient alphas of the nine-item quiz, by enrollment in managed care during the past year

Interview type/survey year	No enrollment	Some enrollment
Sample member interviews		
1998 (Round 23)	0.74	0.68
1999 (Round 26)	0.72	0.70
2000 (Round 29)	0.78	0.74
Proxy interviews		
1998 (Round 23)	0.76	0.75
1999 (Round 26)	0.74	0.73
2000 (Round 29)	0.80	0.81

Table A.12
Coefficient alphas of the nine-item quiz, by institutional utilization[1]

Interview type/survey year	No utilization	Some utilization
Sample member interviews		
1998 (Round 23)	0.75	0.73
1999 (Round 26)	0.73	0.71
2000 (Round 29)	0.80	0.78
Proxy interviews		
1998 (Round 23)	0.78	0.75
1999 (Round 26)	0.72	0.75
2000 (Round 29)	0.82	0.79

[1]Only respondents who were not enrolled in managed care during the past year were included in these analyses.

Table A.13
Coefficient alphas of the nine-item quiz, by part b utilization[1]

Interview type/survey year	No utilization	Some utilization
Sample member interviews		
1998 (Round 23)	0.77	0.73
1999 (Round 26)	0.75	0.71
2000 (Round 29)	0.83	0.78
Proxy interviews		
1998 (Round 23)	0.79	0.76
1999 (Round 26)	0.71	0.75
2000 (Round 29)	0.83	0.80

[1]Only respondents who were not enrolled in managed care during the past year were included in these analyses.

Table A.14
Coefficient alphas of the nine-item quiz, by total reimbursed dollars[1]

Interview type/survey year	$0	$1 – $499	$500 – $4,999	$5,000 or more
Sample member interviews				
1998 (Round 23)	0.76	0.73	0.73	0.75
1999 (Round 26)	0.73	0.71	0.72	0.73
2000 (Round 29)	0.83	0.77	0.77	0.79
Proxy interviews				
1998 (Round 23)	0.78	0.75	0.77	0.76
1999 (Round 26)	0.69	0.75	0.73	0.76
2000 (Round 29)	0.82	0.81	0.80	0.78

[1]Only respondents who were not enrolled in managed care during the past year were included in these analyses.

Table A.15
Coefficient alphas of the nine-item quiz, by covered institutional charges[1]

Interview type/survey year	$0	$1 – $499	$500 – $4,999	$5,000 or more
Sample member interviews				
1998 (Round 23)	0.75	0.73	0.73	0.74
1999 (Round 26)	0.74	0.71	0.70	0.72
2000 (Round 29)	0.80	0.76	0.78	0.79
Proxy interviews				
1998 (Round 23)	0.78	0.73	0.77	0.75
1999 (Round 26)	0.72	0.79	0.73	0.74
2000 (Round 29)	0.82	0.81	0.79	0.78

[1]Only respondents who were not enrolled in managed care during the past year were included in these analyses.

Table A.16
Mean nine-item quiz scores and statistical significance of ANOVAS and t-tests among sample members

Variable	1998	1999	2000
Education	F(2,13172) = 459.53***	F(2,3960) = 135.87***	F(2,12890) = 517.47***
8th grade or less	3.85	3.61	3.97
More than 8th grade, but no college	4.94	4.87	5.27
College	5.66	5.50	6.13
Age category	t(11223) = 13.49***	t(3252) = 8.67***	t(10952) = 13.73***
65-75	5.36	5.32	5.77
Over 75	4.75	4.60	5.12
Income category	t(8477) = –25.76***	t(3079) = –14.30***	t(10282) = –28.20***
$25,000 or less	4.60	4.47	4.89
More than $25,000	5.73	5.59	6.15
Managed care	t(3833) = –23.00***	t(3974) = –12.06***	t(4286) = –22.91***
Some enrollment	5.90	5.78	5.10
No enrollment	4.73	4.64	6.29
Private Supplemental Insurance	t(12263) = –20.40***	t(2069) = –8.34***	t(10051) = 15.59***
Enrolled	5.44	5.31	5.63
Not enrolled	4.58	4.43	4.91
Any institutional utilization[1]	t(7096) = 0.81	t(3194) = 2.21*	t(5797) = 1.25
Some utilization	4.74	4.70	5.12
No utilization	4.70	4.50	5.05
Any Part B utilization[1]	t(1524) = –9.90***	t(3194) = –4.01***	t(1152) = –6.04***
Some utilization	4.81	4.70	5.15
No utilization	4.04	4.16	4.59
Total reimbursed dollars[1]	F(3,10731) = 35.93***	F(3,3192) = 9.14***	F(3,10344) = 23.93***
$0	4.18	4.11	4.61
$1-$499	4.82	4.72	5.15
$500-$4,999	4.91	4.79	5.27
$5,000 or more	4.57	4.51	4.93
Covered institutional charges[1]	F(3,10752) = 4.75**	F(3,3192) = 3.62*	F(3,10362) = 6.72***
$0	4.70	4.51	5.05
$1-$499	4.87	4.88	5.30
$500-$4,999	4.76	4.69	5.10
$5,000 or more	4.60	4.54	4.97

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[1]Respondents who were enrolled in managed care during the past year were excluded from these analyses.

Table A.17
Mean nine-item quiz scores and statistical significance of ANOVAS and t-tests among proxy respondents

Variable	1998	1999	2000
Education	F(2,1492) = 28.76***	F(2,406) = 6.91**	F(2,1345) = 14.47***
8th grade or less	4.05	4.11	4.58
More than 8th grade, but no college	4.95	4.68	5.18
College	5.28	5.53	5.68
Age category	t(1010) = 0.24	t(287) = 1.27	t(926) = 0.31
65-75	4.77	5.00	5.24
Over 75	4.73	4.63	5.18
Income category	t(489) = –9.53***	t(395) = –5.43***	t(466) = –6.74***
$25,000 or less	4.33	4.24	4.78
More than $25,000	5.78	5.86	5.92
Managed care	t(1493) = –4.64***	t(420) = –4.51***	t(1388) = –6.65***
Some enrollment	5.44	5.83	6.29
No enrollment	4.50	4.34	4.84
Private Supplemental Insurance	t(957) = –7.63***	t(201) = –3.19**	t(1279) = 4.95***
Enrolled	4.30	5.61	5.44
Not enrolled	5.34	4.33	4.72
Any institutional utilization[1]	t(1310) = 0.64	t(356) = 0.66	t(1219) = 0.50
Some utilization	4.54	4.40	4.87
No utilization	4.45	4.22	4.79
Any Part B utilization[1]	t(1310) = –3.13**	t(356) = –2.50*	t(1219) = –2.52*
Some utilization	4.60	4.48	4.91
No utilization	3.98	3.59	4.30
Total reimbursed dollars[1]	F(3,1299) = 2.84*	F(3,352) = 3.11*	F(3,1209) = 2.83*
$0	4.20	3.57	4.41
$1-$499	4.41	4.59	4.70
$500-$4,999	4.77	4.29	5.05
$5,000 or more	4.43	4.69	4.96
Covered institutional charges[1]	F(3,1308) = 0.38	F(3,354) = 1.02	F(3,1217) = 2.46
$0	4.45	4.22	4.79
$1-$499	4.43	4.11	4.52
$500-$4,999	4.63	4.27	5.17
$5,000 or more	4.51	4.70	4.85

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[1]Respondents who were enrolled in managed care during the past year were excluded from these analyses.