Differences in the Structure of CAHPS® Measures Among the Medicare Fee-for-Service, Medicare Managed Care, and Privately Insured Populations

Randall H. Bender and Steven A. Garfinkel

Objective. To confirm in a new population, the Medicare fee-for-service population, the factor structure previously found in two Consumer Assessment of Health Plans Study (CAHPS®) field-test surveys with Medicare HMO and adult privately insured populations.

Data Sources. Primary data were collected in the fall of 1998. Survey responses from the Medicare Fee-for-Service CAHPS survey field test were compared to results from the Medicare HMO and adult privately insured field-test studies conducted in the fall of 1996.

Study Design. Respondents for the field-test survey were a random sample of Medicare beneficiaries in five states who had opted for the original Medicare plan (fee-for-service).

Data Collection. Data were collected by a mailed survey with a telephone follow-up survey to those who did not return the mailed survey.

Principal Findings. A confirmatory factor analysis in two different samples of Medicare fee-for-service beneficiaries provided basic support for a previously reported three-factor structure underlying the CAHPS reports and rating items: (1) quality of provider or staff communications; (2) timely access to quality health care; and (3) quality of plan administration. An exploratory factor analysis revealed a variant three-factor structure.

Conclusion. Because of differences in the factor structures among the different populations discussed, caution needs to be exercised in any composite development, based on factor analysis or any other basis, by which cross-population comparisons will be made. Comparisons should only be made on composites representing stable structure across all populations concerned.

Key Words. CAHPS, consumer assessment, factor analysis, health care quality, Medicare

The Consumer Assessment of Health Plans Study (CAHPS®) approach to measuring health plan performance by surveying beneficiaries and reporting the results to prospective beneficiaries, purchasers, and health plan managers has become the standard adopted by the Medicare program, many Medicaid programs, numerous private employers, and all health plans that seek accreditation from the National Committee on Quality Assurance. This research compares the underlying statistical structure of CAHPS measures in different populations defined by age and health plan membership to assess one aspect of the validity of comparisons using CAHPS measures.

The goal in providing CAHPS measures to prospective beneficiaries is to help them choose among competing health plans. The goal for public and private purchasers is to assist them in selecting among competing plans and ensure that different subgroups in their constituent populations are being treated equitably. The goal for health plan managers is to enable them to improve their overall performance by identifying variation among beneficiary populations.

For the Medicare program, comparisons of enrollees in Medicare+ Choice managed care plans with original Medicare beneficiaries in the same locale are a key element in ensuring that beneficiaries have the information they need to make an informed choice. For many private employers and health plan managers who cover both active and retired employees, it is important to know if the two groups are reporting different experiences.

Many statistical issues affect the validity of CAHPS comparisons among plans and populations, but the consistency of the underlying factor structure of composite CAHPS measures is one of the most important. If the same composite measure is being compared for two or more plans or populations, we need some assurance that the relative contribution of the component survey questions to the composite measure is similar for each group. This report summarizes findings from confirmatory factor analysis (CFA) and exploratory factor analysis (EFA) of CAHPS field-test data for three populations that will be compared often as the use of CAHPS becomes more common.

The CAHPS Medicare fee-for-service (M-FFS) field test conducted for the Health Care Financing Administration by Research Triangle Institute

---

Address correspondence to Randall H. Bender, Ph.D., Psychometrician, 3400 Cornwallis Road, P.O. Box 12194, Research Triangle Institute, Research Triangle Park, NC 27709-2194. Steven A. Garfinkel, Ph.D. is Director, Health Services and Policy Research Program, Research Triangle Institute. This article, submitted to *Health Services Research* on April 24, 2000, was revised and accepted for publication on October 25, 2000.

tested survey field procedures and provided data with which to evaluate the psychometric performance of standard CAHPS items in an FFS population. As part of this evaluation (for the complete report see Carman, Keller, and Hays 1999), the underlying factor structure of this population's responses was examined to see if it mirrored that of previously studied populations.

Factor analysis is a statistical methodology that examines a set of response data with a goal of providing a parsimonious account for a set of observed variables with a much smaller set of factors. Such factors may represent attitudes, values, beliefs, or other constructs that are postulated to account for the common association found in a group of items. Here, the underlying factors are content defined, relating to common experiences. For example, the general experience patients have with their doctors drives their responses to the interpersonal communications items.

An EFA often begins with virtually no preconceived model of what, or even exactly how many, factors underlie a set of observed variables. The analyst works interactively with the optimization algorithms to discover the best-fitting, interpretable model. CFA, on the other hand, applies a preconceived factor model to a set of variables to test the fit of the model.

There are two primary reasons to undertake factor analysis in this study. Reporting authorities would like to report the results of CAHPS survey items using a set of composites instead of single items, and factor analysis could be used to identify appropriate unidimensional composites. (Unidimensionality is highly desirable as it improves a scale's performance.) It is also important to know whether the previously found factor structure is also found in this FFS population. If the items group differently, this could both affect how the reporting composites are defined and alert analysts to use caution in making comparisons based on these composites among these populations.

For example, suppose a "time spent with doctor" item was found to relate to other provider variables in an FFS population, but in an HMO population it is found to relate to plan variables. If the FFS structure were then imposed for the purposes of forming reporting composites, any differences between the two populations on the provider composite could actually be related to differences in the underlying factor structure. If the plan items were particularly low, it could partially pull down the provider composite in the HMO population. In such a case individual item reporting is more legitimate.

Because the FFS and managed care delivery systems differ, there is reason for concern about the underlying factor structure; some comparisons will undoubtedly be desirable. Therefore, it is important to undertake a comparison of the underlying factor structures in these populations.

METHODS

The data in these analyses came from a survey conducted between August and December of 1998 on a sample of 3,885 Medicare beneficiaries who were not currently enrolled in Medicare risk- or cost-contract health plans. A sample of Medicare beneficiaries in five states (Pennsylvania, Georgia, Wisconsin, Arizona, and Washington) received surveys in the mail. A total of 2,352 individuals completed surveys (66.1 percent response rate),[1] that is, they indicated that they were on Medicare and answered at least 80 percent of the applicable items.

An alternative form (form B) was fielded to test some new and modified items as well as a different recall period (six months for form A, 12 months for form B) for some of the items. Sixteen percent ($n = 381$) of the respondents received form B, whereas 84 percent ($n = 1{,}971$) received form A. Stratified sampling was used to randomly assign respondents to the two samples while keeping them equivalent on the three demographic attributes gender, race, and age.

The questionnaires used in the M-FFS field test contained items designed to (1) determine if the sample member had one personal health care provider (a doctor or other health care professional) and to obtain his/her rating of that provider; (2) determine if the sample member had used a specialist and, if so, to determine his/her rating of that specialist; (3) obtain information about the sample member's experience with getting care from a specialist; (4) obtain information about the sample member's need for and experience with other types of medical services, such as home nursing or assistance, prescription medicines, etc.; (5) determine the sample member's utilization of other types of health care services and experience in getting those services; (6) obtain the sample member's rating of all health care services provided during the reference period; (7) obtain information about the sample member's experience with the health plan, such as problems with claims, customer service, understanding information about the plan, and determining whether the sample member would recommend the plan to other family members and friends; and (8) obtain demographic information.

The survey was conducted by mail with a telephone follow-up. The protocol included a prenotification letter, an initial survey package followed a week later by a thank you/reminder letter, and a second package sent to all nonrespondents followed by a telephone follow-up with further nonrespondents or a third follow-up package for nonrespondents for whom telephone numbers could not be obtained. The telephone follow-up contributed 9.7 percent to the overall response rate.

The two previous studies were conducted in 1996. The Medicare HMO study analyzed survey field-test data from 965 Medicare beneficiaries in Medicare managed care plans using an earlier version of the CAHPS items. That study was conducted by mail with a telephone follow-up, similar to the protocol for our study. The adult privately insured field-test survey was a telephone survey of 954 (539 completed) members of private health plans and also used an earlier version of the CAHPS items. The majority of the adult privately insured sample and all of the Medicare HMO sample were in some kind of managed care plan; thus, their health care delivery differed substantially from that of the current study's health plan members. In comparing these studies, the adult privately insured and Medicare HMO populations will be considered together as representing primarily managed care populations.

Uses of Factor Analysis for Analyzing Survey Data

Factor analyses were conducted to address three questions about the CAHPS M-FFS questionnaire:

1. Is the factor structure suggested by the M-FFS item responses similar to that found in previous surveys, the Medicare HMO (Sweeny, Williams, Hays, et al. 1997) and adult privately insured field-test surveys (Williams, Burnbauer, Lubalin, et al. 1997), that used similar CAHPS items?
2. Is the factor structure underlying responses equivalent across forms A and B?
3. What underlying structure might be found in a pure EFA of the M-FFS survey data, independent of the other surveys?

The first two study questions are best answered by CFA, which can test whether a specified factor structure fits in a given sample. EFA is the obvious technique to use with the third question.

To test a prespecified factor structure model with CFA one must specify both the number of factors and the factors the items are expected to load on, that is, be most strongly related to. Generally the specified structure loads each item on only one factor, a structure known as "simple structure" (Thurstone 1935). This structure is conceptually clear because each item only represents one underlying construct. In some cases interfactor correlations may be specified, and in a few cases actual values for factor loadings may also be specified. Generally one will simply specify whether factors will be allowed to correlate depending on whether the researcher believes the underlying latent factors represent constructs that should be correlated. The model will estimate

from the data the item factor loadings (how strongly/weakly a given item is related to an underlying factor), the interfactor correlations (how related the underlying constructs are to each other), and the unique variances (the portion of variance among the responses to each item that cannot be accounted for by the hypothesized factor structure). A multivariate statistical model assuming multivariate normal error (Jöreskog and Sörbom 1985) is used to estimate and test the fit of the predesignated factor structure.

EFA has been used with previous CAHPS field-test studies (Williams, Burnbauer, Lubalin, et al. 1997; Sweeny, Williams, Hays, et al. 1997) and was used here to uncover the latent factor structure behind the CAHPS items. This approach requires some judgments on the part of the analyst regarding how many underlying factors (dimensions or constructs) to allow in the structure, whether to allow these factors to correlate, and what statistical method to use in extracting factors. Two shortcomings of this approach are that it (1) does not provide any test of fit for that factor structure, and (2) capitalizes on any random variation and covariation found in the data, whether reliable or meaningful or not. For this reason the best approach, when the sample is large enough, is to use half of the sample to explore the underlying factor structure and test the generalizability of the structure discovered by testing its fit on the reserved half of the sample using CFA.

The Expected Factor Structure

Many of the items in the M-FFS questionnaire had been used in the previous studies with the Medicare HMO and adult privately insured populations and, with some minor exceptions, both studies had roughly the same factor structure. The first factor was called "health care quality" in the adult privately insured study and "perceived provider effectiveness" in the Medicare HMO study. This "provider" factor included items that relate to health care provider communication (e.g., the doctor explained things to me, listened to me, spent time with me, treated me with courtesy and respect). The second factor was called "getting health care when you need it" in the adult privately insured study and "perceived access to care" in the Medicare HMO study. This "access" factor addressed the patient's ability to access timely health care and includes various items addressing timely appointments, tests, and treatment. The third factor was called "quality of health care plan" in the adult privately insured study and "perceived plan responsiveness" in the Medicare HMO report. This "plan" factor incorporated patient perceptions of plan-related issues such as the amount of paperwork, customer service, timeliness of decisions, treatment refusals, and referrals to specialists.

When the two previous studies did not agree on item placement in the factor structure, decisions were based on substantive judgment and were carefully checked with alternative placements in the model-building phase. The same process was used with new CAHPS items 40 (handle claims in a reasonable time), 41 (handle claims correctly), 43 (problem finding information in written materials), and 45 (problem getting help from customer service). Placing these on the plan factor was logical, but other placements were checked in the model-building phase. This approach generated the hypothesized M-FFS factor model shown in Figure 1.

For the reasons just indicated and because this survey's population differed from either of the two previous studies, we needed to check the factor structure. To initially explore the structure on part of the sample while still preserving part of it to test the fit of our model, we randomly split our sample into two halves: a model-building sample used to identify any needed modifications, and a model-testing sample used to test the generalizability of such modifications.

In the model-building procedures we fit the model and examined the estimated model parameters and residual information to determine whether the prespecified factor structure needed to be changed. Items with low factor loadings, or high residual correlations with items on other factors, were moved to see if they fit better on another factor. If such a move resulted in better loading, this was always followed by a check of the overall model fit (it should improve) and overall model structure (it should not change dramatically).

This procedure resulted in moving items 8 (problem with referral to a specialist) and 10 (rate specialist) from factor III (plan) to factor II (access). When the model was estimated[2] with these modifications, the fit improved (see Table 1) and the factor loadings for the shifted items were substantially higher (see Table 2). Table 1 contains some goodness-of-fit indicators[3] for this model on the model-building sample. This adjusted model, M-FFS, would be considered a marginally good-fitting model.

Confirmatory Factor Analysis: Testing the Model

To test for generalizability the M-FFS model structure was fit to the model-testing sample. The fit indices for this sample showed a reasonably good fit, indicating that the model-building procedures had not unduly capitalized on chance.

Table 2 contains the parameter estimates for the model-testing sample. The estimates were generally high (about .6 or higher) and sensible. All of the factor loadings were positive as expected, except item 22, which should be

Figure 1: Hypothesized Factor Structure

The Provider Factor

4. How much problem getting personal doctor with whom you are happy
6. Rate personal doctor
23. How often staff at doctor's office treats you with courtesy
24. How often staff at doctor's office were as helpful as you thought they should be
25. How often doctor listens carefully to you
26. How often doctor explains things in a way you can understand
27. How often doctor shows respect for what you say
28. How often doctor spends enough time with you

The Access Factor

13. How often you get the help you need when you call
15. How often you get an appointment as soon as you want
17. How often you get needed care as soon as you want
20. How much problem you have getting care you believe necessary
21. How much problem you have with delays in health care caused by waiting for approval from health plan
22. How often you wait in the doctor's office for more than 15 minutes past appointment time

The Plan Factor

8. Problem with referral to a specialist in last six months
10. Rate specialist
38. How much problem you have with paperwork for health plan
45. How much problem getting help you need when you call health plan customer service
40. How often health plan handles claims within a reasonable time
41. How often health plan handles claims correctly
43. How much problem finding information in written materials
46. Rate health plan

Table 1: Goodness-of-Fit Indices (GFI) for the M-FFS Model on Various Samples and the Five-Factor Reporting Composites Model

	Model-Building Sample	*Model-Testing Sample*	*Form B Sample*	*CFA of EFA Model*
GFI	.95	.93	.95	.94
GFI adjusted for degrees of freedom	.94	.92	.94	.93
Root mean square residual	.08	.09	.08	.09

negative because of the direction of the scoring for that item. All of the items on factor I had loadings of .6 or higher, the highest being item 25 (doctor listens carefully). Factor II, with four exceptions noted below, had all items loading at .5 or higher; the highest loading was item 17 (got care as soon as wanted). Factor III had all items loading .57 or higher, with the highest being item 38 (problems with paperwork for health plan).

The parameter estimates for the model-testing sample indicate that a few items—20 (problem getting care you believe necessary), 21 (problem with delay in health care caused by waiting for health plan approval), 8 (problem with referral to a specialist),[4] and to some extent 22 (wait more than 15 minutes in waiting room)—do not load well on their specified factors or on any other factor. The association between three of these items (20, 21, and 8) and any factor was limited because of ceiling effects: more than 90 percent of the sample gave the top rating to each. To increase their usefulness these items should be modified to boost response variability.

The estimates of the factor correlations seemed quite reasonable given the features of the FFS health delivery system. The access factor is more highly correlated with the provider factor (.79) than with the plan factor (.46), and the plan factor is not very highly correlated with the provider factor (.32). The correlation between provider and access is so high that the two factors could quite reasonably be combined.

The model was also fit to the sample that received form B of the survey, yielding more proof of the model's generalizability. The factor structure proved to be robust in the face of item editing and a longer reference period (12 months versus 6 months for form A). The factor loadings and correlations for this sample are nearly all higher than those seen on the two form A samples. This may be a result of the longer reference period reducing missing data and thus allowing stronger associations to emerge.

Table 2: Item-Factor Loadings for the M-FFS Model

Item Wording	*I. Model-Building Sample*	*II. Model-Testing Sample*	*III. Form B Sample*
I. Provider Factor			
23. Courtesy of office staff	.61	.63	.44
24. Helpfulness of office staff	.62	.75	.65
25. Careful listening of doctor	.70	.80	.79
26. Understandable doctor explanations	.69	.59	.72
27. Respect from doctor	.73	.75	.82
28. Doctor spends enough time	.73	.73	.74
II. Access Factor			
4. Able to get personal doctor with whom are happy	.54	.63	.89
6. Rate personal doctor	.66	.53	.64
13. Able to get needed help when calling	.81	.59	.80
15. Able to get appointment as soon as want	.56	.51	.55
17. Able to get needed care as soon as want	.78	.69	.61
20. Problems getting care you believe necessary	.56	.41	.51
21. Problems with delays in health care because of wait for approval	.23	.33	.48
22. Wait in doctor's office more than 15 minutes past appointment time	−.46	−.45	−.40
8. Problems with referral to specialist in last six months	.46	.18	.58
10. Rate specialist	.66	.50	.47
III. Plan Factor			
38. Problems with paperwork for health plan	.71	.81	.76
45. Problems getting help needed when calling health plan customer service	.61	.76	.62
40. Claims handled within reasonable time	.66	.72	.64
41. Claims handled correctly	.59	.68	.60
43. Problems finding information in written materials	.55	.57	.53
46. Rate health plan	.64	.76	.70
Factor Correlations			
I. Provider–II. Access	.76	.79	.84
I. Provider–III. Plan	.40	.32	.44
II. Access–III. Plan	.40	.46	.50

Exploratory Factor Analysis of M-FFS Sample

To see if some completely novel factor structure was being missed we also undertook an EFA on the model-building sample. To determine the appropriate number of factors we looked at two popular rules of thumb, the Kaiser-Guttman rule (Guttman 1954; Kaiser 1970) and the scree plot (Gorsuch

1983). The first ten eigenvalues from a principal components analysis were 6.83, 2.47, 1.92, 1.32, 1.00, .91, .86, .79, .70, and .68. This indicates that four or five factors may be required. However, both of these indicators tend to overestimate the appropriate number of factors (Gorsuch 1983; Nunnally and Bernstein 1994), so we fit models with only two and three factors as well.

In examining two- to five-factor models with the model-building sample, we found that the most useful and interpretable result was the three-factor solution. The four-factor solution began to split off factors that were not theoretically significant based solely on item-formatting similarities. The two-factor solution, a split between provider and plan items, was interpretable (the latter being the same plan factor as our M-FFS model), but further factor separation that was both interpretable and substantively useful was possible in the three-factor solution.

Table 3 shows the factor structure and loadings for the preferred three-factor solution. The first factor includes items concerned with the respondents' experiences with their personal doctors and office staff as well as several access items that are under the control of the providers in this delivery system. The largest loading on this factor is the "doctor shows respect" item (27). The reason that the rating of specialist is also on this factor may be because the primary physician is a specialist for a number of the respondents. The second factor loads items concerned with plan-administration issues, with the strongest item being problems with paperwork (38). The third factor includes items related to access, but in a narrower sense. These items seem to mainly be related to gaining access to specialists, led by the strongest item regarding specialist referral (8).

The factor correlations are shown in Table 5. We see that the correlations among the three factors are quite moderate. The provider-access and provider-plan correlations are about .3, whereas the access-plan correlation is about .2; thus, access to health care is a bit more closely correlated with provider than with plan.

To provide some comparison for the fit statistics in the CFA we fit to the model-testing sample the simple structure implied by the structure found in the EFA on the model-building sample. These fit indices (see Table 1) are very much in line with the results of the original CFA fit to the model-building and model-testing samples. This provides additional evidence that the CFA model fits are about as good as one can get for a three-factor simple-structure model because a model whose structure was optimized on this data does not fit dramatically better.

Table 3: EFA: Three-Factor Structure Estimated on the Model-Building Sample

	Factor I	*Factor II*	*Factor III*
6. Rate personal doctor	**.61**	.17	.47
23. Courtesy of office staff	**.61**	.24	.28
24. Helpfulness of office staff	**.62**	.19	.34
25. Careful listening of doctor	**.77**	.31	.13
26. Understandable doctor explanations	**.77**	.18	.17
27. Respect from doctor	**.82**	.13	.20
28. Doctor spends enough time	**.80**	.25	.16
13. Able to get needed help when calling	**.69**	.26	.58
15. Able to get appointment as soon as want	**.49**	.24	.42
22. Wait in doctor's office more than 15 minutes past appointment time	**-.46**	-.19	-.25
10. Rate specialist	**.55**	.32	.51
38. Problems with paperwork for health plan	.18	**.84**	.08
45. Problems getting help needed when calling health plan customer service	.02	**.74**	.43
40. Claims handled within reasonable time	.27	**.72**	.08
41. Claims handled correctly	.29	**.69**	-.07
43. Problems finding information in written materials	.28	**.56**	.17
46. Rate health plan	.29	**.61**	.22
4. Able to get personal doctor with whom are happy	.38	.31	**.55**
17. Able to get needed care as soon as want	.61	.15	**.72**
20. Problems getting care you believe necessary	.37	.07	**.72**
21. Problems with delays in health care because of wait for approval	.03	.00	**.56**
8. Problems with referral to specialist in last six months	.20	.13	**.77**

Note: Highest factor loading for each item is in bold.

DISCUSSION

The factors that emerged for the M-FFS model from the model-building phase have some subtle differences from previous studies. All of the items that remain on the first factor now clearly relate to plan members' interpersonal relations with their providers and office staff. A suitable label might be "quality of provider/staff communications." The second factor now contains all of the items relating to the quality of care received, including issues of access. A good label here might be "quality of and access to health care." The third factor remains clearly the plan factor and could best be labeled "quality of plan administration."

Table 4: Comparison of All Factor Structures*

M-FFS Item Description	*I. Provider*	*II. Access*	*III. Plan*
23. Courtesy of office staff	F E A	H	
24. Helpfulness of office staff	F E		
25. Careful listening of doctor	F E A H		
26. Understandable doctor explanations	F E A		
27. Respect from doctor	F E A		
28. Doctor spends enough time	F E A H		
6. Rate personal doctor	F E ~A ~H		
4. Able to get personal doctor with whom are happy	A	F E H	
8. Problems with referral to specialist in last six months		F E	~A ~H
20. Problems getting care you believe necessary	~H	F E	A
21. Problems with delays in health care because of wait for health plan approval		F E	
17. Able to get needed care as soon as want		F E A H	
13. Able to get needed help when calling	E	F A	
15. Able to get appointment as soon as want	E	F A H	
22. Wait in doctor's office more than 15 minutes past appointment time	E	F A	H
10. Rate specialist	E	F	~A
38. Problems with paperwork for health plan			F E A
45. Problems getting help needed when calling health plan customer service			F E A H
46. Rate health plan			F E A
41. Claims handled correctly			F E ~H
40. Claims handled within reasonable time			F E
43. Problems finding information in written materials			F E

*F = Medicare fee-for-service; H = Medicare HMO; A = adult privately insured; E = indicated EFA three-factor results in this M-FFS study; ~ in front of letter = although the item from the previous study was considered close enough to use for some guidance in determining the hypothesized factor structure, it is not exactly the same wording and in some cases may vary enough in content to be considered a different item.

Factor I labels: M-FFS = quality of interpersonal relations with provider/staff; Medicare HMO = perceived provider effectiveness; adult privately insured = health care quality.

Factor II labels: M-FFS = timely access to quality health care; Medicare HMO = perceived access to care; adult privately insured = getting health care when you need it.

Factor III labels: M-FFS = quality of health plan administration; Medicare HMO = perceived plan responsiveness; adult privately insured = quality of health care plan.

Within each factor items are sorted to facilitate a comparison of the results of the EFA and the CFA.

In the EFA, the first factor is similar to the first factor in our M-FFS model, but its content is a bit broader, including access items that reflect on the provider's quality. The first factor here could probably be best labeled "quality of provider and staff," indicating that it is about more than just communication. The addition of content regarding long waits and problems getting appointments or help on the phone broadens the factor to represent provider quality. The second factor here is identical to the third factor of the M-FFS model (quality of plan administration), but the third factor differs substantially from the second M-FFS model factor, being considerably narrower in substance. A fitting name for this factor might be "access to specialists."

These two different factor structures may lead the reader to wonder which structure is the true underlying factor structure. It would be nice to decide this on the basis of model fit, but with nearly identical model fit results one is hard pressed to declare one structure superior on this basis alone. From that standpoint they are really two equally good ways of looking at these items. However, the structure examined in the CFA was based on that found in two previous studies, and it has now fit fairly well in a third population. That proof of generalization is an important consideration in its favor. The exploratory results have only been fit in this one population and may prove to be less generalizable. Future studies should test both of these structures to see whether either is replicated in other populations or if each population forces subtle modifications of both factor structures. Until the results from such studies are available, we suggest that the structure confirmed by the CFA be considered the best way of understanding the underlying factor structure of these items for the M-FFS population.

Given the results of the two previous studies, the new EFA on this FFS population, and a slight modification of the previously found structure confirmed on this sample, there is a rich set of findings to sift through. Tables 4 and 5 attempt to present all of these findings in a way that allows similarities and differences to emerge and be discussed.

Table 4 shows the differences between the three-factor solution uncovered here, the M-FFS structure, and those found in the two previous studies. Looking over Table 4, one can identify items on which the two previous studies did not agree and see how the M-FFS structure resolved this.

The provider and plan factors are fairly stable across the four analyses and show virtually no disagreement among the studies regarding item placement. In the provider factor, item 23 (office staff courtesy) loaded on the second factor in the EFA of the Medicare HMO field-test sample, but

Table 5: Interfactor Correlations from EFAs of This and Previous Studies

Factor	*I*	*II*	*III*
I. Provider	–		
II. Access	HMO/API/FFS* .44/.27/.32	–	
III. Plan	HMO/API/FFS .38/.39/.27	HMO/API/FFS .31/.27/.19	–

*HMO = Medicare HMO study; API = adult privately insured study; FFS = Medicare fee-for-service study.

that is the sole difference. In the HMO delivery system compared to the FFS system, office staff are less closely tied to providers. The plan factor shows no item placement disagreements.

The access factor, however, shows the most disagreement about item placement among the four studies. Four of the items—8 (specialist referral), 20 (getting the care you believe necessary), 22 (excessive waiting room waits), and 10 (specialist rating)—were associated with the plan factor in one or both of the two primarily managed care populations. Although these items are for the most part access related, the health care delivery plans featured in those two samples exercised more control over these aspects of the patients' care than the primary physician did, and patients' ratings of those items tended to correlate more with other plan-related (plan-controlled) items. The specialist rating item did not appear in the previous studies. However, in the adult privately insured study there was a question about whether specialist care met the patient's needs, which fell on the plan factor.

Two items, 4 (problems getting a satisfactory doctor) and 20 (getting the care you believe necessary), were associated with the provider factors in one of the two managed care populations. However, the provider factor in the adult privately insured sample is less focused, as its label "health care quality" clearly indicates. The placement of item 20 in the Medicare HMO sample may be because of content differences, as that survey's item was not worded the same.

Understanding the differences between the studies requires a close look at the factor labels and some knowledge of the different item wordings. For example, it may appear that the adult privately insured study disagrees with the other two studies about item 4's factor placement, but that study placed the access items that are more concerned with quality on its first factor. Therefore, there is really no disagreement on this item. Comparisons among these studies

are complicated because item wording was sometimes changed, but more significantly the sets of items included differed. Future comparisons of factor structure between populations would be helped by using the same sets of identical items.

Finally, a word about the comparison of the factor structure from the exploratory analysis and that of the M-FFS. The EFA results follow the M-FFS model closely, except the EFA associates four of the items—13 (get the help needed when you called), 15 (get a timely appointment), 22 (excessive waiting room waits), and 10 (specialist rating)—with the provider factor items. In the FFS population, access issues related to a primary physician correlate with all of the other provider items, creating a broader provider quality factor. Left in the access factor for the EFA model is a narrower access factor primarily focusing on access/referral to specialists. Although this structure supports a somewhat different conceptualization of the underlying constructs, the generalization of this structure to other populations has not been tested; therefore, it should be replicated in other samples before placing too much confidence in it.

Turning next to the parameter estimates of factor correlations (see Table 5), one must exercise caution in comparing estimates between these studies, partly because survey differences resulted in factors that have slightly different meanings. In addition, comparing these factor correlations means comparing EFA and CFA results, which is not straightforward because confirmatory models generally impose a simple structure, whereas exploratory analyses generally rotate their factors to only rather distant approximations to simple structure. This means that although we would like to compare the factor correlations for the structure confirmed here on the M-FFS to those found in exploratory analyses in the adult privately insured and Medicare HMO studies, we are unable to do so directly. We could more directly compare the results of the EFAs conducted in all three studies, but survey differences would still complicate this. That said, we will draw what cautious conclusions we can in this situation.

Table 5 shows the rotated factor correlations from the EFA of the sample in this study along with those from the two previous studies. The provider-access factor correlation is somewhat lower in the FFS than the average for the other two, although it falls between them. The shift of several items from the access to provider factor may have reduced this correlation. The provider-plan factor correlation is lower and may reflect a less formal relationship between the respondents' providers and their health care plans

in the FFS delivery system compared to the managed care delivery systems of the previous two studies.

Because the M-FFS factor structure estimated on this sample is so close to those from the two previous studies, we would like to directly compare the factor correlations from the CFA we undertook (see Table 2) to those found in the two previous studies (see Table 5). In view of the comparison difficulties mentioned above, it would be safer to restrict ourselves to comparing, between studies, ratios of correlations within given factor structures. The most interesting ratio is that of the provider-access correlation to plan-access correlation, which is about 1.7 in the M-FFS population and about 1.2 on average in the two managed care populations. This would seem to indicate that access is more closely correlated with provider than with plan in the FFS population than is true in the managed care populations. The latter populations show less relative correlation between the provider and access factors compared to plan and access factors.

CONCLUSIONS

Clearly differences in the factor structure have been found between the managed care populations and the FFS population. Fee-for-service populations have different views from managed care populations of the responsibilities and scopes of influence that providers and plans exercise in delivering their health care and determining access to care. In the FFS population the provider is the locus of the member's health care. The provider is ultimately responsible for all office operations, policies, procedures, and staffing. The provider tends to be the avenue of referral to specialists, although members of this system can access specialists independently. The plan does constrain some access to care insofar as it limits its reimbursement coverage, but options are generally available at a price that allows the member to make the coverage decisions him/herself when signing up for a plan. Thus, the plan is mostly tied up with reimbursement and paperwork concerns, whereas the provider is generally the locus of access and quality of health care concerns.

In managed care systems, access to specialists is controlled by the plan via the gatekeeper provider. The plan in some cases controls the practice in many ways: appointment length, hours of accessibility, and office staffing. Also, the plan determines coverage for additional benefits such as pharmacy benefits or home health care coverage. For the managed care respondent all of these access issues become more tightly associated, and the whole

group is highly associated with other non-access-related plan items. Clearly members of these two major types of delivery system are conditioned by their experience to think about their health care in different ways, focusing on different issues and differentially associating aspects of their health care.

These differences in how respondents in these two populations think about their health care have real effects on the way they evaluate their health care experiences. While an overall framework of three areas—provider, access, and plan—seems to hold for the CAHPS items in all of these populations, the individual placement of items on factors representing these three constructs differs in some cases among these populations. Yet despite these differences in factor structure, comparisons on CAHPS measures will no doubt still be made between these populations. So how can they be done most appropriately given these differences?

First, any composite (or subscale) scores reported must accurately reflect the label given to the composite in all populations compared. This means composites should be restricted to portions of the factor structure on which all of the populations agree. In this case the items related to provider interpersonal communication operate in a similar way in all of the populations examined in this article, as did the items related to how well a plan communicates with plan members and handles claims and paperwork. These two groups could be reported in composites with little concern that these comparisons across populations would not be valid. However, items addressing access to care, specialist access, relations to office staff, appointments, and delay of care relate differently to the underlying constructs discussed and are best compared individually when comparisons between managed care and FFS options are presented. For the most part, comparisons can always be made, but the most legitimate mode of doing so should be carefully thought out. Any other analyses along these lines that can shed further light on the factor structures operating in these or other related populations should be carefully examined for additional insights that can be used in planning fair comparisons.

This conclusion raises difficult issues for reporting because it implies that more measures should ideally be reported when comparisons are being made between FFS and managed care options than when comparisons are limited to managed care plans only. The dilemma stems from constraints on the additional resources needed to publish more measures and from the increased cognitive burden that additional measures impose on Medicare beneficiaries who are trying to make use of the CAHPS information in decision making.

These conclusions are based on data from field tests conducted with a few purposively selected states and health plans. Now that the Health Care

Financing Administration is conducting nationwide CAHPS surveys of all Medicare HMOs and of FFS beneficiaries in all areas of the country, this analysis should be repeated on national data over several years to confirm these factor structures and conclusions. If the national data confirm significant differences between the Medicare managed care and FFS factor structures, trade-offs between increasing the consistency of the comparative data and increasing the cognitive burden on already overburdened Medicare beneficiaries will have to be made.

ACKNOWLEDGMENTS

Thanks to Kristin Carmen, Sheri Fehnel, Lauren Harris-Kojetin, and San Keller, whose comments and insights helped to much improve this article.

NOTES

1. Response rate equaled the number of completed surveys by mail and phone divided by the total number of eligible respondents in the drawn sample.
2. All of the CFAs reported here were estimated using Proc Calis in SAS using an unweighted least squares method applied to a correlation matrix that was constructed using pairwise deletion for missing data. These are not all ideal approaches (Long 1991); however, the large amounts of missing data required the use of pairwise deletion, yielding a correlation matrix that was not compatible with maximum likelihood procedures.
3. There is no single accepted standard for providing goodness-of-fit information; thus, popular estimation programs provide up to 20 different measures of fit. The article includes several measures to provide some sense of the overall fit of the model. The goodness-of-fit index (GFI) is a widely used indicator of how well the model-predicted variance-covariance (correlation) matrix matches that of the observed data. The adjusted goodness-of-fit index (AGFI) is similar but adjusts for parsimony in the model. The root mean square residual (RMR) is the square root of the mean squared residual for the predicted correlation matrix. The GFI and AGFI generally range from 0 to 1 (negative values are evidently possible), with a fit of 1.00 a perfect fit. Generally fits of .90 or higher are considered good (Bentler and Bonett 1980; Hu and Bentler 1995; Schumacker and Lomax 1996). The RMR ranges from 0 to 1 when dealing with a correlation matrix, with a value of 0.00 indicating perfect fit. Rules of thumb for this range from "a small value" to .05 for a good fit (Hu and Bentler 1995; Schumacker and Lomax 1996).
4. Note that we fit an additional model with item 8 back on the plan factor. Although item 8 (referral to specialist) does not load well on the access factor, it loaded there nearly four times stronger than on the plan factor.

REFERENCES

Bentler, P. M., and D. G. Bonett. 1980. "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures." *Psychological Bulletin* 88: 588–606.

Carman, K. L., S. D. Keller, and R. D. Hays. 1999. *Consumer Assessment of Health Plans Study (CAHPS): Medicare Fee-for-service Field Test Analysis Report.* Washington, DC: Research Triangle Institute.

Gorsuch, R. L. 1983. *Factor Analysis, 2nd ed.* Mahwah, NJ: Lawrence Erlbaum.

Guttman, L. 1954. "Some Necessary Conditions for Common Factor Analysis." *Psychometrika* 19 (2): 149–61.

Hu, L., and P. M. Bentler. 1995. "Evaluating Model Fit." In *Structural Equation Modeling,* edited by R. H. Hoyle, 76–99. Newbury Park, CA: Sage Publications.

Jöreskog, K. G., and D. Sörbom. 1985. *LISREL VI: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables, and Least Squares.* Uppsala: University of Uppsala.

Kaiser, H. F. 1970. "A Second Generation Little Jiffy." *Psychometrika* 35: 401–17.

Long, J. S. 1991. *Confirmatory Factor Analysis.* Newbury Park, CA: Sage Publications.

Nunnally, J. C., and I. H. Bernstein. 1994. *Psychometric Theory, 3rd ed.* New York: McGraw-Hill.

Schumacker, R. E., and R. G. Lomax. 1996. *A Beginners Guide to Structural Equation Modeling.* Mahwah, NJ: Lawrence Erlbaum.

Sweeny, S. F., V. S. L. Williams, R. D. Hays, J. Schnaier, J. Lubalin, and D. Hewitt. 1997. *Consumer Assessment of Health Plans Study (CAHPS): Medicare HMO Field Test Analysis Report.* Washington, DC: Research Triangle Institute.

Thurstone, L. L. 1935. *The Vectors of the Mind.* Chicago: University of Chicago Press.

Williams, V. S. L., L. Burnbauer, J. Lubalin, S. F. Sweeny, M. Ardini, and J. Chromy. 1997. *Consumer Assessment of Health Plans Study (CAHPS): Adult Privately Insured Survey (API): Field Test Analysis Report.* Washington, DC: Research Triangle Institute.