

Scoring Manual for the Primary Care Satisfaction Survey for Women (PCSSW)

February, 2004

Center on Women's Healthcare Quality Assessment
2000 West First Street, Winston-Salem, North Carolina 27104

Scoring Manual for the Primary Care

Satisfaction Survey for Women

Table of Contents

| <i>Heading</i> | <i>Page</i> |
|---|-------------|
| Abstract | 3 |
| Introduction | 5 |
| Summary of the Scale Development | 6 |
| Validation (Original Sample) | 7 |
| Sample | 7 |
| Explained Variance | 7 |
| Confirmatory Factor Analysis | 8 |
| Construct Validity | 8 |
| IRT | 11 |
| Results (Validation Sample 1) | 13 |
| Exploratory Factor Analysis | 13 |
| Convergent Validity | 14 |
| Discriminant Validity | 14 |
| Predictive Validity | 15 |
| Contribution of the PCSSW to Overall Satisfaction | 15 |
| Ratings | |
| Latent Variable Modeling | 15 |
| Cross-Validation (Validation Samples 1 and 2) | 16 |
| Sample | 16 |
| Distributional Sensitivity in Samples 1 and 2 | 17 |
| Principal Component Analysis | 17 |
| Confirmatory Factor Analysis | 17 |
| Reliability | 18 |
| Administration Procedures | 18 |
| Paper and Pencil | 18 |
| Computer-Assisted Methods | 18 |
| Web-Based Survey | 18 |
| Multiple Modalities | 19 |
| Norms | 19 |
| Interpretation of scores | 19 |
| Additional Information | 19 |

Abstract

The Primary Care Satisfaction Survey for Women (PCSSW) is a new tool to assess women's satisfaction with primary care. The instrument was developed and tested with research contracts and grants to a consortium of three academic centers from the Department of Health and Human Services, Office on Women's Health; and from the Agency on Healthcare Research and Quality (AHRQ).

The PCSSW is now available for use in health services research and routine quality of care assessment by medical practices. This Manual was developed to guide the use of the instrument and provide key information on its development and validation. A full report of the latter is available in a forthcoming scientific publication by Hudson Scholle, Weisman, Anderson, et al, Women's Health Issues, April, 2004. Benchmarking against the pooled National Centers of Excellence in Women's Health (CoE) data is available by arrangement so that scores may be compared to performance of designated women's health centers under the Office of Women's Health program.

The PCSSW contains three discrete and interpretable scales: satisfaction with Communication during the visit (9 items); Administration and Office Procedures at the visit (5 items); and satisfaction with Care Coordination and Comprehensiveness over the past 12 months (10 items). Developed through careful qualitative work involving women of diverse social and health characteristics across the country, the PCSSW demonstrates factor invariance across subgroups of women defined by age group and by race/ethnicity; strong psychometric properties in the field test; and added explanatory power compared to generic satisfaction measures.

The content of the PCSSW differs from existing satisfaction tools in several ways. The items cover a range of primary care services and experiences and a number of topics specific to women or of particular concern to women. The items also address both a specific visit, and care coordination and comprehensiveness during the past year. Because many women seek health care from more than one professional or site, the latter component of the PCSSW is particularly innovative. The Care Coordination and Comprehensiveness Scale contains the most new content and also demonstrates consistently high performance across the validity assessments. The Communication Scale also performs well. However, both of the visit-specific scales, the Communication Scale and the Administration and Office Procedures Scale, are more like the interpersonal and technical aspects of quality tapped in generic tools such as the Medical Outcomes Study (MOS) Visit Satisfaction Scale. The PCSSW adds sensitivity to satisfaction measurement and can be useful in evaluations of the quality of primary care and in quality improvement programs.

To date, the PCSSW has been used in two multi-center studies assessing patient satisfaction in different organizational settings for women's primary care. In a 10-site survey of women veterans served in Veterans Administration (VA) women's clinics compared with traditional VA primary care clinics, the draft version of the PCSSW was used to provide satisfaction scores in five domains identified in the focus groups: getting care; privacy and comfort; communication; complete care; and follow-up care. In adjusted comparisons, women served in VA women's clinics reported significantly higher satisfaction on all five domains (Bean-Mayberry et al., 2003). In the national evalua-

tion of the quality of care in the clinical care centers of the CoEs, the Care Coordination and Comprehensiveness scale of the PCSSW was used in comparisons between patients served in the CoEs and patients in community comparison samples. In adjusted comparisons, women who were patients in the CoEs received more preventive and counseling services and had higher satisfaction than women in the community, with a moderate effect size (Cohen's d of 0.449) (Anderson et al. 2002). Both of these projects demonstrate that the PCSSW is useful for evaluating patient satisfaction in diverse primary care settings, including programs designed for women.

The combination of the original psychometric development and the replication analysis provide strong evidence that the PCSSW scales are psychometrically valid and reliable and add value over existing generic patient satisfaction measures. The PCSSW is not intended to replace generic satisfaction surveys; instead, it is a focused tool to assess women's health care issues. The PCSSW may be used in either self-administered or telephone formats, and it represents an important new approach for measuring women's satisfaction levels with their primary care in both traditional and women-specific health settings.

I. INTRODUCTION

Today, most health care programs use patient satisfaction surveys to assess the quality of care from patients' perspectives. Many surveys are available, some focusing on satisfaction with a specific health care visit or inpatient episode, and some focusing on health care received over a period of time. A key example of the latter is the Consumer Assessment of Health Plans Study (CAHPS) surveys, which are used by the National Committee for Quality Assurance in the assessment of quality of care received in managed care plans during the past year (Hays et al. 1999). Available patient satisfaction surveys, however, were designed for use in general patient populations and were not intended to focus on primary care issues for women (Weisman et al. 2001).

Women's primary care is characterized by the structural fragmentation of reproductive and general health care, resulting in many women using two physicians (a generalist and an obstetrician-gynecologist) for their regular health care. In addition, there are numerous clinical guidelines for the provision of a range of preventive services specific to women (e.g., Pap smears, mammograms, osteoporosis screening), and research has identified some specific patient-provider communication issues in women's primary care (e.g., better communication on sensitive topics when the provider is female). Because generic patient satisfaction tools are designed to be applicable to all patients regardless of gender, they may not be sensitive to women's concerns about these issues. As one example, the CAHPS adult survey contains questions asking respondents if they "have *one* person you think of as your personal doctor or nurse [emphasis added]" and if they have seen "specialists" in the last 12 months; these questions do not provide an opportunity for women to report *two* regular physicians (a generalist and an obstetrician-gynecologist), and the question about specialists does not include obstetrician-gynecologists in its definition of what a specialist is. As a consequence, those women who rely on two physicians for their regular health care, or who view their obstetrician-gynecologist as their "personal" physician rather than a "specialist," may find these questions to be confusing or ambiguous, with unknown effects on their responses.

It is important to recognize that the argument for a women-specific patient satisfaction survey tool is not based on any assumptions about women being more or less satisfied with their health care than men. In fact, most research finds no significant mean differences between women and men on generic patient satisfaction tools (Hall and Dornan 1990), although some studies find different predictors of satisfaction for women and men (Kolodinsky 1997; Weisman et al. 2000; Weisman et al. 2001). Furthermore, we are not attempting to develop a tool that would enable us to discern gender differences in primary care satisfaction. Rather, we were interested in developing a better measure of *women's* primary care satisfaction. The rationale is that existing patient satisfaction tools may be inadequate for measuring women's satisfaction with alternative models of care delivery or quality improvement initiatives because generic instruments do not tap into the structure of women's primary care and neglect some aspects of the process of women's health care. A tool that captures these dimensions is needed for studies and quality improvement activities in women's primary care.

An alternative approach would have been to develop a set of targeted items specific to women (i.e., not applicable to men) that could be appended to a generic patient satisfaction survey. This is the approach taken in the CAHPS surveys, for example, where special items have been developed for

Medicaid recipients, children, and persons with specific health conditions. The problem with this approach is that it would direct attention only to issues that are unique to women (e.g., reproductive health services) and would not permit reframing existing items to be more sensitive to women's overall experiences.

SUMMARY OF THE SCALE DEVELOPMENT

The PCSSW was developed as a stand-alone survey in that it encompasses both generic and woman-focused aspects of patient satisfaction. The original scale construction employed factor analytic methods and item response theory (IRT) to select the final version of the PCSSW (Horn 1965; Hu and Bentler 1995; Kaiser 1970; Samejima 1997; Velicer 1976). First, an exploratory factor analysis was performed on half of the sample selected at random (with the remaining half serving as a test set in the confirmatory analyses described below). Because the items stemmed from different content domains, the analyses were conducted separately for items referring to a specific visit and for items referring to care over the past year. Four items with excessive missing values or excessive skewness were removed before these analyses were conducted: (1) the health professional's ability to make me feel comfortable during a gynecological (pelvic) exam, (2) the health professional's comfort talking about sensitive issues like sexuality, (3) the health professional's comfort talking about natural or alternative therapies, and (4) child care here if I need it.

II. Validation in the AHRQ Funded Study (Original Sample)

The study consisted of a survey of 1,202 women making primary care visits at sites affiliated with the three participating health centers. The women completed self-administered questionnaires in the sites before and immediately after a primary care visit.

Survey data were collected at primary care sites affiliated with the University of Michigan, the University of Pittsburgh, and Wake Forest University. These sites were selected to produce a diverse sample in terms of geographic region, population density, patient demographics, clinical setting (obstetrics/gynecology, internal medicine and family medicine) and staffing (residents, primary care and specialist physicians, and advanced practice nurses or physician assistants).

2.1 Sample. Women were eligible to participate in the study if they were ages 18 and over, English-speaking, not known or suspected to be cognitively impaired, able to complete the questionnaire without assistance or proxy, and making a primary care visit at the time of the survey. We defined a “primary care visit” as a visit for a routine checkup, gynecological exam, prenatal care, acute care, or routine follow-up care with a doctor or other independent practitioner (advanced practice nurse or physician assistant). Excluded were emergency visits and expedited visits to drop off a lab specimen or to receive a single procedure such as a flu shot, allergy shot, or contraception injection. The sample characteristics are provided in **Table 1**.

The principal factor method with squared multiple correlations as prior communality estimates was used to extract the factors, and an oblique solution was obtained using the Promax factor rotation technique. The number of factors to retain was investigated by examining the scree plots, using Horn’s parallel analysis criterion (Horn 1965), conducting Velicer’s minimum average partial procedure (Velicer 1976), and noting the number of eigenvalues above the average value. Next, an item elimination and selection process was performed for each factor in order to develop indicative scales. Items were assessed by considering the magnitude of their factor loadings, assessing the item-total correlation, and fitting Samejima’s graded IRT model (Samejima 1997) to each set of candidate items that were determined to be indicative of a scale.

2.2 Explained Variance. The two oblique factor solutions explained approximately 70% of the variance in the dataset. Ten items produced standardized loadings greater than .50 on the first factors, the majority of which were directed at some aspects of the communication to and from the health professional. The highest loadings were directed at the ability of the health professional to speak clearly, with slightly lower loadings being produced by items related to listening and taking seriously communication from the patient. Some peripheral loading was produced by items more related to the communication within the practice, such as knowledge of the patient medical history. Taken together, this pattern of loading suggests that the factor be labeled communication. Five items produced loadings greater than .50 on the second factor. All of these items were clearly directed at the quality of the administrative staff and office procedures. One item, directed at the chance to talk with the health professional fully clothed, produced marginal loadings on both factors. The single item suggests a relevance in the scales, but that there were not enough similar items in the factored pool to create its own factor.

For the items pertaining to health care during the past year, the eigenanalysis suggested that a single

factor best described the correlation matrix. This factor explained about 64% of the total variance in the item set. Ten items were removed because they produced lower factor loadings, had lower discriminating ability or inferior psychometric properties. However, several items that we considered to be critical for the content validity of the scale based on our focus groups were retained despite their marginal performance in the psychometric tests. Finally, ten items were combined to form the Care Coordination and Comprehensiveness Scale.

A summary of results for the exploratory factor analysis are presented in the Results section.

2.3 Confirmatory Factor Analysis. Having settled on the items for each scale, a confirmatory factor analysis was performed on both the random half of the sample used in the exploratory factor analysis and the remaining test sample. This showed that the factors explained 70% of the total variance in the visit-specific items and 66% of the variance for the past-year items. The Tucker and Lewis reliability coefficient was 0.94 for the visit-specific item set and 0.86 for the past-year item set, suggesting excellent to good model fit. Comparison of the factor loadings from the initial sample to the test sample provided evidence that the factors were stable, although the second factor loadings for the visit-specific item set exhibited a relative alteration, suggesting that reliabilities may be lower for this scale than the other.

Factor analytic methods and item response theory (IRT) were used to identify the factor structure of the PCSSW and to select items based on their performance. After scales were constructed, bivariate analyses and multiple regression techniques were used to examine the associations between PCSSW scales and other variables in order to assess the validity of the PCSSW. Convergent validity, discriminant validity, and predictive validity were examined. To determine whether the PCSSW adds in explanatory power to generic satisfaction surveys, the relationship of the PCSSW and standard satisfaction measures (the MOS Visit Satisfaction scale and the CAHPS rating of overall health care) was examined.

2.4 Construct Validity. To assess the validity of the PCSSW scales, several analyses were conducted. For each PCSSW scale and two comparison generic patient satisfaction measures (the MOS Visit Satisfaction scale and a CAHPS overall rating item, described below), we present the mean and range. In addition, for the PCSSW and MOS scales, we show the percent with the highest possible rating (as a measure of a potential ceiling effect) and the coefficient alpha as a measure of internal consistency and reliability. To assess convergent validity with existing generic patient satisfaction tools, we calculated the correlation of the PCSSW scales with the MOS scale (a visit-specific scale) and the CAHPS item (which refers to care during the past year), using unadjusted Pearson correlations. To assess discriminant validity, we computed means for each of several known groups (adjusted for site, age, and perceived health status) and p-values from t-tests comparing the means. In addition, we present the proportion of variance in the criterion item explained by the satisfaction item, as measured by eta-squared (η^2). Eta-squared is the proportion of the sum of squares attributable to the criterion variable divided by the total corrected sum of squares (Becker 1999). An eta-squared close to 0 implies that the two groups are difficult to distinguish, and an eta-squared close to 1 implies a clear difference between the two groups. Predictive validity was assessed the same way.

To compare the PCSSW scales' ability to capture the variance in satisfaction to the generic MOS

Visit Satisfaction Scale, we conducted linear regressions with the overall visit quality rating and the CAHPS rating of the overall quality of care during the past year as the dependent variables. In the regression models, we first entered site and patient covariates (age, race/ethnicity, education, perceived health status) and then the satisfaction scales, with separate regressions for the generic MOS Visit Satisfaction Scale and for the PCSSW scales. The p-value and proportion of the variance explained represents the contribution of each satisfaction scale. The following measures were used to assess the construct validity of the PCSSW.

- *Generic satisfaction tools.* We used three generic measures of satisfaction with outpatient care for our convergent validity comparisons: the MOS Visit Satisfaction scale, the rating of all health care during the past year from the CAHPS, and a single item on visit satisfaction.
- The *MOS Visit Satisfaction* scale consists of nine items tapping multiple dimensions of care (e.g., access, time spent with provider, communication, technical quality, interpersonal quality) as well as a rating of the overall visit. It uses a 5-point excellent-to-poor response set (rated 5 to 1) that is summed to give a score. The instrument has been widely used in research on primary care and has been found to discriminate between types of practice settings and health plans and to predict such outcomes as returning for follow-up care and whether patients will change physicians within six months (Davies and Ware 1991; Rubin et al. 1993).
- The *CAHPS item* asks respondents to rate “all of your health care in the last 12 months from all doctors and other health professionals at this office or clinic.” The item is rated on a scale from 0 (worst health care possible) to 10 (best health care possible). This item is used for benchmarking satisfaction in health plans and typically is reported as percent of enrollees who rate their plan with a score of 8 or higher (NCQA 1998). In the context of this project, this item reflects the woman’s overall assessment of the quality of care at the site in the past year.
- *Continuity of care.* Having a regular source of health care or a regular provider (i.e., site and provider continuity) is known to be associated with higher levels of patient satisfaction (Aharony and Strasser 1993; Cleary and Mc Neil 1988; Donaldson 2001). Measures of these constructs were used to define known groups for assessment of discriminant validity. Women’s patterns of continuous care at the site and with a regular health care professional were measured as follows: (1) for those using the site as their usual source of care, having longer tenure at this site (two years or longer), versus shorter tenure (less than two years); and (2) for women who have a regular health care professional, seeing the regular health professional at the visit, versus seeing another health care professional. We hypothesized that women who had longer tenure at their usual source of care would have higher ratings on both the Communication and the Care Coordination and Comprehensiveness scales (as well as the MOS Visit Satisfaction scale and the CAHPS rating). We also hypothesized that women who saw their regular doctor on the day of the visit would have higher ratings on the Communication scale (as well as the MOS scale) compared with women who did not see their regular doctor (women without a regular provider were excluded).

- Comprehensiveness of care.* The comprehensiveness of preventive care, as an important aspect of the technical quality of care, was hypothesized to be associated with higher satisfaction with care. Although studies linking comprehensiveness of care with patient satisfaction are sparse, the hypothesis is supported by some literature (Cleary and McNeil 1988; Orlando and Meredith 2002; Schauffler et al. 1996; Sitzia and Wood 1997). Three variables reflecting comprehensiveness were defined for the discriminant validity analyses. *Comprehensiveness of preventive screening services* was defined as the number of age-appropriate screening services provided at the visit. For all women, these included blood pressure check, Pap smear, and physical breast exam. For women ages 50 and over, mammogram, blood cholesterol test, and colon cancer screening also were included. The sum of services received was dichotomized for analysis: a high score was three or more services, and a low score was 0-2 services. *Comprehensiveness of preventive counseling* was defined as the number of age-appropriate counseling topics discussed during non-illness visits. For all women, topics were: smoking or quitting smoking; nutrition or diet; physical fitness or exercise; alcohol or drug use; calcium intake and risk of osteoporosis; violence in the home or family or relationship problems; sexual function or problems; work or financial problems; stress management; and alternative therapies, such as herbal products or massage therapy. Additional topics for women ages 18-44 included preventing unwanted pregnancy or planning a pregnancy. Additional topics for women ages 45 and older also included menopause or hormone replacement. The sum of counseling topics was dichotomized for analysis: a high score was counseling on at least one topic at the visit, and a low score was no counseling. Finally, we asked women to rate the *completeness of their visit*: “At your visit today, did you get everything that you thought you needed?” This was scored yes or no. We hypothesized that women who had more comprehensive services based on each of these measures would have higher ratings on both visit-based and past-year measures of satisfaction.
- Behavioral intentions.* Measures of behavioral intent are viewed as determinants of subsequent behavior and are often used in patient satisfaction studies as proxies for post-visit outcomes when these outcomes cannot be observed directly. Patient satisfaction has been found to be correlated with intention to return to the provider, adhere to providers’ recommendations, recommend the provider to others, and similar items (Aharony and Strasser 1993; Ware and Davies 1983; Ware and Hays 1988). To assess predictive validity using behavioral intentions, we asked women to respond “definitely yes,” “probably yes,” “probably not,” and “definitely not” to the following four questions: “Thinking about any health care advice or recommendations you received today from your health professionals do you plan to follow this advice?” “Based on your experience at this visit, do you plan to return to this office or clinic for care?” “Based on your experiences at this visit, would you want to see the same health care professional you saw today again?” “Based on your experiences at this visit, would you recommend this office or clinic to a family member or friend?” We compared satisfaction scores of women who rated each item “definitely yes” versus all other responses. We hypothesized that women with definite intentions to follow the health professional’s advice from today’s visit and to want the same health care professional would have higher visit-based satisfaction scores. Further, we expected that women with definite intentions to return to the office or to recommend the office would have higher scores on both the visit-based and past-year ratings of satisfaction.

- *Self-efficacy for preventive care.* We hypothesized that women who received higher quality preventive care and who were more highly satisfied with their care should have higher self-efficacy for preventive care. The relationship of satisfaction levels to health-related outcomes has not been well studied (Aharony and Strasser 1993; Cleary and McNeil 1988), but a logical hypothesis for primary care is that a higher quality of both the technical and process aspects of care (especially communication between providers and patients) should result in both higher satisfaction levels and higher levels of patient knowledge and self-efficacy for health promotion and disease prevention. To measure self-efficacy for preventive care for assessment of predictive validity, we used six items to measure how sure women were that they knew when to have a mammogram, when to have the next Pap smear, how often to do a breast self-exam, when to have the next gynecologic exam, when to get a cholesterol test, and the steps to take to prevent getting osteoporosis. Each item was scored on a four-point scale from “very sure” (4) to “not at all sure” (1). The items were summed. Women who scored 20 or higher on this scale were compared with those with lower scores. We hypothesized that women with greater self-efficacy would have higher ratings on the PCSSW Communication scale and the MOS scale.

2.5 IRT. Items were assessed by considering the magnitude of their factor loadings, assessing the item-total correlation, and fitting Samejima’s graded IRT model (Samejima 1997) to each set of candidate items. IRT is a method for characterizing the relationship between a person’s responses to specific survey items and her standing on an overall construct. IRT models provide Item Characteristic Curves (ICC), also referred to as trace lines. These parametric curves describe the probability (on a scale of 0 to 1.0) that a particular respondent will choose a given item response category given her underlying satisfaction level, where satisfaction is conceptualized as an unbounded continuous latent variable with a mean of zero and a standard deviation of one. By considering the parameters that define these curves, it becomes possible to discover items that discriminate poorly between respondents with differing degrees of satisfaction. Samejima’s model introduces constraints to the trace lines and assumes the item categories are on an ordinal scale. The trace line for the lowest category will approach a probability of 1.0 as the respondent’s satisfaction level increases in the negative direction, and it will approach zero as the respondent’s satisfaction increases positively. Conversely, the trace line for the highest category will approach a probability of 1.0 as the level of satisfaction increases, and will approach zero as satisfaction decreases. Categories in between are constrained to have trace lines that reach a peak and decrease to a probability of zero in either direction.

The point where an item response category trace line would reach 1.0 at a particular level of satisfaction indicates a high level of consistency between levels of a specific item response and satisfaction.

The ideal for ordinal response sets, like those in the PCSSW, would be for respondents with low satisfaction to have a high probability of selecting low category responses and for respondents with high satisfaction to have a high probability of selecting high category responses. The set of parameters that determine the trace lines are reported in Table 3. The discrimination index (denoted as “a” in Table 3) is a measure of how well the item response continuum may differentiate levels of the latent construct (satisfaction). Items with low index values have substantial overlap among response category trace lines. Items with high index values have trace lines with little or no overlap.

Other parameters derived from the trace lines are called location parameters (b), and the number of location parameters is equal to the number of response categories minus 1. These location parameters assess item difficulty, which is defined as the point along the measurement of the latent variable (satisfaction) where respondents are likely to choose a response. By convention, the parameter “ b_1 ” denotes the point along the continuum for which there is a 50% probability of selecting the lowest response; “ b_2 ” denotes the place where there is a 50% probability of selecting either the lowest or next-to-lowest response; and so on (as shown in Table 3). Thus, items with high b_1 parameters are better at discriminating among respondents with low satisfaction.

In addition to the ICC of the individual items, IRT modeling assesses the contribution of each item in determining the precision with which patient satisfaction can be measured. Unlike classical test theory, in which precision is viewed as associated with item reliability, IRT views precision as a function of the level of the latent construct (satisfaction) itself and can be extracted from the IRT model by considering an information function plotted as an Information Curve. The concept of “information” is akin to the certainty with which the underlying construct is being measured. Satisfaction items with high information, compared with satisfaction items with low information, contribute more to the certainty (precision) with which satisfaction is being measured within the scale. The information of the overall scale is the reciprocal of the standard error of the estimate of theta. At a given satisfaction level, items with higher information contribute more to the overall precision of the scale. Information curves can be used to indicate which items are providing a high level of information and therefore should be retained in the scale. However, items with lower information are not necessarily candidates for removal because they provide at least some information and may perform well on other aspects of psychometric testing, such as item discrimination.

The final selection of items was based on consideration of the quantitative and qualitative characteristics of the items and the subscales. That is, we wanted to include items that were highly reliable, that were sensitive to the full range of the latent variables, and that contributed to the overall reliability of each subscale. At the same time, we wanted items that reflected the full range of conceptual content in each of the domains. The final set of items was selected based on all of these considerations.

Third, once having settled on the items for each scale, a new factor analysis was repeated on the exploratory sample to obtain the factor loadings of the new model. In order to assess the stability and generalizability of the proposed factor structure, the same analysis was conducted on the test sample, and results were compared. The χ^2 test and the Tucker Reliability Index (Hu and Bentler 1995) were also assessed to evaluate the fit of the new factor model on the test set.

2.6 RESULTS (Validation Sample 1)

2.6.1 Exploratory Factor Analysis. The results of the analysis of visit-specific items will be discussed first, followed by the results on items rating experiences during the past year. **Table 2** shows the results of the exploratory factor analysis.

In the analysis of the visit-specific items, two oblique factors were retained explaining approximately 70% of the variance in the data. **Table 3** shows the estimated parameters from the IRT models. The discrimination parameters suggest that all items are indicative of their corresponding scale, although some items possess better discriminating ability than others. Items Q11i, Q11j, Q11k, Q11l, and Q11h form a cluster with high discriminating properties. Item Q11i (“the health professional’s ability to explain things clearly”), in particular, has a high slope parameter ($a=5.09$), which indicates that it discriminates well among respondents with high versus low satisfaction. It also has the largest negative location parameter ($b_1 = -2.42$), which suggests that this item is the best at discriminating among respondents with low satisfaction. As an example, **Figure 1** shows the ICC curves for item Q11i. **Figure 2** shows the information curves for the Communication Scale items, with a cluster of similar performing items distinguishable from less reliable items. This model strongly suggests that these items should be retained as components of a scale assessing a trait identified as satisfaction with communication.

Of the ten items that loaded on the Communication factor, eight items were retained in the final scale. Based on the pattern of item-total correlations, factor loadings, and ICCs, as well as the face validity of the items, two items were dropped: Q11n (“The chance to get everything I need at this visit”) and Q11m (“My health professional’s knowledge of my medical history”). Generally, these items had lower loadings on the initial factor, were highly correlated with other items, and performed less well in the IRT analyses; they also did not fit as well conceptually with other items in the scale. Several items that performed less well in the analyses were included in the final scale because of their importance to the content validity of the scale.

A second scale, Administration and Office Procedures, was formed with five items loading on the second factor identified among the visit-specific items (see Table 2). One item that loaded equally on both of the two initial factors, Q11f (“The chance to talk to my health professional with my clothes on”), was grouped with the second factor. This item was retained because of its conceptual importance for this tool (as evidenced by the focus group results) and placed with the Administration and Office Procedures scale because it had marginally better performance and was more interpretable with the items on visit procedures. Table 3 shows that the items in this scale have somewhat lower discriminating ability compared to the Communication items. Item Q11f, in particular, does not possess as much discriminating ability as the others (-1.79), although it also has the largest negative location parameter ($b_1 = -2.60$) and thus appears to be best at distinguishing among individuals with low satisfaction. **Figure 3** shows the information curves for the Administration and Office Procedures items and illustrates this pattern.

For the items pertaining to health care during the past year, the factor analysis suggested that a single factor solution, as shown in Table 2, explained about 64% of the variance in the item set. The initial factor analysis identified three items with somewhat lower loadings (Q12j, Q12k, and Q12m), and the IRT model parameters (Table 3) and information curves (**Figure 4**) suggested that three items (Q12j, Q12k, and Q12l) performed less well in terms of their discriminating ability. Based on these results, Q12j (“the chance to see the same health professional at each visit”) and Q12k (“the chance to see a health professional of the gender I prefer”) were dropped. Another item, Q12g (“how well my health information is kept private”) was dropped because it was less relevant to the construct despite its reasonable psychometric properties. Several items that we considered to be critical for the

content validity of the scale based on our focus groups were retained despite their marginal performance in the psychometric tests. For example, Q121, “the chance to get both gynecological and general health care here,” had a high loading (.81) in the exploratory factor analysis. The IRT analyses suggested that this item had lower overall discriminating ability but may be better at discriminating among those with low satisfaction. This item was retained in the scale because of its importance for content validity. Finally, ten items were selected for the Care Coordination and Comprehensiveness Scale.

After deciding on the final set of items for each scale, a confirmatory factor analysis was performed on both the random half of the sample used in the exploratory factor analysis and the remaining test sample, as shown in **Table 4**. Results found that the two day-of-visit factors explained 70% of the total variance in the visit-specific items and 66% of the variance for the past-year item set. The Tucker and Lewis reliability coefficient was 0.94 for the visit-specific item set and 0.86 for the past-year item set, suggesting excellent to good model fit. Comparison of the factor loadings from the initial sample to the test sample provided evidence that the factors were stable, although the second factor loadings for the visit-specific item set exhibited a relative alteration, suggesting that reliabilities may be lower for this scale than the other.

2.6.2 Convergent Validity. Each of the PCSSW scales has high internal consistency, with coefficient alphas of 0.96 for the Communication Scale, 0.88 for the Administration and Office Procedures Scale, and 0.95 for the Care Coordination and Comprehensiveness Scale. The Communication Scale appears to be more subject to a ceiling effect than the other scales and the MOS Visit Satisfaction Scale, with 26.1% of respondents giving the highest possible rating, compared to only 10.2 – 13.4% on the other scales.

The PCSSW has very good convergent validity. Correlations with the MOS Visit Satisfaction Scale (0.67 to 0.73, $p < .001$ for all) are somewhat higher than correlations with the CAHPS overall rating (0.42 to 0.61, $p < .001$). The Care Coordination and Comprehensiveness Scale has higher correlations with both generic measures (0.73 for the MOS Visit Satisfaction Scale and 0.61 for the CAHPS item).

2.6.3 Discriminant Validity. The ability of the PCSSW scales to distinguish among groups generally expected to have higher satisfaction based on previous literature, and a comparison of this discriminant validity to that of the MOS is examined in **Table 5**. The Care Coordination and Comprehensiveness Scale discriminated among groups based on length of time at the clinical site, with a mean of 38.7 for women who had used the site for two or more years, compared with 37.6 for those with a shorter tenure at the site ($p = .02$), though the proportion of variance in length of time explained by the scale was minimal. All of the measures showed significantly higher satisfaction ratings when the regular doctor was seen at the visit.

All three PCSSW scales and the generic tools showed large differences in satisfaction based on the comprehensiveness of the visit. For example, women who said they “got everything they needed today” at their visit had an adjusted mean score of 33.5 on the PCSSW Communication Scale, compared with 24.3 for women who said “no” to this question ($p < .0001$), with this scale explaining 9% of the variance in the latter outcome. Women who received more clinical preventive services at the

visit also had higher satisfaction ratings on all PCSSW scales and the MOS Visit Satisfaction Scale, although the effect sizes were smaller (Anderson et al, 2002). The Care Coordination and Comprehensiveness Scale differentiated among women who did and did not receive counseling on preventive topics (mean scores of 39.2 and 37.2, $p < .0001$, respectively).

2.6.4 Predictive Validity. All of the satisfaction measures were significantly associated with women's behavioral intentions and self-efficacy for preventive care, and the Care Coordination and Comprehensiveness Scale had the highest effect size for intention to return and intention to recommend the office or clinic to friends. As shown in **Table 6**, the Communication Scale had the strongest effect on women's intentions to follow advice received at the visit and on their desire to see the same health care professional again.

2.6.5 Contribution of the PCSSW to Overall Satisfaction Ratings. In regression analyses examining the relationship of the generic MOS Visit Satisfaction scale and the PCSSW scales to overall ratings of the quality of care, the PCSSW scales contribute significant explanatory power over and beyond the generic MOS Scale. For example, the Communication Scale explains 26% of the variance in the overall quality rating, after patient demographic characteristics and the MOS Scale are entered into the model, and it explains 12% of the variance when the other two PCSSW scales are also in the model. The Care Coordination and Comprehensiveness Scale explains 14% of the variance in overall visit quality and 7% of the variance in the overall CAHPS rating after patient characteristics and the MOS Scale are entered into the model, and it explains 1% and 5% of the variance, respectively, when the other two PCSSW scales are also in the model.

2.6.6 Latent Variable Modeling. We also used multiple group structural equation modeling to compare the similarity of the factor structure across different race/ethnic groups (white, black, and other) and age groups (18-35, 36-54, and 55 and over) that could be defined in this sample. Although the development of the items for the PCSSW was conducted in focus groups stratified by age group and by race/ethnicity to ensure that the items were meaningful to women of all ages and ethnicities, testing for factor invariance provides evidence of PCSSW properties for these subgroups. The form of the structural equation model is shown in **Figure 5**, which graphically illustrates a set of linear relationships between observed (represented by boxes) and latent, hidden, or error variables (represented by circles) suggested by the factor analysis. Straight arrow links represent linear relationships between modeled variables and correspond to regression coefficients, labeled in the figure as beta1 through beta24. Curved arrows represent covariances between variables.

For the multiple group model fit on the different race/ethnic subgroups, the minimum discrepancy score divided by the degrees of freedom (CHISQ/DF) was 5.0, slightly above the recommended rules of thumb of 2 or 3 (Carmines and McIver 1981). The Bentler-Bonnet normed fit index (NFI) was .964, and the Comparative Fit Index (CFI) was .971, above the recommended .90 rule of thumb (Bentler and Bonnet 1980), suggesting fit was good enough that it cannot be improved substantially above the baseline independence model. The root mean square error of approximation (RSMEA) was .06, slightly above the recommended value of .05, the suggested threshold indicating a model with a reasonable error of approximation (Browne and Cudeck 1993). The unconstrained multiple-group model shows an almost identical fit (CHISQ/DF = 5.2, NFI = .965, RSMEA = .060) with a very trivial decrease in Aikake Information Criteria (.01%) when compared to the constrained

model. Taken together, these results support factor invariance for the race/ethnic groups at the current level of precision afforded by the sample, with deterioration in model fit more likely due to slight model misspecification than to the addition of constraints.

Similarly, the multiple group fit on the different age groups had a CHISQ/DF of 5.11, a NFI of 0.964, a CFI of .970, and an RSMEA of .06. The unconstrained model had a similar fit (CHISQ.DF = 5.37, NFI = 0.964, CFI = .971, RSMEA = .06), with a trivial increase in the Akaike Information Criteria (0.5%). These results also suggest factor invariance across age groups.

III. Cross-Validation Meta Analysis and Distribution Sensitivity (Validation Samples 1 and 2)

This section formally compares the psychometric evidence in support for the PCSSW across the original sample (Sample 1) above and the CoE evaluation study (Sample 2). The aim is not to replicate the entire range of psychometric tests, but rather to explore the limits of generalizability. Three specific aspects of the measurement quality of the PCSSW will be considered: Distributional Sensitivity, Structural Validity and Reliability. Each analysis will include estimates for the AHRQ data and the CoE data and is labeled accordingly.

3.1 Sample. All of the 15 CoEs in operation in 2001 participated in this evaluation under IRB approval from each CoE and the survey center. Women 18 years of age and older who had made at least one primary care visit at the CoE within the prior year were eligible for the survey. Excluded from the study were women who had no visits to the CoE during the past year, or whose most recent CoE visit was solely for an emergency visit; dropping off a specimen; a single procedure such as contraceptive injection, flu shot, a mammogram, or allergy shot; or a visit with an allied health service such as physical therapy. Also excluded were patients who did not see a doctor, nurse practitioner, nurse midwife, or physician's assistant at the most recent visit.

A target of 200 completed surveys was sought for each CoE using the following methods: a sampling frame was assembled of all patient visits during the last 3 months (for two sites this time frame was extended due to patient volume), and a random sample of up to 400 names was selected for telephone contact and eligibility screening. A recruitment database containing names and identifying information for potentially eligible participants was kept separate, and these data were deleted from the recruitment database after the survey was completed or the callback protocol was fulfilled. The telephone interview was conducted using computerized telephone interviewing (CATI) at the University of South Carolina Survey Research Lab and required an average of 15 minutes to complete. Attempted telephone calls were made at different times of day and on different days of the week to reach women who were away from home regularly. A minimum of 15 call attempts was made. If the selected respondent was not at home or otherwise unable to complete the interview at the time of the initial contact, a callback time was scheduled and repeated attempts, as necessary, were made to complete the interview. The overall response rate (defined as the number of completed interviews out of the total number of completed and partial interviews and refusals) was 70.7% and varied across sites from 57.7% to 84.7%. **Table 7** presents the sample characteristics of the study sample.

3.2 Distributional Sensitivity in Samples 1 and 2. **Tables 8 and 9** present the descriptive statistics

for the items in the PCSSW and in the AHRQ and CoE samples, respectively. Both tables show that the items of the PCSSW have means somewhat higher than their theoretical midpoints, but are not so high that they threaten the scales with ceiling effects. Moreover, the level of skew is consistent with that typically found in satisfaction measures. The standard deviations also show a moderate degree of variance among patients. Taken together, the results suggest that the patients in both samples use the entire range of the response surface.

3.3 Principal Component Analysis. **Tables 10** and **11** present the rotated factor structure for the items of the PCSSW. Analysis of the eigenvalues and comparisons of the rotated factor structures suggests that three factors capture most of the reliable variance reflected in the correlations matrix. The final structure is relatively clear, although there is some amount of cross-loading among the items. However, the patterns of loading in the rotated factor structure in both samples parallel those seen in the original analysis and the three factors support the scaling models defined in the original study. The pattern of cross loading may be an indication of the correlated nature of the underlying structure.

3.4 Confirmatory Factor Analysis. **Tables 12** and **13** present the standardized loading for a maximum likelihood confirmatory factor analysis of the scales designed in the original study. In this model, we fixed the cross-loadings for the different scales to zero and permitted the factor themselves to be intercorrelated with one another. The results of the analysis supported the structure of the scales designed in the original study. Although the χ^2 test showed a significant difference between the observed and model covariance matrix (expected due to large sample size), the patterns of residual covariance and the goodness-of-fit statistics were consistent. In the CoE study, the BBNFI = .92; CFI = .92 indicated a good fit to the covariance matrix. In the AHRQ study, the BBNFI = .91; CFI = .92 also showed a good fit to the covariance matrix. Taken together, the results showed that there was little reliable variance left to be explained after accounting for the three correlated factors in either sample. The results also show a high level of intercorrelation among the three factors in both samples. In the AHRQ study, Administration and Office Procedures and Communication scales correlated at .70 ($p < .01$), Care Coordination and Communication scales correlated at .81 ($p < .01$), and Care Coordination and Administration and Office Procedures correlated at .73 ($p < .01$). In the CoE study, Administration and Office Procedures and Communication correlated at .74 ($p < .01$), Care Coordination and Communication scales correlated at .74 ($p < .01$), and Care Coordination and Administration and Office Procedures scales correlated at .70 ($p < .01$). The confirmatory factor analysis verified the simple structure of the principal components analysis and showed that the three factors are themselves highly intercorrelated and can be reasonably expected to emerge from a single superordinate global satisfaction construct.

3.5 Reliability. **Tables 14** and **15** present the reliability analysis of the communication subscale for the AHRQ and CoE studies, respectively. Both datasets show the communication subscale to be highly internally consistent with alpha exceeding the .93 and are therefore relatively unaffected by random measurement error. The reliability analysis for the Administration and Office Procedures scale is shown in **Tables 16** and **17**. The alpha is above .83, again indicating the subscale is relatively free of random measurement error.

Finally, **Tables 18** and **19** present the reliability analysis for the Care Coordination and Comprehen-

siveness subscale, which again shows the subscales to have all alphas greater than .90.

IV. Administration Procedures

4.1 Paper and Pencil. The PCSSW can be administered using a number of different modalities. The relatively short length of the PCSSW makes it suitable for the standard paper-and-pencil methods. Because of its straightforward format, staff in clinics may be able to administer the instrument with little or no interruption of their normal routine. The instrument should take no more than 10 minutes to complete and is easily adapted to a scanning format that can be read directly into a data file. The instrument can be mailed to patients, but the fact that the questions refer to specific visits does raise the possibility of error if there is too long a lag between the visit and the responses to the questionnaire. Users should be aware that response rates for mail-out questionnaires are typically very low, and it may be very time-consuming and costly to obtain even a marginally reasonable response rate using the mail-out methodology alone.

4.2 Computer-Assisted Methods. Like all other paper-and-pencil questionnaires, the PCSSW is easily adapted to a computer-assisted telephone method. Unlike most academic studies that typically involve hundreds of questions with complex skip-out patterns, the PCSSW's 24 questions are easy to read to a person over the phone with little risk of confusing the respondent. A computer-assisted version of the PCSSW should not take longer to complete than a self-administered paper-and-pencil version. Our previous work has demonstrated that computer-assisted methodology will not alter the psychometric operation of satisfaction scales (see: McBride J., Anderson RT, and Bahnson J : Using a Hand-Held Computer to Collect Data in an Orthopedic Outpatient Clinic. A Randomized Trial of Two Survey Methods. Medical Care 1999; 37(7): 647-651)

4.3 Web-Based Survey. The PCSSW is also well suited to administration in a web-based survey format. Again, the relatively short length makes for a short response time. The reference to today's visit may pose problems for respondents who do not make their responses immediately after they leave the clinic. On the other hand, such limitations do not totally invalidate the responses given on the web. Providing a web-based response method may help with patients who do not have the time and/or patience to fill it out in the busy clinic.

4.4 Multiple Modalities. Although there is a fair amount of discussion in the academic literature regarding differences between administration modes, the effects are generally minor and probably do not seriously threaten the use of a satisfaction measure for assessing quality of care in practice. For people interested in using the PCSSW for evaluation of a clinic, we contend that whatever error is associated with the different modalities is worth trading off for the benefit of maximizing the response rate. The challenges of collecting responses for patients in various health states and in various health-related emotional states suggest the need for circumspection in deciding on the appropriate mix of methods to be used in applied science. Of course, the individual user will have to make the final decision regarding the strategy.

Norms

The ideal use of the PCSSW would include a method of comparing the scores to other benchmark data. Norms for the PCSSW are provided in **Table 20**. Standard deviations on the individual items for the scales provided in this report can be obtained from The Center on Women's Healthcare Quality Assessment (WHQA).

Interpretation of scores

The three scores of the PCSSW can be interpreted as rank order data and can for more practical purposes be treated as continuous data. The measures can be administered over time to the same subjects and analyzed with a within-subject test of mean differences, but such a design requires linking the measures to a patient identifier, which may have patient confidentiality implications. The measures can also be administered to the clinic population over time without patient identifiers in a between-groups test of means.

The individual items of the PCSSW can be examined as a content area or topic that may need quality improvement. One very useful method is to examine areas where any form of disapproval exists, and to evaluate potential problems that are manifested in these scores.

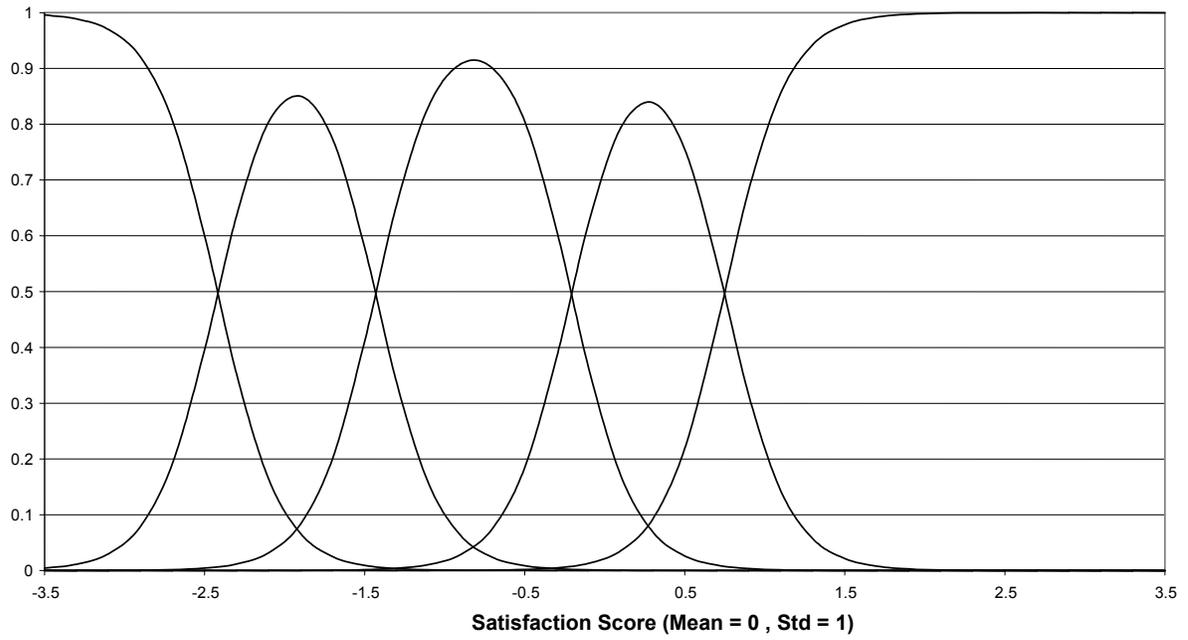
Additional Information

For additional information on the PCSSW please contact:

The Center on Women's Healthcare Assessment (WHQA): www.WHQA.org

Probability

Figure 1. Samejima Item Characteristic Curves for Item Q11i



Trace lines showing the probability of selecting the lowest to highest response at a given level of satisfaction are displayed in order

Figure 2. Information Curves for Communication Items

I(theta)

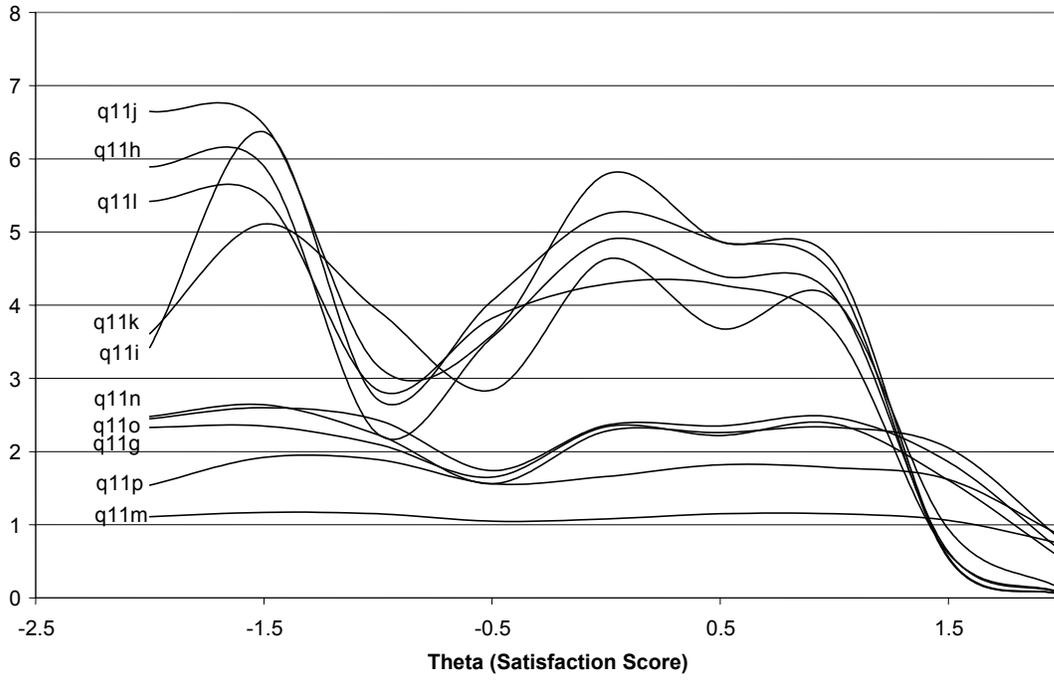


Figure 3. Information Curves for Administration and Office Procedures Items

I(theta)

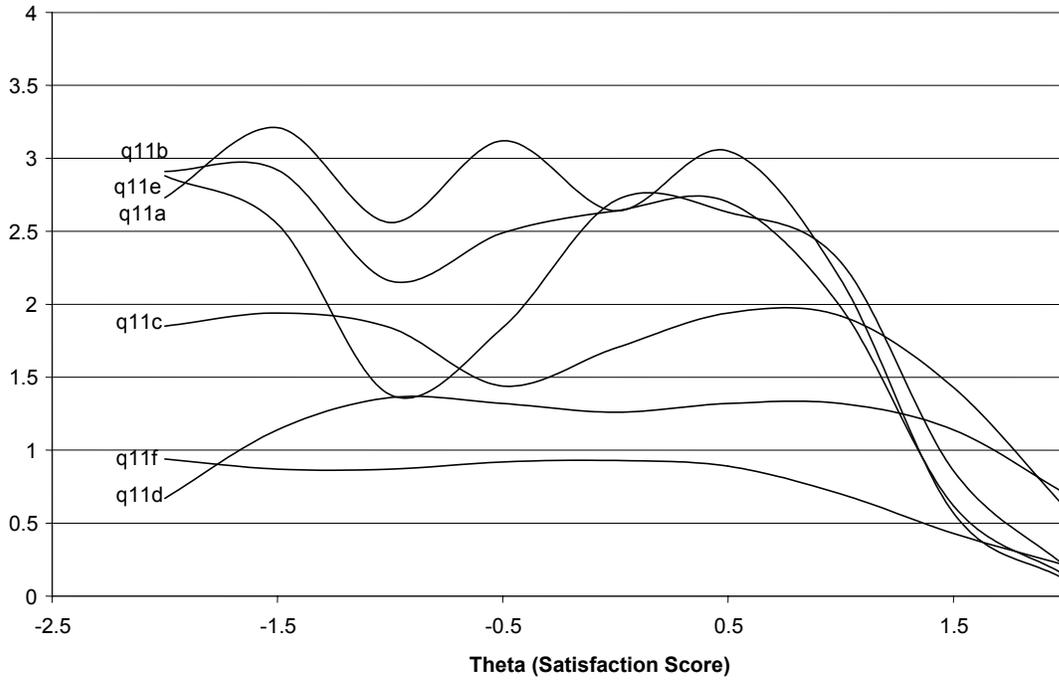


Figure 4. Information Curves for Care Coordination and Comprehensiveness Items

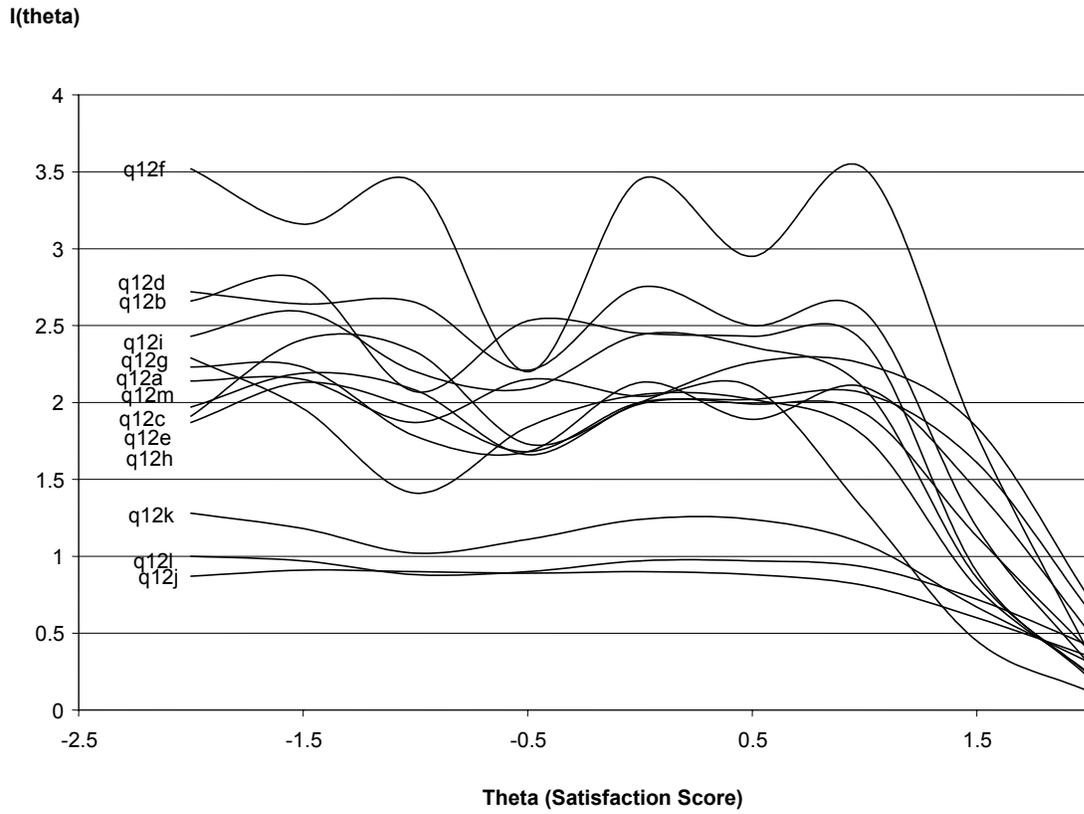


Figure 5. Structural Equation Model to Test Invariance Across Age and Race/ethnic Groups (n = 1,202)

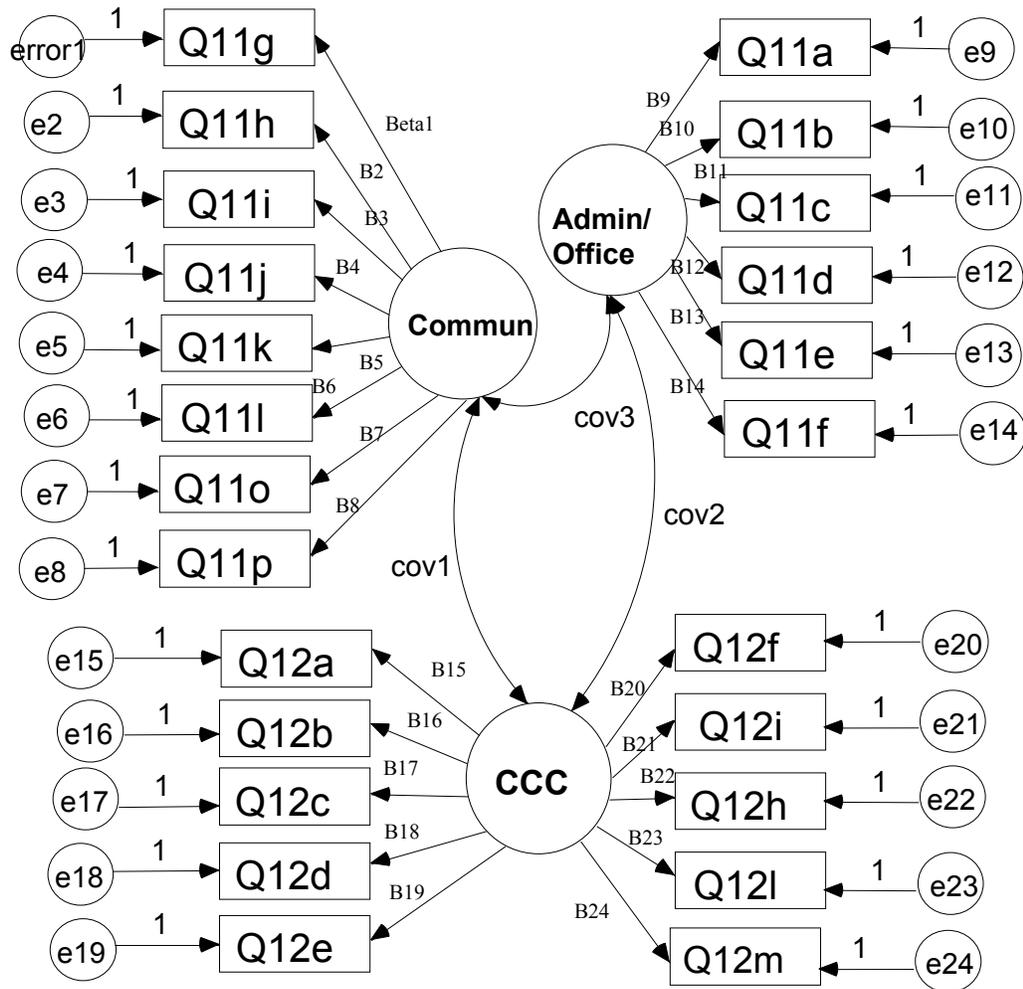


Table 1. Characteristics of the Pooled Study Sample (n = 1,202)

% or Mean (Standard Deviation)

Demographic and Health Characteristics

| | |
|---|---------------|
| Age (mean, SD) | 41.81 (16.56) |
| Education (%): | |
| High school or less | 30.35 |
| Some college | 31.28 |
| College graduate | 18.80 |
| Graduate school | 19.56 |
| Household income (%): | |
| \$20,000 or less | 25.95 |
| \$20,001 - \$50,000 | 30.71 |
| \$50,001 - \$75,000 | 15.47 |
| \$75,001 and over | 16.97 |
| Refused/Don't know | 0.90 |
| Health insurance ^a (%): | |
| Private | 62.52 |
| Medicaid | 18.86 |
| Medicare | 11.52 |
| Other/None | 7.10 |
| Race/ethnicity (%): | |
| White, non-Hispanic | 66.53 |
| Black, non-Hispanic | 23.48 |
| Other/Multi-ethnic ^b | 9.99 |
| Perceived health status ^c (%): | |
| Excellent | 13.42 |
| Very Good | 35.00 |
| Good | 34.33 |
| Fair | 14.50 |
| Poor | 2.75 |
| Pregnant in past year (%) | 18.58 |

Table 1. Characteristics of the Pooled Study Sample (continued)

Health Care Utilization

| | |
|--|--------------|
| Number of office visits, past year (mean, SD) | 7.79 (8.24) |
| Length of time coming to clinical site (%): | |
| First time today | 10.57 |
| Less than one year | 20.80 |
| 1 –2 years | 16.31 |
| 2 years or more | 52.33 |
| Clinical site is usual source of care (%) | 76.44 |
| Main reason for visit (%): | |
| Followup care | 31.65 |
| New health problem | 26.53 |
| Routine exam | 25.78 |
| Prenatal or postpartum care | 16.04 |
| Type(s) of regular provider (%): | |
| Generalist physician + Obgyn | 40.20 |
| Generalist physician | 25.93 |
| Obgyn | 10.05 |
| Other health professional | 3.89 |
| No regular provider | 19.93 |
| Regular provider (for those who have one) is at this site (%) | <u>78.25</u> |

a “Other” health insurance includes military, CHAMPUS, Tri Care, or the VA. Respondents reporting more than one source of health insurance were coded hierarchically in the following order: Medicaid, Medicare, private, other, none.

b “Other” race/ethnicity includes those with Hispanic ancestry, other racial/ethnic identification, and multi-ethnic identification.

c This item is from the SF-36: “In general, would you say your health is...” (Ware and Scherbourne 1992). When used as a covariate in analyses, it is dichotomized to contrast those reporting “fair” or “poor” health with all others.

Table 2. Exploratory Factor Loadings (n = 601)*

| First Item Set: Today's Visit | | Loadings Factor 1 | Loadings Factor 2 |
|---|--|----------------------|----------------------|
| <i>Items loading on Factor 1 (Communication)</i> | | | |
| Q11i | My health professional's ability to explain things clearly | 0.97 | -0.06 |
| Q11j | My health professional's ability to help me feel comfortable talking about my concerns | 0.95 | -0.02 |
| Q11l | My health professional's ability to take what I say seriously | 0.94 | -0.01 |
| Q11k | The chance to ask all my questions | 0.92 | 0.01 |
| Q11h | My health professional's ability to answer questions in a sensitive and caring way | 0.92 | 0.02 |
| Q11o | My health professional's willingness to explain different options for my care | 0.78 | 0.11 |
| Q11p | My health professional's interest in how my life affects my health | 0.76 | 0.06 |
| Q11g | The amount of time I had to talk with my health professional | 0.67 | 0.25 |
| Q11n | The chance to get everything I need at this visit | 0.72 | 0.20 |
| Q11m | My health professional's knowledge of my medical history | 0.59 | 0.17 |
| <i>Items loading on Factor 2 (Administration & Office Procedures)</i> | | | |
| Q11a | The courtesy of the office staff | 0.10 | 0.76 |
| Q11b | The staff's flexibility in scheduling my appointment around my needs | 0.08 | 0.78 |
| Q11c | Privacy when talking to the receptionist | 0.06 | 0.66 |
| Q11d | How well the staff kept me informed about the waiting time | -0.03 | 0.72 |
| Q11e | Help with scheduling my next visit | 0.07 | 0.75 |
| <i>Item loading both factors</i> | | | |
| Q11f | The chance to talk to my health professional with my clothes on | 0.36 | 0.40 |
| <i>% variance explained by both factors</i> | | | 70% |
| <i>Inter-factor correlation</i> | | .61 | |

Table 2. Exploratory Factor Loadings (continued)

| <u>Second Item Set: Care During the Past 12 Months</u> | | <u>Loadings Factor 1</u> |
|--|--|--------------------------|
| Q12f | How well my health care fits my stage of life | 0.87 |
| Q12d | The health professionals' interest in my mental and emotional health | 0.86 |
| Q12c | The information I get about healthy living (such as diet and exercise) | 0.83 |
| Q12i | How well the health professionals explain the results of tests or procedures | 0.83 |
| Q12a | The health professional's focus on prevention | 0.82 |
| Q12b | The health professional's knowledge of women's health issues | 0.82 |
| Q12e | Help with finding information resources in women's health | 0.81 |
| Q12l | The chance to get both gynecological and general health care here | 0.81 |
| Q12h | Information about how to get the results of my tests | 0.80 |
| Q12m | My overall trust in the health professionals here | 0.69 |
| Q12g | How well my health information is kept private | 0.79 |
| Q12k | The chance to see a health professional of the gender I prefer | 0.76 |
| Q12j | The chance to see the same health professional at each visit | 0.64 |
| <i>% variance explained by factor</i> | | 64% |

* Results of a principal factor analysis with a Promax rotation

Table 3. Item Response Theory Results: Samejima Graded Model Item Parameters and Standard Errors

| Abbreviated Item Content | a | b ₁ | b ₂ | b ₃ | b ₄ |
|---|-------------|----------------|----------------|----------------|----------------|
| Today's Visit Items | | | | | |
| Factor 1: Communication | | | | | |
| Q11g: Amount of time to talk | 3.01 (0.14) | -2.25 (0.17) | -1.27 (0.13) | 0.09 (0.05) | 1.10 (0.04) |
| Q11i: Explain things clearly | 5.09 (0.28) | -2.42 (0.18) | -1.43 (0.06) | -0.21 (0.05) | 0.75 (0.03) |
| Q11j: Help me feel comfortable talking | 5.13 (0.28) | -2.08 (0.15) | -1.38 (0.08) | -0.16 (0.04) | 0.76 (0.03) |
| Q11k: Chance to ask all of my questions | 4.29 (0.19) | -1.76 (0.10) | -1.22 (0.06) | -0.09 (0.05) | 0.83 (0.04) |
| Q11l: Take what I say seriously | 4.52 (0.24) | -2.02 (0.18) | -1.41 (0.07) | -0.23 (0.04) | 0.74 (0.03) |
| Q11h: Sensitive and caring answers | 4.73 (0.27) | -2.10 (0.28) | -1.49 (0.08) | -0.19 (0.04) | 0.76 (0.03) |
| Q11p: Interest in how my life affects my health | 2.58 (0.14) | -1.74 (0.14) | -0.97 (0.09) | 0.31 (0.05) | 1.32 (0.05) |
| Q11o: Willingness to explain different options | 3.08 (0.15) | -1.92 (0.16) | -1.14 (0.09) | 0.17 (0.04) | 1.17 (0.04) |
| Q11m: Knowledge of my medical history | 2.01 (0.11) | -2.02 (0.17) | -1.04 (0.08) | 0.36 (0.05) | 1.42 (0.06) |
| Q11n: The chance to get everything I need | 3.07 (0.16) | -1.95 (0.13) | -1.23 (0.09) | 0.18 (0.05) | 1.25 (0.04) |
| Factor 2: Administration and Office Procedures | | | | | |
| Q11a: Courtesy of the office staff | 3.50 (0.21) | -2.41 (0.21) | -1.43 (0.07) | -0.38 (0.04) | 0.65 (0.04) |
| Q11b: Flexibility in scheduling my appointment | 3.22 (0.18) | -2.09 (0.15) | -1.44 (0.08) | -0.26 (0.04) | 0.67 (0.04) |
| Q11c: Privacy when talking to the receptionist | 2.63 (0.14) | -1.98 (0.14) | -1.11 (0.07) | 0.28 (0.05) | 1.15 (0.06) |
| Q11d: Informed about the waiting time | 2.12 (0.12) | -1.29 (0.09) | -0.67 (0.06) | 0.44 (0.05) | 1.32 (0.07) |
| Q11e: Help with scheduling next visit | 3.18 (0.18) | -2.02 (0.14) | -1.66 (0.10) | -0.09 (0.04) | 0.78 (0.05) |
| Q11f: Talk with my clothes on | 1.79 (0.12) | -2.60 (0.24) | -2.02 (0.16) | -0.55 (0.07) | 0.56 (0.06) |

Table 3. Item Response Theory Results: Samejima Graded Model Item Parameters and Standard Errors (continued)

| Abbreviated Item Content | a | b ₁ | b ₂ | b ₃ | b ₄ |
|--|-------------|----------------|----------------|----------------|----------------|
| Care During Past Year Items | | | | | |
| Q12a: Focus on prevention | 2.83 (0.13) | -2.20(0.15) | -1.38 (0.08) | -0.03 (0.04) | 1.08 (0.05) |
| Q12b: Knowledge of women's health issues | 3.24 (0.16) | -2.36 (0.19) | -1.49 (0.09) | -0.29 (0.04) | 0.80 (0.04) |
| Q12c: Information about healthy living | 2.78 (0.13) | -1.88 (0.10) | -1.09 (0.06) | 0.18 (0.04) | 1.18 (0.05) |
| Q12d: Interest in my mental and emotional health | 3.21 (0.16) | -2.02 (0.12) | -1.14 (0.07) | -0.03 (0.04) | 0.92 (0.04) |
| Q12e: Information resources in women's health | 2.91 (0.14) | -1.75 (0.10) | -1.02 (0.07) | 0.28 (0.04) | 1.23 (0.05) |
| Q12f: Care fits my stage of life | 3.70 (0.17) | -1.98 (0.12) | -1.11 (0.06) | 0.09 (0.04) | 1.04 (0.04) |
| Q12g: Information is kept private | 2.78 (0.13) | -2.28 (0.17) | -1.70 (0.10) | -0.24 (0.05) | 0.77 (0.05) |
| Q12h: Get the results of my tests | 2.71 (0.13) | -1.87 (0.12) | -1.16 (0.20) | 0.05 (0.05) | 0.96 (0.05) |
| Q12i: Explain results of tests or procedures | 3.01 (0.14) | -1.98 (0.12) | -1.28 (0.07) | -0.13 (0.04) | 0.79 (0.05) |
| Q12j: Same health professional at each visit | 1.73 (0.10) | -2.11 (0.16) | -1.25 (0.09) | -0.13 (0.06) | 0.92 (0.07) |
| Q12k: Health professional of the gender I prefer | 2.08 (0.11) | -2.47 (0.18) | -1.67 (0.11) | -0.14 (0.05) | 0.81 (0.06) |
| Q12l: Both gynecological and general health care | 1.84 (0.10) | -2.46 (0.20) | -1.58 (0.11) | -0.07 (0.06) | 1.03 (0.07) |
| Q12m: Overall trust in health professionals | 2.82 (0.15) | -2.48 (0.20) | -1.59 (0.09) | -0.46 (0.05) | 0.52 (0.04) |

a = discrimination index (see text)

b₁, ..., b₄ = item difficulty parameters (see text)

Table 4. Factor Analysis using Maximum Likelihood Factor Analysis on both the Initial and Test Data Sets

| | Loadings Factor 1 | | Loadings Factors 2 | |
|---|--------------------------------------|----------------|-------------------------------|-----------------|
| First Item Set: Today's Visit | Standardized Regression Coefficients | | | |
| Factor 1: Communication | INITIAL N= 601 | TEST N= 601 | INITIAL N= 601 | TEST N = 601 |
| Q11j: Help me feel comfortable talking | 0.97 | 0.97 | -0.02 | -0.03 |
| Q11i: Explain things clearly | 0.98 | 0.92 | -0.04 | 0.02 |
| Q11l: Take what I say seriously | 0.94 | 0.90 | 0.01 | 0.02 |
| Q11h: Sensitive and caring answers | 0.93 | 0.90 | 0.03 | 0.05 |
| Q11k: Chance to ask all of my questions | 0.92 | 0.89 | 0.03 | -0.02 |
| Q11o: Willingness to explain different options | 0.70 | 0.76 | 0.17 | 0.11 |
| Q11g: Amount of time to talk | 0.67 | 0.72 | 0.26 | 0.17 |
| Q11p: Interest in how my life affects my health | 0.68 | 0.68 | 0.11 | 0.14 |
| Factor 2: Administration and Office Procedures | | | | |
| Q11c: Privacy when talking to the receptionist | 0.06 | -0.04 | 0.64 | 0.83 |
| Q11e: Help with scheduling next visit | 0.05 | 0.09 | 0.78 | 0.76 |
| Q11b: Flexibility in scheduling my appointment | 0.08 | 0.08 | 0.80 | 0.72 |
| Q11d: Informed about the waiting time | -0.02 | -0.02 | 0.69 | 0.72 |
| Q11a: Courtesy of the office staff | 0.08 | 0.20 | 0.79 | 0.66 |
| Q11f: Talk with my clothes on | 0.33 | 0.36 | 0.42 | 0.40 |
| <i>% Variance Explained by Both Factors</i> | INITIAL = 71% ,TEST = 70% | | | |
| <i>Interfactor Correlation</i> | INITIAL = .60, TEST = .60 | | | |
| | Loadings Factor1 | | | |
| Second Item Set | Standardized Regression Coefficients | | | |
| | INITIAL N=601 | TEST N= 601 | | |
| Q12f: Care fits my stage of life | 0.88 | 0.89 | | |
| Q12d: Interest in my mental and emotional health | 0.88 | 0.85 | | |
| Q12b: Knowledge of women's health issues | 0.83 | 0.85 | | |
| Q12e: Information resources in women's health | 0.83 | 0.83 | | |
| Q12i: Explain results of tests or procedures | 0.81 | 0.82 | | |
| Q12a: Focus on prevention | 0.83 | 0.82 | | |
| Q12c: Information about healthy living | 0.85 | 0.82 | | |
| Q12m: Overall trust in health professionals | 0.78 | 0.79 | | |
| Q12h: Get the results of my tests | 0.77 | 0.79 | | |
| Q12l: Both gynecological and general health care | 0.65 | 0.66 | | |
| <i>% Variance Explained by Factor</i> | INITIAL = 66%, TEST = 66% | | | |

Table 5. Discriminant Validity of the PCSSW and Generic Scales (adjusted means and standard er-

rors; n = 1,202)*

| | PCSSW | | | MOS Visit Satisfaction | CAHPS Quality of Care |
|------------------------------|---------------|------------------------|----------------------------------|------------------------|-----------------------|
| | Communication | Administration/ Office | Coordination & Comprehensiveness | | |
| Length of time at this place | | | | | |
| 2 years or longer (n=629) | 33.34(0.43) | 23.15(0.32) | 38.70(0.55) | 35.42(0.44) | 8.58(0.10) |
| Less than 2 years (n=573) | 32.79(0.43) | 23.17(0.31) | 37.55(0.55) | 34.81(0.44) | 8.42(0.10) |
| p-value (% variance) | NS (0%) | NS (0%) | .02 (0%) | NS (0%) | NS (0%) |
| Saw Regular Doctor Today | | | | | |
| Yes (n=721) | 33.62(0.40) | 23.41(0.30) | 38.82(0.52) | 35.52(0.42) | 8.60(0.09) |
| No (n=473) | 32.03(0.46) | 22.67(0.34) | 36.81(0.59) | 34.35(0.48) | 8.35(0.11) |
| p-value (% variance) | <.0001 (1%) | .01 (0%) | <.0001 (1%) | .0050(1%) | .0070(1%) |
| Got everything needed | | | | | |
| Yes (n=1,138) | 33.49(0.37) | 23.27(0.28) | 38.50(0.48) | 35.37(0.39) | 8.58(0.09) |
| No (n=53) | 24.26(0.91) | 20.50(0.69) | 29.37(1.19) | 29.19(0.96) | 7.17(0.21) |
| p-value (% variance) | <.0001 (9%) | <.0001(1%) | <.0001 (5%) | <.0001(3%) | <.0001 (4%) |
| Counseling topics | | | | | |
| At Least 1 (n=745) | 33.82(0.40) | 23.36(0.30) | 39.21(0.51) | 35.63(0.42) | 8.66(0.09) |
| None (n=457) | 31.72(0.45) | 22.79(0.33) | 36.17(0.57) | 34.19(0.46) | 8.26(0.10) |
| p-value (% variance) | <.0001 (2%) | .04 (0%) | <.0001 (3%) | .0003 (1%) | <.0001 (2%) |
| Preventive services | | | | | |
| High (3 or more) (n=343) | 33.80(0.48) | 23.51(0.35) | 39.53(0.62) | 35.49(0.49) | 8.63(0.11) |
| Low (n=859) | 32.75(0.40) | 23.01(0.29) | 37.57(0.51) | 34.95(0.41) | 8.46(0.09) |
| p-value (% variance) | .01 (0%) | NS (0%) | .0005 (1%) | NS (0%) | .06 (0%) |

PCSSW= Primary Care Satisfaction Survey for Women

MOS=Medical Outcomes Study

CAHPS= Consumer Assessment of Health Plans

* Means are adjusted for site, age, education, and perceived health status.

** "At your visit today, did you get everything that you thought you needed?"

*** % variance is the percent of the variance in the satisfaction scale that is explained by group membership. This % is also referred to as the eta-squared (see methods).

Table 6. Predictive Validity: Relationship of the Satisfaction Measures to Behavioral Intentions and Self-efficacy for Preventive Care (adjusted means and standard errors; n = 1,202)*

| | PCSSW | | | MOS Visit Satisfaction | CAHPS Quality of Care |
|------------------------------------|--------------------------|------------------------|----------------------------------|------------------------|-----------------------|
| | Communication | Administration/ Office | Coordination & Comprehensiveness | | |
| Plan to return to office | | | | | |
| Definitely Yes (n=1,096) | 33.73(0.37) ¹ | 23.55(0.28) | 39.10(0.47) | 35.91(0.37) | 8.68(0.08) |
| Other (n= 105) | 26.72(0.68) | 19.50(0.51) | 28.80(0.85) | 27.57(0.68) | 6.88(0.15) |
| p-value (% variance) | <.0001(10%) | <.0001(6%) | <.0001(12%) | <.0001(12%) | <.0001(11%) |
| Recommend this of- fice/clinic | | | | | |
| Definitely Yes (n=958) | 34.49(0.35) | 23.97(0.27) | 40.09(0.44) | 36.64(0.36) | 8.83(0.08) |
| Other (n=240) | 27.56(0.48) | 19.90(0.37) | 30.35(0.61) | 29.98(0.49) | 7.25(0.11) |
| p-value (% variance) | <.0001(19%) | <.0001(12%) | <.0001(22%) | <.0001(21%) | <.0001(18%) |
| Plan to follow advice | | | | | |
| Definitely Yes (n=973) | 33.92(0.37) | 23.53(0.28) | 39.11(0.49) | 35.91(0.39) | 8.64(0.09) |
| Other (n=228) | 29.40(0.56) | 21.40(0.42) | 34.16(0.73) | 31.72(0.58) | 7.92(0.13) |
| p-value (% variance) | <.0001(6%) | <.0001(3%) | <.0001(5%) | <.0001(5%) | <.0001(3%) |
| Want to see same profes- sional | | | | | |
| Definitely Yes (n=983) | 34.24(0.36) | 23.64(0.28) | 39.43(0.47) | 36.12(0.38) | 8.73(0.08) |
| Other (n=215) | 27.48(0.51) | 20.87(0.40) | 31.97(0.67) | 30.32(0.54) | 7.51(0.12) |
| p-value (% variance) | <.0001(17%) | <.0001(5%) | <.0001(12%) | <.0001(11%) | <.0001(9%) |

PCSSW= Primary Care Satisfaction Survey for Women

MOS=Medical Outcomes Study

CAHPS= Consumer Assessment of Health Plans

* Means are adjusted for site, age, education, and perceived health status.

Table 7: Selected Demographic and Descriptive Characteristics of CoE and Commonwealth Fund (CWF) Samples

| Characteristic | COE Sample (n=3,111) | | CWF Sample (n=2,075) | |
|---|--------------------------|-------------------------|--------------------------|-------------------------|
| | Mean or Proportion | 95% Confidence Interval | Mean or Proportion | 95% Confidence Interval |
| Mean age | 45.24 (range = 18-94) | (44.66, 45.81) | 45.41 (range = 18-97) | (44.56, 46.26) |
| <i>Ethnicity</i> | | | | |
| White, non-Hispanic | 0.552 | (0.54, 0.56) | 0.729 | (0.71, 0.75) |
| African-American, non-Hispanic | 0.242 | (0.23, 0.25) | 0.125 | (0.11, 0.14) |
| Hispanic | 0.117 | (0.11, 0.12) | 0.086 | (0.08, 0.10) |
| Asian/Pacific Islander | 0.039 | (0.04, 0.04) | 0.031 | (0.03, 0.03) |
| Other | 0.051 | (0.05, 0.05) | 0.029 | (0.02, 0.04) |
| <i>Marital status:</i> | | | | |
| Married/living with partner | 0.544 | (0.54, 0.55) | 0.566 | (0.54, 0.59) |
| Single | 0.204 | (0.20, 0.21) | 0.187 | (0.17, 0.21) |
| Widowed/separated/divorced | 0.250 | (0.24, 0.26) | 0.247 | (0.23, 0.27) |
| <i>Employment status:</i> | | | | |
| Employed | 0.628 | (0.62, 0.64) | 0.588 | (0.57, 0.61) |
| Not employed | 0.372 | (0.36, 0.38) | 0.412 | (0.39, 0.43) |
| <i>Children under 18 years in household</i> | 0.374 | (0.37, 0.38) | 0.400 | (0.38, 0.42) |
| <i>Education:</i> | | | | |
| Less than high school | 0.084 | (0.08, 0.09) | 0.186 | (0.17, 0.21) |
| High school/ some college | 0.356 | (0.35, 0.36) | 0.604 | (0.58, 0.63) |
| College graduate/ more | 0.560 | (0.55, 0.57) | 0.210 | (0.19, 0.23) |
| <i>Income:</i> | | | | |
| \$10,000 or less | 0.153 | (0.15, 0.16) | 0.126 | (0.11, 0.14) |
| \$10,001 to \$20,000 | 0.127 | (0.12, 0.13) | 0.170 | (0.15, 0.19) |
| \$20,001 to \$30,000 | 0.110 | (0.10, 0.12) | 0.120 | (0.10, 0.13) |
| \$30,001 to \$40,000 | 0.106 | (0.10, 0.11) | 0.154 | (0.14, 0.17) |
| \$40,001 to \$50,000 | 0.093 | (0.09, 0.10) | 0.106 | (0.09, 0.12) |
| \$50,001 to \$75,000 | 0.154 | (0.15, 0.16) | 0.124 | (0.11, 0.14) |
| \$75,001 to \$100,000 | 0.114 | (0.11, 0.12) | 0.044 | (0.04, 0.05) |
| \$100,001 or above | 0.143 | (0.14, 0.15) | 0.040 | (0.03, 0.05) |

Table 8: Descriptive Statistics for the Women's Primary Care Satisfaction Survey AHRQ Data

| | M | SD | Min | Max | N |
|--|------|------|------|------|------|
| 1: Visit today | | | | | |
| a. The courtesy of the staff..... | 4.06 | 0.89 | 1.00 | 5.00 | 1198 |
| b. The staff's flexibility in scheduling my appointment around my needs..... | 4.01 | 0.95 | 1.00 | 5.00 | 1180 |
| c. Privacy when talking to the receptionist | 3.60 | 1.00 | 1.00 | 5.00 | 1183 |
| d. How well the staff kept you informed about the waiting time..... | 3.31 | 1.21 | 1.00 | 5.00 | 1184 |
| e. Help with scheduling my next visit..... | 3.92 | 0.93 | 1.00 | 5.00 | 1130 |
| f. The chance to talk to my health professional with my clothes on... | 4.12 | 0.94 | 1.00 | 5.00 | 1187 |
| g. The amount of time I had to talk with my health professional..... | 4.12 | 0.92 | 1.00 | 5.00 | 1197 |
| h. My health professional's ability to answer questions in a sensitive and caring way... | 4.32 | 0.85 | 1.00 | 5.00 | 1199 |
| i. My health professional's ability to explain things clearly..... | 4.33 | 0.84 | 1.00 | 5.00 | 1197 |
| j. My health professional's ability to help me feel comfortable talking about my concerns..... | 4.32 | 0.86 | 1.00 | 5.00 | 1197 |
| k. The chance to ask all of my questions... | 4.26 | 0.91 | 1.00 | 5.00 | 1197 |
| l. My health professional's ability to take what I say seriously..... | 4.34 | 0.85 | 1.00 | 5.00 | 1198 |
| m. My health professional's willingness to explain different options for my care... | 4.07 | 0.95 | 1.00 | 5.00 | 1188 |
| n. My health professionals interest in how my life affects health | 3.93 | 1.02 | 1.00 | 5.00 | 1192 |
| 2. During the last 12 months. | | | | | |
| a. The health professional's focus on prevention..... | 3.84 | 0.94 | 1.00 | 5.00 | 1190 |
| b. The health professional's knowledge of women's health issues..... | 4.02 | 0.91 | 1.00 | 5.00 | 1195 |
| c. The information I get about healthy living (such as diet and exercise)..... | 3.70 | 1.03 | 1.00 | 5.00 | 1184 |
| d. The health professional's interest in my mental and emotional health..... | 3.86 | 1.01 | 1.00 | 5.00 | 1191 |
| e. Help with finding information resources in women's health..... | 3.62 | 1.06 | 1.00 | 5.00 | 1165 |
| f. How well my health care fits my stage of life..... | 3.81 | 0.98 | 1.00 | 5.00 | 1189 |
| g. How well the health professional explain the results of tests or procedures... | 3.95 | 1.00 | 1.00 | 5.00 | 1186 |
| h. The chance to see the same health prof. at each visit..... | 3.97 | 0.97 | 1.00 | 5.00 | 1185 |
| i. The chance to get both gynecological and general health care here..... | 3.87 | 1.00 | 1.00 | 5.00 | 1153 |
| j. My overall trust in the health professionals here..... | 4.18 | 0.91 | 1.00 | 5.00 | 1194 |

Table 9: Descriptive Statistics for the Women's Primary Care Satisfaction Survey Center of Excellence Data

| | M | SD | Min | Max | N |
|---|------|------|------|------|------|
| 1: Visit today | | | | | |
| a. The courtesy of the staff..... | 3.70 | 0.93 | 1.00 | 5.00 | 3095 |
| b. The staff's flexibility in scheduling my appointment around my needs..... | 3.40 | 1.08 | 1.00 | 5.00 | 3083 |
| c. Privacy when talking to the receptionist | 3.24 | 1.01 | 1.00 | 5.00 | 3039 |
| d. How well the staff kept you informed about the waiting time..... | 3.20 | 1.14 | 1.00 | 5.00 | 3033 |
| e. Help with scheduling my next visit..... | 3.49 | 1.00 | 1.00 | 5.00 | 2926 |
| f. The chance to talk to my health professional with my clothes on..... | 3.72 | 0.93 | 1.00 | 5.00 | 3032 |
| g. The amount of time I had to talk with my health professional..... | 3.58 | 0.97 | 1.00 | 5.00 | 2879 |
| h. My health professional's ability to answer questions in a sensitive and caring way | 3.94 | 0.90 | 1.00 | 5.00 | 3012 |
| i. My health professional's ability to explain things clearly..... | 3.47 | 1.04 | 1.00 | 5.00 | 2799 |
| j. My health professional's ability to help me feel comfortable talking about my concerns | 3.57 | 1.05 | 1.00 | 5.00 | 2865 |
| k. The chance to ask all of my questions... | 3.45 | 1.04 | 1.00 | 5.00 | 2760 |
| l. My health professional's ability to take what I say seriously..... | 3.58 | 0.98 | 1.00 | 5.00 | 2990 |
| m. My health professional's willingness to explain different options for my care... | 3.59 | 1.07 | 1.00 | 5.00 | 2977 |
| n. My health professionals interest in how my life affects health | 3.66 | 1.13 | 1.00 | 5.00 | 2943 |
| 2. During the last 12 months. | | | | | |
| a. The health professional's focus on prevention..... | 3.78 | 0.97 | 1.00 | 5.00 | 2755 |
| b. The health professional's knowledge of women's health issues..... | 3.89 | 0.95 | 1.00 | 5.00 | 3049 |
| c. The information I get about healthy living (such as diet and exercise)..... | 3.67 | 1.03 | 1.00 | 5.00 | 3087 |
| d. The health professional's interest in my mental and emotional health..... | 3.95 | 0.96 | 1.00 | 5.00 | 3086 |
| e. Help with finding information resources in women's health..... | 3.95 | 0.93 | 1.00 | 5.00 | 3088 |
| f. How well my health care fits my stage of life..... | 3.94 | 0.94 | 1.00 | 5.00 | 3088 |
| g. How well the health professional explain the results of tests or procedures... | 3.85 | 0.99 | 1.00 | 5.00 | 3089 |
| h. The chance to see the same health prof. at each visit..... | 3.92 | 0.97 | 1.00 | 5.00 | 3083 |
| i. The chance to get both gynecological and general health care here..... | 3.78 | 0.98 | 1.00 | 5.00 | 3039 |
| j. My overall trust in the health professionals here..... | 3.66 | 1.04 | 1.00 | 5.00 | 3000 |

Table 10: Varimax Rotated Factor Structure Matrix for the Women's Primary Care Satisfaction Survey AHRQ Data

| | 1 | 2 | 3 |
|--|------|------|------|
| 1: Visit today | | | |
| a. The courtesy of the staff..... | 0.23 | 0.34 | 0.74 |
| b. The staff's flexibility in scheduling my appointment around my needs..... | 0.24 | 0.28 | 0.76 |
| c. Privacy when talking to the receptionist | 0.27 | 0.16 | 0.75 |
| d. How well the staff kept you informed about the waiting time..... | 0.24 | 0.11 | 0.75 |
| e. Help with scheduling my next visit..... | 0.25 | 0.28 | 0.75 |
| f. The chance to talk to my health professional with my clothes on..... | 0.25 | 0.47 | 0.50 |
| g. The amount of time I had to talk with my health professional..... | 0.31 | 0.72 | 0.38 |
| h. My health professional's ability to answer questions in a sensitive and caring way... | 0.31 | 0.84 | 0.29 |
| i. My health professional's ability to explain things clearly..... | 0.35 | 0.84 | 0.25 |
| j. My health professional's ability to help me feel comfortable talking about my concerns... | 0.35 | 0.85 | 0.24 |
| k. The chance to ask all of my questions... | 0.35 | 0.82 | 0.23 |
| l. My health professional's ability to take what I say seriously..... | 0.36 | 0.82 | 0.24 |
| m. My health professional's willingness to explain different options for my care... | 0.53 | 0.64 | 0.24 |
| n. My health professionals interest in how my life affects health | 0.61 | 0.55 | 0.20 |
| 2. During the last 12 months. | | | |
| a. The health professional's focus on prevention..... | 0.72 | 0.36 | 0.27 |
| b. The health professional's knowledge of women's health issues..... | 0.69 | 0.44 | 0.25 |
| c. The information I get about healthy living (such as diet and exercise)..... | 0.80 | 0.26 | 0.23 |
| d. The health professional's interest in my mental and emotional health..... | 0.80 | 0.34 | 0.19 |
| e. Help with finding information resources in women's health..... | 0.79 | 0.18 | 0.31 |
| f. How well my health care fits my stage of life..... | 0.80 | 0.31 | 0.26 |
| g. How well the health professional explain the results of tests or procedures... | 0.64 | 0.38 | 0.35 |
| h. The chance to see the same health prof. at each visit..... | 0.56 | 0.40 | 0.31 |
| i. The chance to get both gynecological and general health care here..... | 0.53 | 0.31 | 0.35 |
| j. My overall trust in the health professionals here..... | 0.58 | 0.54 | 0.28 |

Table 11: Varimax Rotated Factor Structure Matrix for the Women's Primary Care Satisfaction Survey COE Data

| | 1 | 2 | 3 |
|--|------|------|------|
| 1: Visit today | | | |
| a. The courtesy of the staff..... | 0.30 | 0.18 | 0.70 |
| b. The staff's flexibility in scheduling my appointment around my needs..... | 0.22 | 0.22 | 0.71 |
| c. Privacy when talking to the receptionist | 0.08 | 0.22 | 0.72 |
| d. How well the staff kept you informed about the waiting time..... | 0.14 | 0.25 | 0.73 |
| e. Help with scheduling my next visit..... | 0.27 | 0.22 | 0.73 |
| f. The chance to talk to my health professional with my clothes on..... | 0.60 | 0.24 | 0.45 |
| g. The amount of time I had to talk with my health professional..... | 0.68 | 0.27 | 0.36 |
| h. My health professional's ability to answer questions in a sensitive and caring way... | 0.84 | 0.31 | 0.21 |
| i. My health professional's ability to explain things clearly..... | 0.83 | 0.32 | 0.18 |
| j. My health professional's ability to help me feel comfortable talking about my concerns... | 0.83 | 0.35 | 0.19 |
| k. The chance to ask all of my questions... | 0.81 | 0.33 | 0.21 |
| l. My health professional's ability to take what I say seriously..... | 0.81 | 0.35 | 0.19 |
| m. My health professional's willingness to explain different options for my care... | 0.71 | 0.44 | 0.23 |
| n. My health professionals interest in how my life affects health | 0.61 | 0.52 | 0.26 |
| 2. During the last 12 months. | | | |
| a. The health professional's focus on prevention..... | 0.33 | 0.71 | 0.27 |
| b. The health professional's knowledge of women's health issues..... | 0.55 | 0.60 | 0.20 |
| c. The information I get about healthy living (such as diet and exercise)..... | 0.24 | 0.77 | 0.22 |
| d. The health professional's interest in my mental and emotional health..... | 0.36 | 0.74 | 0.21 |
| e. Help with finding information resources in women's health..... | 0.27 | 0.73 | 0.30 |
| f. How well my health care fits my stage of life..... | 0.41 | 0.66 | 0.29 |
| g. How well the health professional explain the results of tests or procedures... | 0.38 | 0.57 | 0.35 |
| h. The chance to see the same health prof. at each visit..... | 0.39 | 0.52 | 0.26 |
| i. The chance to get both gynecological and general health care here..... | 0.47 | 0.50 | 0.26 |
| j. My overall trust in the health professionals here..... | 0.56 | 0.54 | 0.25 |

Table 12: Confirmatory Factor Analysis for the Women's Primary Care Satisfaction Survey AHRQ Data (Standardized Coefficients)

| | 1 | 2 | 3 |
|--|------|------|------|
| 1: Visit today | | | |
| a. The courtesy of the staff..... | | 0.83 | |
| b. The staff's flexibility in scheduling my appointment around my needs..... | | 0.82 | |
| c. Privacy when talking to the receptionist | | 0.73 | |
| d. How well the staff kept you informed about the waiting time..... | | 0.68 | |
| e. Help with scheduling my next visit..... | | 0.81 | |
| f. The chance to talk to my health professional with my clothes on..... | | 0.66 | |
| g. The amount of time I had to talk with my health professional..... | 0.83 | | |
| h. My health professional's ability to answer questions in a sensitive and caring way... | 0.94 | | |
| i. My health professional's ability to explain things clearly..... | 0.94 | | |
| j. My health professional's ability to help me feel comfortable talking about my concerns... | 0.95 | | |
| k. The chance to ask all of my questions... | 0.90 | | |
| l. My health professional's ability to take what I say seriously..... | 0.92 | | |
| m. My health professional's willingness to explain different options for my care... | 0.82 | | |
| n. My health professionals interest in how my life affects health | 0.77 | | |
| 2. During the last 12 months. | | | |
| a. The health professional's focus on prevention..... | | | 0.83 |
| b. The health professional's knowledge of women's health issues..... | | | 0.85 |
| c. The information I get about healthy living (such as diet and exercise)..... | | | 0.82 |
| d. The health professional's interest in my mental and emotional health..... | | | 0.86 |
| e. Help with finding information resources in women's health..... | | | 0.81 |
| f. How well my health care fits my stage of life..... | | | 0.88 |
| g. How well the health professional explain the results of tests or procedures... | | | 0.80 |
| h. The chance to see the same health prof. at each visit..... | | | 0.72 |
| i. The chance to get both gynecological and general health care here..... | | | 0.68 |
| j. My overall trust in the health professionals here..... | | | 0.81 |

Table 13: Confirmatory Factor Analysis for the Women's Primary Care Satisfaction Survey COE Data (Standardized Coefficients)

| | 1 | 2 | 3 |
|--|------|------|------|
| 1: Visit today | | | |
| a. The courtesy of the staff..... | | 0.72 | |
| b. The staff's flexibility in scheduling my appointment around my needs..... | | 0.67 | |
| c. Privacy when talking to the receptionist | | 0.60 | |
| d. How well the staff kept you informed about the waiting time..... | | 0.67 | |
| e. Help with scheduling my next visit..... | | 0.60 | |
| f. The chance to talk to my health professional with my clothes on..... | | 0.67 | |
| g. The amount of time I had to talk with my health professional..... | 0.78 | | |
| h. My health professional's ability to answer questions in a sensitive and caring way... | 0.90 | | |
| i. My health professional's ability to explain things clearly..... | 0.90 | | |
| j. My health professional's ability to help me feel comfortable talking about my concerns... | 0.92 | | |
| k. The chance to ask all of my questions... | 0.89 | | |
| l. My health professional's ability to take what I say seriously..... | 0.88 | | |
| m. My health professional's willingness to explain different options for my care... | 0.78 | | |
| n. My health professionals interest in how my life affects health | 0.69 | | |
| 2. During the last 12 months. | | | |
| a. The health professional's focus on prevention..... | | | 0.56 |
| b. The health professional's knowledge of women's health issues..... | | | 0.72 |
| c. The information I get about healthy living (such as diet and exercise)..... | | | 0.51 |
| d. The health professional's interest in my mental and emotional health..... | | | 0.61 |
| e. Help with finding information resources in women's health..... | | | 0.54 |
| f. How well my health care fits my stage of life..... | | | 0.71 |
| g. How well the health professional explain the results of tests or procedures... | | | 0.61 |
| h. The chance to see the same health prof. at each visit..... | | | 0.56 |
| i. The chance to get both gynecological and general health care here..... | | | 0.48 |
| j. My overall trust in the health professionals here..... | | | 0.75 |

Table 14: Reliability Analysis of the Communication Subscale AHRQ data

| | Scale Mean if Item Deleted | Corrected Item- Total Correlation | Squared Multi- ple Correlation | Alpha if Item De- leted |
|--|-------------------------------|--------------------------------------|-----------------------------------|-------------------------------|
| l: Visit today | | | | |
| g. The amount of time I had to talk with my health professional | 29.62 | 0.81 | 0.69 | 0.96 |
| h. My health professional's ability to answer questions in a sensitive | 29.42 | 0.90 | 0.87 | 0.96 |
| i. My health professional's ability to explain things clearly... | 29.41 | 0.91 | 0.88 | 0.96 |
| j. My health professional's ability to help me feel comfortable talking | 29.43 | 0.92 | 0.89 | 0.96 |
| k. The chance to ask all of my questions... | 29.49 | 0.89 | 0.82 | 0.96 |
| l. My health professional's ability to take what I say | 29.41 | 0.90 | 0.84 | 0.96 |
| m. My health professional's willingness to explain different options for | 29.68 | 0.82 | 0.73 | 0.96 |
| n. My health professionals interest in how my life affects health | 29.82 | 0.77 | 0.68 | 0.97 |

Table 15: Reliability Analysis of the Communication Subscale COE Data

| | Scale Mean if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Alpha if Item De- leted |
|--|-------------------------------|--|------------------------------------|-------------------------------|
| l: Visit today | | | | |
| g. The amount of time I had to talk with my health professional..... | 27.14 | 0.75 | 0.60 | 0.96 |
| h. My health professional's ability to answer questions in a sensitive and caring way... | 26.87 | 0.88 | 0.80 | 0.95 |
| i. My health professional's ability to explain things clearly..... | 26.87 | 0.87 | 0.79 | 0.95 |
| j. My health professional's ability to help me feel comfortable talking about my con- cerns | 26.88 | 0.89 | 0.83 | 0.95 |
| k. The chance to ask all of my questions... | 26.97 | 0.88 | 0.78 | 0.95 |
| l. My health professional's ability to take what I say seriously..... | 26.90 | 0.87 | 0.78 | 0.95 |
| m. My health professional's willingness to explain different options for my care... | 27.04 | 0.84 | 0.73 | 0.95 |
| n. My health professionals interest in how my life affects health | 27.17 | 0.79 | 0.65 | 0.96 |

Table 16: Reliability Analysis of the Office Administration Subscale AHRQ Data

| | Scale Mean if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Alpha if Item De- leted |
|--|-------------------------------|--|------------------------------------|-------------------------------|
| 1: Visit today | | | | |
| a. The courtesy of the staff..... | 19.01 | 0.76 | 0.60 | 0.85 |
| b. The staff's flexibility in scheduling my appointment around my needs..... | 19.06 | 0.75 | 0.60 | 0.85 |
| c. Privacy when talking to the receptionist | 19.46 | 0.69 | 0.49 | 0.86 |
| d. How well the staff kept you informed about the waiting time..... | 19.74 | 0.65 | 0.44 | 0.87 |
| e. Help with scheduling my next visit..... | 19.15 | 0.75 | 0.58 | 0.85 |
| f. The chance to talk to my health professional with my clothes on..... | 18.96 | 0.57 | 0.35 | 0.88 |

Table 17: Reliability Analysis of the Office Administration Subscale COE Data

| | Scale Mean if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Alpha if Item De- leted |
|---|-------------------------------|--|------------------------------------|-------------------------------|
| 1: Visit today | | | | |
| a. The courtesy of the staff..... | 17.07 | 0.67 | 0.44 | 0.83 |
| b. The staff's flexibility in scheduling my appointment around my needs..... | 17.36 | 0.64 | 0.43 | 0.83 |
| c. Privacy when talking to the receptionist | 17.53 | 0.60 | 0.37 | 0.84 |
| d. How well the staff kept you informed about the waiting time..... | 17.58 | 0.66 | 0.44 | 0.83 |
| e. Help with scheduling my next visit..... | 17.28 | 0.71 | 0.51 | 0.82 |
| f. The chance to talk to my health professional with my clothes on..... | 17.05 | 0.59 | 0.36 | 0.84 |

Table 18: Reliability Analysis of the Care Coordination and Comprehensiveness Subscale AHRQ data

| | Scale Mean if Item De- leted | Corrected Item-Total Correlation | Squared Multiple Correlation | Alpha if Item De- leted |
|--|---------------------------------------|--|------------------------------------|-------------------------------|
| 2. During the last 12 months. | | | | |
| a. The health professional's focus on prevention..... | 33.29 | 0.77 | 0.62 | 0.93 |
| b. The health professional's knowledge of women's health issues..... | 32.91 | 0.78 | 0.63 | 0.93 |
| c. The information I get about healthy living (such as diet and exercise)... | 33.39 | 0.73 | 0.60 | 0.93 |
| d. The health professional's interest in my mental and emotional health.... | 33.27 | 0.78 | 0.65 | 0.93 |
| e. Help with finding information resources in women's health..... | 33.39 | 0.75 | 0.60 | 0.93 |
| f. How well my health care fits my stage of life..... | 33.26 | 0.79 | 0.63 | 0.93 |
| g. How well the health professional explain the results of tests or procedures... | 33.24 | 0.71 | 0.52 | 0.93 |
| h. The chance to see same health professional at each visit..... | 33.20 | 0.65 | 0.45 | 0.93 |
| i. The chance to get both gynecological and general health care here..... | 33.05 | 0.69 | 0.52 | 0.93 |
| j. My overall trust in the health professionals here..... | 32.94 | 0.78 | 0.65 | 0.93 |

Table 19: Reliability Analysis of the Care Coordination and Comprehensiveness Subscale COE data

| | Scale Mean if Item De- leted | Corrected Item-Total Correlation | Squared Multiple Correlation | Alpha if Item De- leted |
|--|---------------------------------------|--|------------------------------------|-------------------------------|
| 2. During the last 12 months. | | | | |
| a. The health professional's focus on prevention..... | 33.32 | 0.71 | 0.53 | 0.92 |
| b. The health professional's knowledge of women's health issues..... | 32.99 | 0.76 | 0.61 | 0.92 |
| c. The information I get about healthy living (such as diet and exercise).... | 33.44 | 0.71 | 0.56 | 0.92 |
| d. The health professional's interest in my mental and emotional health.... | 33.33 | 0.76 | 0.62 | 0.91 |
| e. Help with finding information resources in women's health..... | 33.44 | 0.72 | 0.54 | 0.92 |
| f. How well my health care fits my stage of life..... | 33.31 | 0.74 | 0.57 | 0.92 |
| g. How well the health professional explain the results of tests or procedures... | 33.31 | 0.68 | 0.48 | 0.92 |
| h. The chance to see same health professional at each visit..... | 33.24 | 0.63 | 0.42 | 0.92 |
| i. The chance to get both gynecological and general health care here..... | 33.10 | 0.66 | 0.49 | 0.92 |
| j. My overall trust in the health professionals here..... | 33.01 | 0.76 | 0.64 | 0.92 |

Table 20. PCSSW Norms for CoE Evaluating Program by Age Group

| AGE | Communication Mean (Std) | Administration and Office Pro- cedures Mean (Std) | Care Coordi- nation Mean (Std) | N |
|------------|-------------------------------------|--|---|----------|
| 18-34 | 30.53 (6.90) | 20.62 (4.59) | 35.86 (7.63) | 964 |
| 35-44 | 30.55 (7.28) | 20.50 (4.69) | 35.96 (8.43) | 612 |
| 45-54 | 31.17 (7.18) | 20.65 (4.92) | 36.71 (8.47) | 690 |
| 55-65 | 31.10 (6.85) | 20.87 (4.90) | 36.94 (8.34) | 422 |
| 65+ | 30.16 (5.97) | 21.42 (3.93) | 36.59 (7.28) | 423 |

REFERENCES

- Aharony L, and Strasser S. 1993. "Patient Satisfaction: What We Know About and What We Still Need to Explore." Medical Care Review 50:49-79.
- Anderson RT, Weisman CS, Scholle SH, Henderson JT, Oldendick R, and Camacho F. 2002. "Evaluation of the Quality of Care in the Clinical Care Centers of the National Centers of Excellence in Women's Health." Women's Health Issues 12(6):309-326.
- Bean-Mayberry, BA, Chang CH, McNeil MA, Whittle JC, Hayes PM, Scholle SH. 2003. "Patient Satisfaction in Women's Clinics and Traditional Primary Care Clinics in the Veteran's Administration." Journal of General Internal Medicine, 18(3):175-181.
- Becker, LA, 1999. Becker, LA, 1999. "Measures of Effect Size (Strength of Association)." November 8, 1999. Available at: http://web.uccs.edu/lbecker/SPSS/glm_effectsize.htm. Accessed August, 2002.
- Bentler PM, and Bonnet DG. 1980. "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures." Psychological Bulletin 88:588-606.
- Browne MW, and Cudeck R. 1993. "Alternative Ways of Assessing Model Fit." Pp. 136-162 in KA Bollen and JS Long (eds.), Testing Structural Equation Models. Newbury Park, CA: Sage.
- Carmines EG, and McIver JP. 1981. "Analyzing Models with Unobserved Variables." In GW Bohrnstedt and EF Borgatta (eds.). Social Measurement: Current Issues. Beverly Hills, CA:Sage.
- Cleary PD, and McNeil BJ. 1988. "Patient Satisfaction as an Indicator of Quality Care." Inquiry 25:25-36.
- Davies AR, and Ware JE. 1991. GHAA's Consumer Satisfaction Survey and User's Manual. Washington, DC: Group Health Association of America.
- Donaldson MS. 2001. "Continuity of Care: A Reconceptualization." Medical Care Research and Review 58(3):255-290.
- Hall, J. and Dornan, MC. 1990. "Patient sociodemographic characteristics as predictors of satisfaction with medical care: A meta-analyses." Soc.Sci.Med., 30(7):811-818.
- Hays RD, Shaul JA, Williams VSL, et al. 1999. "Psychometric Properties of the CAHPS™ 1.0 Survey Measures." Medical Care 37(3) supplement: MS22-MS31.

- Horn, J.L, 1965. "A rationale and test for the number of factors in factor analysis. *Psychometrika*," 30:179-185.
- Hu L, and Bentler PM. 1995. "Evaluating Model Fit." In RH Hoyle (ed.), Structural Equation Modeling: Concepts, Issues, and Applications. Thousand Oaks, CA: Sage.
- Kaiser HF. 1970. "A Second Generation Little Jiffy." Psychometrika. 35:401-415.
- Kolodinsky J. 1997. "Gender Differences in Satisfaction with Primary Care Physicians in a Managed Care Health Plan." Women and Health 26:67-86.
- National Committee for Quality Assurance. 1998. HEDIS 1999 Volume 3: HEDIS Protocol for Administering CAHPS 2.0H Survey. Washington, D.C.: NCQA.
- Orlando M, and Meredith LS. 2002. "Understanding the Causal Relationship Between Patient-reported Interpersonal and Technical Quality of Care for Depression." Medical Care 40(8):696-704.
- Rubin HR, Gandek B, Rogers WH, et al. 1993. "Patient's Ratings of Outpatient Visits in Different Practice Settings: Results from the Medical Outcomes Study." Journal of the American Medical Association 270(7):835-840.
- Samejima F. 1997. "Models for Items with Polytomous Response Formats: Graded Response Model." In WJ Linden and RK Hambleton (eds.), Handbook of Modern Item Response Theory. New York: Springer-Verlag.
- Schauffler HH, Rodriguez T, and Milstein A. 1996. "Health Education and Patient Satisfaction." Journal of Family Practice 42(1):62-68.
- Sitzia J, and Wood N. 1997. "Patient Satisfaction: A Review of Issues and Concepts." Social Science and Medicine 45(12):1829-1843.
- Velicer, W.F. 1976, "The relation between factor score estimates, image scores, and principal component scores". Educational and Psychological Measurement 36:149-159.
- Ware, JE. and Davies, AR. 1983. "Behavioral consequences of consumer dissatisfaction with medical care." Evaluation and Program Planning, 6:291-297.
- Ware, JE and Hays,RD. 1988. "Methods for measuring patient satisfaction with specific medical encounters" Medical Care 26:393-402.

- Ware JE, and Sherbourne CD. 1992. "The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual Framework and Item Selection." Medical Care 30(6):473-483.
- Weisman, CS, and Henderson JT. 2001. "Managed Care and Women's Health: Access, Preventive Services, and Satisfaction." Women's Health Issues 11(3):201-215.
- Weisman CS, Henderson JT, Schiffrin E, Romans M, and Clancy CM. 2001. "Gender and Patient Satisfaction in Managed Care Plans: Analysis of the 1999 HEDIS/CAHPS 2.0H Adult Survey." Women's Health Issues 11(5):401-415.
- Weisman CS, Rich DE, Rogers J, et al. 2000. "Gender and Patient Satisfaction with Primary Care: Tuning in to Women in Quality Measurement." Journal of Women's Health and Gender-based Medicine 9(6):657-665.