Development and Validation of an Instrument to Measure Factors Related to Colorectal Cancer Screening Adherence

Sally W. Vernon, Ronald E. Myers, and Barbara C. Tilley

School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas 77030 [S. W. V.]; Jefferson Medical College, Thomas Jefferson University, Philadelphia, Pennsylvania 19107 [R. E. M.]; and Biostatistics and Research Epidemiology, Henry Ford Health Sciences Center, Detroit, Michigan 48202 [B. C. T.]

Abstract

This report describes the development and refinement of a set of scales for use in research on predictors of colorectal cancer screening adherence. The study population included 2693 of 4490 eligible white male automotive employees who answered a mailed questionnaire (60% response rate) on beliefs and attitudes related to colorectal cancer and screening. Exploratory and confirmatory factor analyses and multitrait scaling analysis were used to evaluate the construct validity of scales developed to measure salience and coherence, perceived susceptibility, worries about screening, screening efficacy, social influence, and intention. Analyses supported the construct validity of scales for salience and coherence, perceived susceptibility, and worries about screening. Four items originally assigned to the salience and coherence construct loaded on a separate factor that appeared to measure self-efficacy. There was no empirical support for scales measuring screening efficacy and social influence, and there was limited empirical support for a scale measuring intention. Confirmatory factor analysis of the scales measuring salience and coherence, self-efficacy, perceived susceptibility, and worries about screening showed a similar factor structure in white men and African-American men. Multitrait scaling analysis showed support for internal consistency reliability of those scales in white men (42) and in African-American men (46), and there was some support for the factor structure in those two subgroups. Future studies should evaluate the psychometric properties of these and similar scales in diverse population subgroups.

Introduction

To date, there are few published reports of systematic attempts to develop and validate instruments used to predict cancer screening adherence. Several reports have been published on scale development for mammography screening adherence (1–5), and there is recent work on scale development to predict adherence to cancer control regimens (6, 7). A number of studies of colorectal cancer screening adherence have used constructs from the Health Belief Model, as well as from other theoretical models to explain adherence, but scant information was provided on the reliability and validity of the psychosocial measures used in those studies. Although Myers and colleagues (8–11) developed a theory-based predictive model for cancer screening adherence, published data on the psychometric properties of the scales were limited. Even when investigators use the same construct, comparison of findings across studies is difficult, because different operational definitions often are used.

The purpose of this report is to describe the development and refinement of a set of items and scales to be used in research on predictors of colorectal cancer screening adherence.

Materials and Methods

Study Population. The target population for development and validation of the instrument described here consisted of employees from 28 worksites who worked in the pattern-making areas of an automobile manufacturer and who had been found to be at increased risk of colorectal cancer mortality compared with the United States white male population (12). The 28 worksites were part of a randomized clinical trial of behavior change interventions designed to increase colorectal cancer screening with digital rectal examination, fecal occult blood testing, and flexible sigmoidoscopy (recommended tests and procedures varied depending on the employee’s age and past screening history) and to change eating behaviors. Prior to initiating the interventions in year 1 of the study, a baseline self-report questionnaire was mailed to 5042 eligible current and former employees. The survey included questions on attitudes and beliefs related to colorectal cancer screening and eating behaviors. Nonrespondents were sent a reminder postcard after 2 weeks and were telephoned after four weeks. A second survey was mailed to employees who said they had not received the survey or who did not remember receiving it. Overall, 58% of all employees returned the baseline questionnaire. As reported elsewhere (13), respondents were more likely than nonrespondents to be older, married, currently working, not current smokers, to have undergone colorectal cancer screening in the 2 years prior to the survey, and to have a history of colorectal polyps.

The study population for this report consisted of 2693 of 4490 eligible white males who responded to the baseline survey (a 60% response rate) and who were free of colorectal cancer at baseline (60 white males, 43 of whom responded to the survey,
were excluded due to a reported diagnosis of colorectal cancer). Women \((n = 197)\) and African-American men \((n = 170)\) were excluded from the primary analysis to maintain homogeneity, but confirmatory factor analysis and multitrait scaling analysis were conducted on data from survey responders in these two subgroups \((n = 42 \text{ and } 56, \text{ respectively})\). Other nonwhite men \((n = 125)\) were excluded from all analyses.

**Background on Theory-based Framework and Instrument Development.** The items and scales for the current project came from two sources. Several constructs included in both research efforts described below have a long history in research on adherence with recommended health actions, including perceived susceptibility, perceived severity, benefits/costs or perceived utility from the Health Belief Model \((14-19)\), behavioral intention from the Theory of Reasoned Action \((20-22)\), and subjective norms or social influence from Social Cognitive Theory \((23)\).

In the late 1980s, researchers from the University of California, Los Angeles, under the auspices of the University of California, Los Angeles Cancer Control Science Program, developed a questionnaire to predict adherence to cancer control regimens including smoking cessation, cancer rehabilitation, adherence to follow-up for abnormal Pap tests, and adherence to a low-fat diet to prevent breast cancer \((24)\). The questionnaire was validated in four randomized trials of those outcomes, and seven constructs were identified that composed the Adherence Determinants Questionnaire, *i.e.*, interpersonal aspects of care, perceived utility, severity, susceptibility, subjective norms, intentions, and supports/barriers \((6)\). Other constructs measured in relation to adherence included health value, social desirability, and general adherence \((6)\).

Concurrently, researchers at Fox Chase Cancer Center \((Philadelphia, PA)\) were developing a questionnaire to study adherence to colorectal cancer screening recommendations, and the University of California, Los Angeles questionnaire was used to augment that effort \((8)\). In addition to the constructs listed above, three additional constructs were measured: a two-item construct called salience and coherence based on Antonovsky’s \((25)\) concept sense of coherence \((e.g., \text{“having a stool blood test is an important thing for me to do” and “doing annual stool blood testing makes sense to me”})\); a five-item construct measuring the efficacy of colorectal cancer screening \((e.g., \text{“stool blood testing is an effective way to find colorectal cancer early”}; \text{Ref. 9})\), and locus of control \((26)\).

Multitrait scaling analysis \((27)\) was used to form scales and to evaluate reliability and item discrimination across scales in two studies of adherence to colorectal cancer screening conducted at Fox Chase Cancer Center. One study was conducted in a population of 577 predominantly white male employees of a chemical manufacturing company, who were offered colorectal cancer screening with fecal occult blood testing and flexible sigmoidoscopy by their employer \((8)\). Internal consistency reliability was moderate to low for scales measuring salience and coherence \((0.74)\), intention \((0.81)\), perceived susceptibility \((0.41)\), social influence \((0.73)\), benefits/costs \((0.78)\), efficacy of colorectal cancer screening \((0.60)\), general \((self)\) efficacy \((0.57)\), and severity \((0.47)\). In another study, substantially the same questionnaire was evaluated in a sample of 501 men and women who were members of a health maintenance organization and who were offered free fecal occult blood testing by the health maintenance organization \((9)\). Internal consistency reliability was moderate to low for scales measuring salience and coherence \((0.78)\), intention \((0.67)\), perceived susceptibility \((0.60)\), social influence \((0.74)\), benefits/costs \((0.72)\), efficacy of colorectal cancer screening \((0.65)\), and general efficacy \((0.43)\). With the exceptions of perceived susceptibility and social influence, the other scales were correlated 0.30 or higher, indicating marginal discriminant validity in the health maintenance organization population. A scale measuring severity was dropped, because the items were uncorrelated in that study population.

**Refinement of the Instrument for the Current Study.** On the basis of the analyses described above, constructs and items were evaluated for relevance in the current study, *i.e.*, a predominantly male, blue-collar, employed population at increased risk of death from colorectal cancer. Constructs that were retained included salience and coherence, perceived susceptibility, perceived efficacy of colorectal cancer screening \((a\ construct\ that\ might\ also\ be\ conceptualized\ as\ benefits; \text{Ref. 28})\), social influence, and intention. The construct of general \((self)\) efficacy, *i.e.*, a belief in one’s ability or competence to undertake and complete a behavior of interest \((23, 29)\), was subsumed conceptually under salience and coherence \((8, 9)\). We added a measure of worries or fears about colorectal cancer screening, a construct that might also be conceptualized as barriers \((28)\). Perceived severity of the disease was not included, because it showed low internal consistency reliability and low variance in the studies of colorectal cancer screening described above. In a review of the literature on the Health Belief Model, Janz and Becker \((18)\) concluded that perceived severity may be of low relevance in relation to preventive health behaviors.

Thirty items, measured using a four-point Likert-type format from “strongly agree” to “strongly disagree,” were included in the baseline questionnaire to measure these six constructs. The Appendix lists the constructs, definitions, and the *a priori* assignment of items to scales.

Because of respondent burden reported during the baseline survey, we were interested in selecting a smaller subset of items for the follow-up surveys. Thus, we conducted a preliminary analysis of available data from the baseline survey, including item nonresponse, item variability, clarity of wording, redundancy of item content, internal consistency reliability, item-scale correlations, and an exploratory factor analysis. On the basis of these analyses, we eliminated items 4, 5, 8, 13, 18, 19, 26, 29, and 30. Items retained in the final statistical analysis reported here included 21 items \(\text{(shown in italics in “Appendix”).}\)

**Final Statistical Analysis.** The statistical analysis focused on the assessment of construct validity of the scales through exploratory and confirmatory factor analysis, multitrait scaling analysis, and correlation of scale scores with demographic variables and with past screening behavior.

The approach taken in the factor analysis was to divide the study population \(\text{(i.e., white men who were free of colorectal cancer at the time of the baseline survey)\ into two samples using a random number generator}.\) Exploratory factor analysis using the Statistical Analysis System procedure FACTOR \((30)\) was done on 21 items using sample 1 \(n = 1103\). Factors that accounted for at least 5% of the total variance \(\text{(eigenvalue for component of interest divided by total eigenvalues of the correlation matrix)\ were included in the varimax rotation}.\) After the rotation, items with factor loadings of 0.5 or less were excluded from the set of items used in the confirmatory factor analysis.

Sample 2 \(n = 1141\) was used to confirm the factor structure identified using sample 1. Because a high percentage of men \(\text{(approximately 25%)\ in the study population had a personal history of colorectal polyps and because this characteristic could differentially affect responses to the attitude and}}\)
belief items, separate confirmatory factor analyses also were conducted for subgroups of men with and without a personal history of polyps in both sample 1 and sample 2. Confirmatory factor analysis also was done for women (n = 42) and for African-American men (n = 56) who responded to the baseline survey.

Confirmatory factor analysis was conducted using the Statistical Analysis System procedure CALAIS (30), specifying the factor structure as identified in the exploratory analysis. Bentler’s (31) CFI and Bentler and Bonett’s (32) NNFI were used to assess the goodness of fit of the model. CFI and NNFI are less sensitive to sample size than many other measures. Values of CFI and NNFI of 0.9 or greater indicate a good fit of the factor model. We also reviewed test statistics for each factor loading (all should be >1.96), as described in detail by Hatcher (33).

Both the a priori assignment of items to constructs and constructs identified through the factor analysis were evaluated in light of the confirmatory factor analysis using the MAP (27). This software program was developed for the Medical Outcomes Study (34) to evaluate construct validity in scale development when multiple traits are assessed using a single method of measurement. Similar methods have been recommended for use in conjunction with exploratory factor analysis (35). The MAP software provides data on item and scale descriptive statistics (e.g., mean, SD, and variance), scale internal consistency estimates, item-scale correlations adjusted for overlap, and correlations among scales. Item convergence is assessed by item-scale correlations; a correlation, adjusted for overlap of 0.40 or greater is interpreted as support for item convergence (30). Item discrimination is supported if an item correlates higher with the designated scale than with the other scales under study and if this correlation is significantly larger than with other scales in the multitrait multi-item matrix. Following Hays et al. (27), an item was considered a “success” in the item discrimination analysis when the correlation between an item and its hypothesized scale was more than 2 SE higher than its correlation with other scales. It was considered a “probable” scaling error if its correlation with the hypothesized scale was within 2 SE of another scale and a “definite” scaling error if its correlation with the hypothesized scale was more than 2 SE below its correlation with another scale. Internal consistency reliability was estimated by Cronbach’s coefficient α. A coefficient of 0.70 or greater was used to identify scales with adequate internal consistency reliability. Scale-scale correlations, adjusted for attenuation (unreliability of measurement), using methods of Steiger (36), should be less than unity to support the distinctiveness of the scale (27). We hypothesized that intention would be correlated positively with salience and coherence, self-efficacy, and perceived susceptibility.

The construct validity of the scales was assessed further by correlating scale scores with each other and with age, education, and past screening behavior. We expected that the scales would be associated positively with past screening behavior and that age and education would be uncorrelated with the scales. Age and education were analyzed as continuous variables using the Pearson correlation coefficient. Past colorectal cancer screening behavior, defined as having been screened at least once in the 2 years prior to the baseline survey (coded 1) versus having not been screened in the 2 years prior to the baseline survey (coded 0), was analyzed in relation to scale scores using point biserial correlation. In addition, mean scale scores for men screened at least once in the 2 years prior to the baseline survey were compared with those for men who were not screened in the 2 years prior to the baseline survey; differences were tested for statistical significance using t tests adjusted for unequal variances. Due to concerns about respondent burden, test-retest reliability was not evaluated.

Analyses reported here were conducted ignoring the clustering effect (worksites were the sampling units). To assess the effect of clustering on the factor analysis, we conducted separate factor analyses by worksite for worksites with more than 100 white male employees. With minor exceptions, the factor structure in those six worksites was similar to the factor structure identified in the exploratory factor analysis of sample 1. The clustering effect was evaluated for mean scale score differences using ratio estimates (37). Standard errors and results of statistical tests taking clustering into account were similar to the unadjusted standard errors and statistical tests, so the unadjusted values are reported here.

**Results**

**Factor Analysis.** As shown on Table 1, the exploratory factor analysis based on 21 items resulted in five factors. Four of the seven items (i.e., items 2, 7, 11, and 12) loading on factor 1 were assigned a priori to the salience and coherence scale (see “Appendix”). Both items measuring the intention construct.
The three-item measure of perceived susceptibility showed good item convergence and discrimination and good internal consistency reliability in MAP runs based on both \textit{a priori} and factor-derived scales (Table 3, MAP run 1). The three-item \textit{a priori} worries scale showed poor item convergence and discrimination (Table 3, MAP run 1). Internal consistency reliability was lowered by inclusion of item 25 in the worries scale compared with the factor-derived scale (Table 3, MAP runs 1 and 2).

The three-item \textit{a priori} scales for efficacy of screening and social influence showed poor item convergence and discrimination and low internal consistency reliability (Table 3). Because those two scales were not supported by the factor analysis, no MAP runs were done on factor-derived versions.

MAP run 1 provided some support for the construct validity of the \textit{a priori} two-item intention scale; item convergence and internal consistency were good, but item discrimination was not (Table 3). Item 15 correlated more strongly with the total score of the salience and coherence scale (0.72) than with the intention scale (0.65). Omitting items 10 and 15 from the salience and coherence scale, however, did not greatly affect internal consistency reliability for that scale; Cronbach’s \( \alpha \) was 0.88 for a four-item salience and coherence scale (items 2, 7, 11, and 12).

On the basis of the results of the confirmatory factor analysis and of the multitrait scaling analysis, the following scales will be retained for predictive analyses of adherence to colorectal cancer screening: salience and coherence (items 2, 7, 11, and 12), self-efficacy (items 6, 14, 21, and 28), perceived susceptibility (items 1, 17, and 24), worries (items 16 and 23), and intention (items 10 and 15). Although there was limited empirical support for the intention scale, behavioral intention is an important theoretical construct (21) that is conceptually distinct. The items in the intention scale have face validity, and exclusion of those two items did not compromise the psychometric properties of the salience and coherence scale. Descriptive statistics for the final versions of the scales are shown in Table 4.

**Construct Validity.** Table 5 shows correlation coefficients between the final versions of the scales and between the scales and age, education, and past screening behavior. Perceived susceptibility and worries showed low correlation with the other scales, indicating good discriminant validity. As expected from the factor analysis, salience and coherence was strongly and positively correlated with intention, and it also showed a strong positive correlation with self-efficacy. Likewise, intention and self-efficacy were strongly correlated, but these coefficients did not approach unity. The demographic variables of age and education showed low correlation with all scales (Table 5).

The correlation coefficients between the scales and past screening behavior was moderate for the salience and coherence, self-efficacy, and intention scales (Table 5). Past screening status showed low correlation with perceived susceptibility and was uncorrelated with the worries scale. That analysis was consistent with results based on comparison of mean scores that showed statistically significant differences between groups for all scales but worries (Table 6).

With the exceptions of the intention (\( \alpha = 0.67 \)) and worries (\( \alpha = 0.42 \)) scales in African-American men, coefficients were 0.70 or greater for women and for African-American men; however, the patterns for item convergence and discrimination were not as consistent as for white men.
Table 3  Multitrait scaling analysis for scales based on a priori assignment and on confirmatory factor analysis (n = 2244)

<table>
<thead>
<tr>
<th>Scales*</th>
<th>Item numbersa</th>
<th>Internal consistency reliability</th>
<th>Item convergenceb</th>
<th>Item discriminationc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salience and coherence</td>
<td>1 (2, 6, 7, 11, 12, 14, 21, 28)</td>
<td>0.88</td>
<td>0.58-0.71</td>
<td>(2', 11', 12')</td>
</tr>
<tr>
<td></td>
<td>2 (7, 10, 11, 12, 15)</td>
<td>0.91</td>
<td>0.65-0.84</td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>1</td>
<td>Not applicabled</td>
<td>0.82</td>
<td>0.65-0.70</td>
</tr>
<tr>
<td></td>
<td>2 (6, 14, 21, 28)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived susceptibility</td>
<td>1 (1, 17, 24)</td>
<td>0.79</td>
<td>0.55-0.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (1, 17, 24)</td>
<td>0.79</td>
<td>0.55-0.70</td>
<td></td>
</tr>
<tr>
<td>Worries/concerns</td>
<td>1 (16, 23, 25)</td>
<td>0.64</td>
<td>0.33-0.53</td>
<td>(25°)</td>
</tr>
<tr>
<td></td>
<td>2 (16, 23)</td>
<td>0.72</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>Efficacy of screening</td>
<td>1 (3, 20, 27)</td>
<td>0.42</td>
<td>0.17-0.35</td>
<td>(27°)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Not applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social influence</td>
<td>1 (9, 22)</td>
<td>0.57</td>
<td>0.41</td>
<td>(22&quot;)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Not applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intention</td>
<td>1 (10, 15)</td>
<td>0.79</td>
<td>0.65</td>
<td>(15&quot;)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Not applicable</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Scoring procedure: Items were scored on a four-point scale from "strongly disagree" to "strongly agree." Item scores were summed, and the total score was divided by the number of items in the scale to standardize scale scores for differences in the number of scale items. Because item values ranged from 1 to 4, scale scores also ranged from 1 to 4. Persons with missing data on some of the scale items were assigned a scale score as follows: for three-item and four-item scales, persons with two or more responses were assigned a scale score; for two-item scales, persons had to respond to both items or be deleted from the analysis.

* Questionnaire items are listed in the Appendix.

* Range of item-scale correlation coefficients for multitrait scaling analysis runs.

* Numbers designate the multitrait scaling analysis run for different groupings of items and scales corresponding to: a priori assignment (1) and confirmatory factor analysis based on the 15-item, four factor model (2).

* Counted as a "probable" scaling error; the correlation between the item and another scale was within 2 SE of its correlation with its hypothesized scale.

* Scale was not included in the multitrait scaling analysis run.

* Counted as a "definite" scaling error; the correlation between the item and its hypothesized scale was more than 2 SE below its correlation with another scale.

Table 4  Descriptive statistics for final versions of the scales for colorectal cancer screening adherence

<table>
<thead>
<tr>
<th>Item numbers</th>
<th>Meand</th>
<th>SD</th>
<th>Median</th>
<th>Skew</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salience and coherence</td>
<td>(2, 7, 11, 12)</td>
<td>3.58</td>
<td>0.57</td>
<td>4.00</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>(6, 14, 21, 28)</td>
<td>3.13</td>
<td>0.74</td>
<td>3.25</td>
</tr>
<tr>
<td>Perceived susceptibility</td>
<td>(1, 17, 24)</td>
<td>2.43</td>
<td>0.73</td>
<td>2.33</td>
</tr>
<tr>
<td>Worries</td>
<td>(16, 23)</td>
<td>3.06</td>
<td>0.80</td>
<td>3.00</td>
</tr>
<tr>
<td>Intention</td>
<td>(10, 15)</td>
<td>3.44</td>
<td>0.79</td>
<td>4.00</td>
</tr>
</tbody>
</table>

* Scale means are standardized for the number of items included in the scale.

Discussion
In general, results of the factor analysis supported the a priori assignment of items to scales. Notable exceptions were the failure of the social influence items to factor together and the lack of discrimination between the items measuring salience and coherence and intention. It also was the case that four items from the longer version of the salience and coherence scale formed a factor that appeared to measure self-efficacy. The multitrait scaling analysis generally supported these results, in that the social influence scale showed low internal consistency reliability, and the self-efficacy scale showed high reliability. Although two of the three items measuring the efficacy of screening factored together in the exploratory factor analysis, the loadings did not meet our cutoff for inclusion in the confirmatory factor analysis, and multitrait scaling analysis showed little support for the a priori scale.

Although intention, salience and coherence, and self-efficacy are conceptually distinct, they were relatively highly correlated with each other, and the correlation coefficients with past screening behavior were of similar magnitude. Theoretically, intention is an important variable, because it is hypothesized to be an immediate and necessary precursor of behavior (21). As such, it may be important to understand the patterns of association between attitudes and beliefs and intention to develop effective interventions to influence behavior change. Moreover, there is empirical evidence that predictors or correlates differ for intention and for behavior, e.g., completion of screening (10, 38–40). The usefulness of the scales should be evaluated further in studies to predict prospective adherence to colorectal cancer screening recommendations. If they are found to predict adherence, they may be useful in identifying subgroups of persons who require encouragement (e.g., those with low intention) or whose concerns about screening need to be addressed (e.g., those who score high on worries about screening).

Janz and Becker (18) suggested that there is overlap in the...
Table 5. Correlation coefficients for final versions of the scales for colorectal cancer screening adherence (n = 2244)

<table>
<thead>
<tr>
<th>Scales and other factors</th>
<th>Intention</th>
<th>Salience and coherence</th>
<th>Self-efficacy</th>
<th>Perceived susceptibility</th>
<th>Worries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salience and coherence</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>0.55</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived susceptibility</td>
<td>0.27</td>
<td>0.27</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worries</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.15</td>
<td>-0.21</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.03</td>
<td>0.06</td>
<td>0.15</td>
<td>0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>Education</td>
<td>0.08</td>
<td>0.11</td>
<td>0.06</td>
<td>0.04</td>
<td>-0.01</td>
</tr>
<tr>
<td>Past screening status*</td>
<td>0.42</td>
<td>0.38</td>
<td>0.37</td>
<td>0.14</td>
<td>0.01</td>
</tr>
</tbody>
</table>

* Scale-scale correlation coefficients were corrected for attenuation (36).

* Past colorectal cancer screening status was defined as "ever screened" in the two years prior to the baseline survey versus "never screened" in the 2 years prior to the baseline survey. Point biserial correlations.

As noted by Champion (1) in discussing breast cancer screening, even when investigators used the same construct, operational definitions often varied from study to study, and evidence to support the reliability and validity of the measures was rarely provided. We found this circumstance to pertain to the literature on colorectal cancer screening as well. Although a number of studies investigated the association between psychosocial constructs and colorectal cancer screening behavior (9, 10, 38, 39, 41-49), only a few provided any information on the reliability or validity of the measures used (9, 42).

The similarity of the factor structure across subgroups of white men with and without polyps indicates that these scales could be useful both for screening and for surveillance because disease status does not appear to affect response patterns. There was less consistent evidence for acceptable psychometric properties of the scales for women and for African-American men. Although internal consistency reliability was adequate for most scales, coefficients were, in general, lower, and the patterns of item convergence and discrimination in the multitrait scaling analysis were less consistent than for white men. Furthermore, there was limited support for the factor structure identified for white men in those two subgroups.

Strengths of our study include a large sample size, permitting stable estimates in most subgroups of the psychometric characteristics evaluated, and a study population that included men with and without a personal history of polyps so that the measures could be evaluated for their usefulness in screening and surveillance. The fact that the study population was identified as being at increased risk for colorectal cancer may limit the generalizability of the results. Likewise, because the program offered multiple colorectal cancer screening procedures, the questionnaire items were not specific to the type of screening procedure, e.g., fecal occult blood testing and flexible sigmoidoscopy. It may be the case that beliefs and intentions about colorectal cancer screening differ for different types of procedures. Although the wording might affect responses to some scales (e.g., salience and coherence and intention), it would not affect responses to others (e.g., perceived susceptibility and worries). In circumstances when only one test or procedure is recommended, the survey questions could be modified to be specific to the procedure. Other limitations of the study were the potential for nonresponse bias, because only 60% of the study population completed the questionnaire on which the analysis was based, and our decision not to assess test-retest reliability. The implications of nonresponse for estimates of reliability and validity of the constructs measured are not known, but a likely consequence is that the distribution of scores was skewed toward the more favorable end of the response continuum, affecting the generalizability of results to the type of person who responds to surveys. This assumption is consistent with findings by others that persons who completed colorectal cancer screening were more likely to complete questionnaires (53, 54-58). Several studies found that persons with healthier lifestyles tended to participate in worksite health promotion programs (53-55), but the evidence was not consistent on this point (56). Studies conducted at worksites found no differences in health status between participants and nonparticipants in health promotion programs (53, 55-58). Because researchers conducting similar surveys will inevitably face the
same constraint, the effects of nonresponse on the psychometric characteristics of the scales, if any, would be a worthwhile focus for future research efforts. Future research efforts would benefit from attention directed to clarifying conceptual definitions, with particular reference to specifying the overlap in content domains for constructs from different behavior change models. There also is a need to re-evaluate the content validity of items used to measure clearly defined constructs and to systematically evaluate the reliability and validity of the operational definitions. Finally, future studies should evaluate more thoroughly the reliability and validity of these and similar scales in other population groups including women, minorities, and persons at average risk.

Acknowledgments
We acknowledge the support of the following persons: Brenda McGowan for secretarial assistance in preparing the manuscript, Collete Miesse for assistance with the reference list, and Donna Mott and Toni Chociemski for assistance with computer programming and data management support. We thank Karen Glanz for her comments on an earlier version of the manuscript.

Appendix

Attitude and Belief Items Related to Colorectal Cancer and Screening Included in the Baseline Surveya

Salience and Coherence: Perceptions about the Technical Effectiveness, Practical Convenience, Personal Benefit, and Whether or Not the Behavior Is Actively Encouraged by Significant Others.

(2) Doing colorectal cancer screening makes sense to me. (Sense)
(6) Arranging my schedule to go through colorectal screening is an easy thing to do. (Arrange)
(7) I think the benefits of colorectal screening outweigh any difficulty I might have in going through the tests. (Benefits)
(8) I think that going through colorectal screening is too much trouble for what I'll get out of it. (Difficult)
(11) Going through screening is an important thing for me to do. (Important)
(12) I believe that colorectal screening can help to protect my health. (Protect)
(14) Finding time to go through colorectal screening would be difficult for me to do. (Time)
(18) I'll be just as healthy if I avoid being tested for colorectal cancer. (Present)
(19) I believe that colorectal screening is an easy way to find colorectal cancer early. (Cure)
(21) Going through colorectal screening would be difficult for me to do. (Difficult)
(22) Doing colorectal screening does not make sense to me. (Sense)
(28) I think that going through colorectal screening would be an easy thing for me to do. (Easy)

Intention: Expressed Intent to Undertake a Specific Behavior.

(10) I do not intend to go through colorectal screening. (Don’t intend)
(15) I intend to undergo colorectal screening. (Do intend)
(29) In the future I intend to have a colorectal screening test given by the worksite program. (Intend)
(30) In the future I intend to have a colorectal screening test outside the worksite program given by my private physician, medical clinic, hospital, or other health professional. (Intend)

Perceived Susceptibility: Subjective Personal Risk of Developing Colorectal Cancer or Polyps.

(1) I believe that the chance I might develop colorectal cancer is high. (Chance of cancer)
(5) I think that compared to other persons my age, I am at lower risk for colorectal cancer. (Chance)
(17) I think it is very likely that I will develop colorectal cancer or polyps. (Change)
(24) I believe that the chance that I will develop colorectal polyps is high. (Chance of polyps)

Worries and Fears (Barriers): Perceived Negative Components or Consequences of Undertaking and Completing the Behavior.

(16) I am afraid of having an abnormal screening test result. (Afraid)
(23) I am worried that screening will show that I have colorectal cancer or polyps. (Worried)
(25) I am bothered by the possibility that screening might be physically uncomfortable. (Discomfort)

Efficiency of Screening (Benefits): Perceived Positive Components or Consequences of Undertaking and Completing the Behavior.

(5) I believe that if I had a normal screening test result, I wouldn't have to worry about developing colorectal cancer. (Not worried)
(20) I think that when colorectal polyps are found and removed, colorectal cancer can be prevented. (Prevent)
(27) I believe that when colorectal cancer is found early, it can be cured. (Cure)

Social Influence: Role of Social Norms and Interaction with Members of One's Social Network.

(4) My close friends think I should go through colorectal screening. (Family thinks)
(9) I want to do what members of my immediate family think I should do about colorectal screening. (Family think)
(13) I want to do what my close friends think I should do about colorectal screening. (Close friends think)
(12) Members of my immediate family think I should go through colorectal screening. (Family thinks)

References

a Items are listed by a priori assignment to constructs. Items in italics were included in the final scale development analysis. Numbers and names in parentheses correspond to those listed on Table 1.

b Item was reverse coded.
Measuring Colorectal Cancer Screening Adherence

Development and validation of an instrument to measure factors related to colorectal cancer screening adherence.

S W Vernon, R E Myers and B C Tilley

Cancer Epidemiol Biomarkers Prev 1997;6:825-832.

Updated version  Access the most recent version of this article at: http://cebp.aacrjournals.org/content/6/10/825

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.