



Inpatient (IN) Survey Psychometrics Report

Background

One of the hallmarks of Press Ganey's surveys is their scientific basis: Our products incorporate the best characteristics of survey design. Our surveys are developed by conducting patient focus groups, reviewing surveys from health care facilities across the country, soliciting feedback from physicians, nurses and administrators, reviewing current professional and scientific publications on health care delivery, and utilizing the latest research on survey statistics and design.

The Inpatient Survey was first developed in 1987. In 1997, it was revised to ensure that it continued to accurately reflect patients' experiences and provide useful data to our clients. The 49-question survey was fully tested in 1997 and has a robust factor structure and high subscale reliability. However, the environment for measuring patient perceptions has changed. Not all of these questions continue to be vital to understanding patient views, and hospitals have additional needs for incorporating standardized measurement such as HCAHPS within their patient survey.

In identifying a standard question set, which is the basis of comparative benchmarking, Press Ganey strives to meet the current needs of the greatest number of clients and reflect the changing needs of health care delivery. We would like to support our clients in a way that provides them with the best information possible for quality improvement purposes in a broad sense, as well as targeted quality improvement related to HCAHPS efforts.

Press Ganey has spent two years conducting research to understand the best way to meet client needs for measuring patient-centeredness in light of the HCAHPS environment. During the testing of the HCAHPS integrated survey in the Summer of 2005 and the implementation of the HCAHPS Dry Run during the Spring and Summer of 2006, Press Ganey tested shorter sets of standard questions that produced section and overall scores similar to the full Press Ganey survey. Press Ganey has evaluated its full survey to determine which questions could be removed and still provide a full quality improvement tool.

As a result of these analyses, eleven questions are being removed from the standard question set to take effect in early 2007. Questions were removed after individual consideration if they were found to be non-universal, to not reflect the current structure of the health care experience, to be highly correlated with another item and/or to have a lower factor loading than the other questions in the same section.



We have heard from clients that they would like to have the same set of standard questions on all of their Inpatient surveys, regardless of the HCAHPS initiative, and that they would like to customize their surveys more extensively. The new format for the integrated survey along with the reduced question set allows for all of these goals to be realized. The result is a standard question set that gathers data on the most important aspects of inpatient care while allowing maximum flexibility with customization to a hospital's specific needs.

The Testing Process

Effective questionnaires have three important attributes: focus, brevity, and clarity. Questions should focus directly on a specific issue or topic, be as brief as possible while still conveying the intended meaning, and be expressed as simply and as clearly as possible. The original 49 questions met these criteria and their **face**, **content**, and **consensus validities** were established in the original development work and all subsequent re-validations that ensured that the instrument remained accurate over time.

Our analyses led us to remove the 11 questions from the standard survey. Please see the following page for a list of the items that were removed and the rationale for doing so.

Table 1. Items Removed From the Standard Question Set

Item Removed	Rationale for Removal
Rating of pre-admission process (if any)	This question is non-universal; not all patients have this experience.
How well things worked (TV, call button, lights, bed, etc.)	This question has the lowest factor loading for the Room section meaning that it contributes the least to the Room Section concept.
If you were placed on a special/restricted diet, how well it was explained	This question is non-universal; not all patients have this experience. There is also potential for patients to misinterpret this question. Patients could interpret any change in diet as 'special' and this question might imply that they should have received more extensive education than is typically warranted.
Concern shown for your comfort during tests or treatments	This question is highly correlated with other Tests & Treatments questions and the pain control question.
Skill of the person who took your blood (e.g., did it quickly, with minimal pain)	This question is highly correlated with the question regarding the courtesy of the person who took your blood and has a lower factor loading than the courtesy question.
Skill of the person who started the IV (e.g., did it quickly, with minimal pain)	This question is highly correlated with the question regarding the courtesy of the person who started the IV and has a lower factor loading than the courtesy question.
Helpfulness of the people at the information desk	This question is non-universal; many hospital entrances do not include an information desk.
Information given to your family about your condition and treatment	Hospitals are limited in the information they can share with others about the patient due to HIPAA concerns.
Help with arranging home care services (if needed)	This question is non-universal; not all patients have this experience. This question could be misinterpreted by patients to imply that assistance with home care should have occurred even when unnecessary.

Table 1 (continued). Items Removed From the Standard Question Set

Staff sensitivity to the inconvenience that health problems and hospitalization can cause	This question is highly correlated with other questions in the Personal Issues section. Because of the global nature of this question, it is difficult to prescribe a specific action to address this. Removal of this question does not adversely affect the reliability or validity of the Personal Issues scale.
Overall cheerfulness of the hospital	This question is highly correlated with other questions in the Overall Assessment section. Because of the global nature of this question, it is difficult to prescribe a specific action to address this. Removal of this question does not adversely affect the reliability or validity of the Overall Assessment scale.

The remaining 38 questions were tested using our standard mail-out/mail-back methodology to eliminate selection and acquiescence biases.

As with other Press Ganey questionnaires, a Likert-type response format was used with the following categories: very poor, poor, fair, good, and very good. Because this format is balanced and parallel—unlike a “poor” to “excellent” format—responses can be quantified and used statistically without violating methodological assumptions. Also, variability in patients’ responses with this format allows for the identification of opportunities to improve, unlike “yes/no” response formats.

The validation of the 38-item Inpatient tool is described in this document. The analyses were based on 565,557 surveys returned to Press Ganey during the first quarter of 2005, representing 1,436 hospitals.

Results stemming from the test of the reduced Inpatient Survey showed a robust scale that consisted of ten internally consistent subscales: Admission, Room, Meals, Nurses, Tests and Treatments, Visitors and Family, Physician, Discharge, Personal Issues, and Overall Assessment.

Questionnaire Psychometrics

The accuracy of a questionnaire is assessed by measuring its validity and reliability. Validity is the degree to which a questionnaire measures what it was designed to measure. Reliability is the degree to which survey data are consistent across respondents or across surveys. The Inpatient instrument was once again found to be psychometrically sound across a wide variety of tests of validity and reliability as described below.

Response Patterns and Variance.

Measures of central tendency (i.e., mean, median, and mode) and variability (standard deviation, standard error) were examined for all questions. For a subsample of returned surveys, response frequencies and patient comments were evaluated for patterns of missing data and question-wording, or ambiguity, problems. Although no problems were encountered, these steps are important for ensuring the clarity of questions and the absence of instrument bias and error.

As mentioned above, it is important for variation to exist among each question’s responses. If every respondent answered “yes” or circled “4” for a particular question, there would be no variability and no opportunity to correlate that issue with others. If all respondents rated a service a “5,” it would suggest high service quality but little opportunity for improvement. In general, questions with little or no variation



should be omitted, and questions allowing facilities room to improve should be maintained. None of the questions had little or no variation; subsequently, all questions were included in further analyses.

Measure Redundancy.

As discussed above, it also is important for a survey instrument to be concise. A correlation matrix of all test questions was examined to find questions that were too highly associated with one another. If two questions are highly intercorrelated, it suggests that they may be measuring the same issue or concept and, therefore, redundant. Removing one of the two questions removes this redundancy and leaves a more parsimonious scale or questionnaire. None of the questions was found to be too highly correlated with other questions; subsequently, all questions were included in additional analyses.

Construct Validity.

Factor analysis is a technique used to identify factors that statistically explain the variation among responses to a questionnaire. In other words factor analysis helps to identify which questions belong together, confirming a questionnaire’s **construct validity**, or structure.

Ideally, the factor analysis should place questions in groupings similar to the sections of the questionnaire. Questions that are highly correlated with one another typically represent a common dimension or concept. For example, “Speed of admission process” and “Courtesy of the person who admitted you” are more likely to define an “Admission” dimension than “Nurses’ attitude toward your requests” or “Time the physician spent with you.”

A factor analysis identified nine factors that accounted for 73% of the total variance in patient responses. A principle component extraction was performed with Promax oblique rotation. Overall evaluation questions were omitted from the factor analysis due to their (predicted) high intercorrelations with other items. These factors paralleled the structure of sub-scales on the questionnaire. Factor loadings in Table 2 represent the correlations between factors and variables

Table 2. Item Content and Primary Factor Loadings

Survey Section and Survey Item	Factor Loadings
Admission	
Speed of admission process	0.95
Courtesy of the person who admitted you	0.84
Room	
Pleasantness of room decor	0.78
Room cleanliness	0.82
Courtesy of the person who cleaned your room	0.66
Room temperature	0.75
Noise level in and around room	0.74
Meals	
Temperature of the food (cold foods cold, hot foods hot)	0.92
Quality of the food	0.95
Courtesy of the person who served your food	0.58



Table 2 (continued). Item Content and Primary Factor Loadings

Nurses	
Friendliness/courtesy of the nurses	0.90
Promptness in responding to the call button	0.83
Nurses' attitude toward your requests	0.92
Amount of attention paid to your special or personal needs	0.87
How well the nurses kept you informed	0.78
Skill of the nurses	0.83
Tests & Treatments	
Waiting time for tests or treatments	0.48
Explanations about what would happen during tests/treatments	0.64
Courtesy of the person who took your blood	0.93
Courtesy of the person who started the IV	0.90
Visitors & Family	
Accommodations and comfort for visitors	0.93
Staff attitude toward your visitors	0.83
Physician	
Time physician spent with you	0.88
Physician's concern for your questions and worries	0.95
How well physician kept you informed	0.94
Friendliness/courtesy of physician	0.93
Skill of physician	0.82
Discharge	
Extent to which you felt ready to be discharged	0.80
Speed of discharge process after told you could go home	0.91
Instructions given about how to care for yourself at home	0.72
Personal Issues	
Staff concern for your privacy	0.56
How well your pain was controlled	0.72
Degree to which hospital staff addressed your emotional needs	0.82
Response to concerns/complaints made during your stay	0.75
Staff effort to include you in decisions about your treatment	0.77
Extraction Method: Principle Component. Rotation Method: Promax.	

Convergent and Discriminant Validities.

Although factor analysis is a popular method of establishing the construct validity of a questionnaire, other methods are available. For example, one of the assumptions of questionnaire and scale construction is that an individual item from a scale should be well correlated with the other items in that scale.

Researchers suggest a minimum "item-to-scale" correlation of .30.



Table 3 shows the average and range of correlations between each question and its parent section. These correlations are “corrected” in the sense that the item of interest is omitted from its section score when correlations between the two are calculated. Corrections are performed to avoid inflated or spuriously positive correlations.

Table 3. Item Analyses and Reliability Estimates

Subscale	Alpha	Average Corrected Item-Scale Correlations	Range of Corrected Item-Scale Correlations	Average Item-Non-Scale Correlations	Range of Item-Non-Scale Correlations
Admission	.77	.62	.62	.40	.33-.48
Room	.83	.63	.56-.71	.42	.29-.53
Meals	.80	.66	.55-.72	.41	.28-.50
Nurses	.95	.84	.79-.88	.53	.36-.73
Tests and Treatments	.84	.68	.61-.71	.50	.41-.63
Visitors and Family	.84	.72	.72	.52	.38-.67
Physician	.95	.85	.78-.91	.45	.32-.54
Discharge	.78	.62	.60-.66	.46	.33-.62
Personal Issues	.92	.79	.72-.84	.57	.39-.76
Overall Assessment	.94	.87	.84-.90	.59	.45-.77

An item should be correlated with its own scale (**convergent validity**) as well as correlated more with its own scale than with other scales (**discriminant validity**). That is, item-to-scale correlations should be higher than item-to-*non*-scale correlations. For example, a “Meals” question should be more highly correlated to its section than to the “Nurses” section (or any other section for that matter).

As Table 3 confirms, the Inpatient questionnaire demonstrates both convergent validity and discriminant validity. Because these are subclasses of construct validity, the questionnaire’s effectiveness at measuring patients’ perceptions is confirmed across multiple tests.



Criterion, or Predictive, Validity.

Predictive validity is defined as the ability of an instrument to predict outcomes that theoretically should be tied to the construct measured by the instrument. In consumer satisfaction, the perception of a satisfying experience is expected to be linked to two types of behavioral outcomes in an individual: re-purchase behavior (coming back at a future date to purchase another product or service) and positive word of mouth (recommending a product to family and friends).

Within the health care industry, limitations on patients' choice of providers and health services make it difficult to assess the relationship between satisfaction and re-purchase behavior. Patients may be very satisfied but not return to a provider or institution because their insurance plans either don't cover the service or require that they receive the medical care from a different provider. An estimate of the relationship between satisfaction and positive word of mouth can be obtained by asking patients about their intent to recommend a hospital. The predictive validity of a patient satisfaction instrument then can be estimated by the degree to which individual items on the instrument predict the patients' intentions to recommend.

In this regard the questionnaire shows high levels of predictive validity. A series of simple regression analyses revealed that each of the individual items is a significant predictor (at the .001 level) of patients' reported likelihood to recommend the facility (beta's ranged between .35 and .85). Additionally, multiple regression analysis revealed that, collectively, all items are significant predictors of patients' reported likelihood to recommend the hospital, $F(37, 565519) = 46373.744, p < .001, R^2 = .75$ (adjusted $R^2 = .75$). In other words, the revised instrument explains approximately 75% of the variance in patients' likelihood to recommend the hospital they visited. According to Carey¹ "A good survey will account for 50% or more of the variance in global evaluations..." The Inpatient Survey has surpassed this high standard.

Reliability.

Reliability testing is a method of evaluating the internal consistency of a questionnaire. The traditional statistic used to illustrate the degree of consistency among the items of a scale or questionnaire is Cronbach's alpha. A set of questions with no internal consistency—no reliability—has an alpha of 0.0 indicating that the questions within the scale may not be measuring the same issues. A set of questions with perfect internal consistency has an alpha value of 1.0.

All ten scales exceeded the stringent .70 standard for reliable measures: Reliability estimates range from .78 to .95 (see Table 2). The Cronbach's alpha for the entire questionnaire is .97, confirming the instrument's high internal consistency and reliability.

Readability.

According to the Flesch-Kincaid Index, which is based on the average number of syllables per word and words per question, the Inpatient questionnaire tests between a fifth- and sixth-grade reading level.

¹ Carey, R. G. (1999). How to choose a patient survey system. *The Joint Commission Journal of Quality Improvement*, 25, 20-25.