

Use of an Open-Ended Question to Supplement a Patient Satisfaction Questionnaire in a Medical Residents' Clinic

Bruce D. Bialor, MD; Joseph L. Musial, MPA; Graciela E. Rojas, MD; and Mark J. Fagan, MD

Abstract

Objectives: To determine (1) the proportion of responses to an open-ended question related to patient satisfaction that could be categorized into 1 or more of 9 previously developed domains of outpatient care and (2) whether any other important aspects of care could be identified by adding the open-ended question to a satisfaction questionnaire.

Study Design: A 3-month observational study was done at the internal medicine clinic of an urban teaching hospital.

Patients and Methods: As part of a patient satisfaction study, 511 visitors were asked after their visit, "What are the 1 or 2 things that are most important to you when you see a doctor?" The responses were categorized independently by 2 raters into 1 or more of the 9 domains. When these 2 raters disagreed, the responses were read to a third rater. When either all 3 raters disagreed, or at least 1 rater thought a new domain was mentioned, those responses were categorized by consensus. Interobserver reliability between raters 1 and 2 was calculated by using Cohen's κ statistic.

Results: The 355 responses were categorized as follows: 303 (85.4%) identified one or more domains that were part of the previously developed taxonomy, 9 (2.5%) identified a new domain, 11 (3.1%) identified both old and new domains, and 32 (9.0%) could not be categorized. Cohen's κ was 0.57 ($P < .001$). Cultural sensitivity and physician honesty were the additional domains identified, by 1.1% and 4.5% of respondents, respectively.

Conclusions: The previously developed taxonomy of domains can be used in this setting to categorize the large majority of open-ended responses. Such responses can identify important aspects of care that were either previously unidentified or were already identified but given low ratings. This information then can help improve quality of care.

(*Am J Manag Care* 1999;5:1542-1549)

From the Department of Internal Medicine, Division of General Internal Medicine, Wayne State University School of Medicine, Detroit, MI (BDB, JLM, GER); and the Department of Medicine, Division of General Internal Medicine, Rhode Island Hospital, Brown University School of Medicine, Providence, RI (MJF).

This study was supported by the Department of Medicine, Rhode Island Hospital, Providence, RI.

Address correspondence to: Bruce D. Bialor, MD, Wayne State University Health Center, 4201 St. Antoine, Pod 5C, Detroit, MI 48201. E-mail: bbialor@intmed.wayne.edu.

Patient satisfaction has become an important outcome in ambulatory medicine. It is being used as a measure of the quality of care in many managed care settings,¹ and it has been shown to predict the likelihood of staying with the same physician.² Some managed care organizations use patient satisfaction ratings to provide financial incentives to physicians.³ Various patient satisfaction questionnaires (PSQs) have been developed; most use Likert-type scales with closed-ended questions, either about a specific medical encounter^{2,4-6} or more global aspects of outpatient medical care.^{7,8}

One advantage of using such a format is that it facilitates analysis of both individual items and the questionnaire as a whole (eg, with respect to reliability and validity measures). It also facilitates comparisons between or among settings, providers, and specialties. Its major disadvantage is that it is too restrictive from the patient's point of view; patients rarely get the opportunity to describe in an open-ended fashion what is most important to them when they see a physician.

Clearly, it would be desirable to give patients the opportunity to respond to open-ended questions, because that would enable providers to ascertain what is most important from the patient's perspective. These responses can provide unique information that may not be obtainable from a more conventional PSQ. An open-ended approach to measuring patient satisfaction also would create an opportunity to find out whether patients' needs are being met, and if not, how this can be improved. Investigators and managed care organizations may be hesitant to obtain such information, however, because analyzing these types of responses can be a complex task.

Few previous authors have attempted, in an open-ended fashion, to assess patients' opinions about what is important to them when they see a physician, with respect to either a specific medical encounter or more global aspects of outpatient medical care. One group asked 3 open-ended questions to 202 predominantly older male veterans after their clinic visits; the authors then used the responses to develop a taxonomy of components of patient satisfaction.⁹ Another group, in a community-based study of an older fee-for-service population, used open-ended methods (ie, focus groups) to help them develop an expansive list of elements of care thought to be important to high-quality office-based care. This list formed the basis for a questionnaire that asked study participants to rate the importance of these elements on a 4-point scale. In addition to using open-ended methods to develop their questionnaire, the authors defined a taxonomy of 9 domains of care and then classified each element into 1 of the 9 domains. Despite the expansiveness of both the questionnaire and the domains they defined, the authors acknowledged that certain elements may have been inadvertently neglected by not inviting open-ended comments from participants.⁸

Given the potential advantages of asking an open-ended question and the fact that such taxonomies exist,^{8,9} 2 questions arise. First, can such taxonomies be used to categorize these types of responses from patients in settings different from

those used by the previous researchers thereby making the process of analyzing the responses less complex and more feasible? One can make the empiric argument that most aspects of care that patients think are important will likely fall into one of the categories or domains in either taxonomy; responses that fit into these categories or domains would be unlikely to provide information not already obtainable from a more conventional PSQ. Second and more important, however, if the taxonomy does not capture all of the information, can open-ended responses identify important aspects of care that were previously not identified? These aspects of care may or may not be included in a more conventional PSQ.

We specifically asked patients an open-ended question as part of a patient satisfaction study we conducted in the internal medicine clinic of an urban teaching hospital. We used the taxonomy of 9 domains of care to categorize the responses to the open-ended question.⁸ Our first objective was to determine the proportion of responses that could be categorized into one or more of these domains. If this proportion was high, we thought it would encourage others to ask an open-ended question with their satisfaction surveys, because the responses that could not be categorized (and therefore might identify new areas of importance) would make up a relatively small proportion of the total. This would make the analysis of the responses more feasible. Our second objective was to determine from the responses that could not be categorized whether we could identify any new domains. Identification of new domains would suggest that open-ended responses provide information that a closed-ended questionnaire alone may not provide, thereby suggesting an important role for this type of question in satisfaction surveys. If a practical approach to analyzing responses to open-ended questions could be developed, it would be an important tool for both managed care organizations, in their ongoing quality improvement efforts, and investigators in this area.

... METHODS ...

Study Setting and Data Collection

The study was conducted over a 3-month period in 1994 in the internal medicine clinic of an urban teaching hospital. Patients who were scheduled to see a resident and had not completed the survey earlier in the study period were asked by a bilingual

research assistant to anonymously complete pre- and postvisit surveys, which were available in both English and Spanish. Verbal informed consent was obtained from the participants, and the hospital's institutional review board approved the study protocol. The previsit survey asked for demographic, socioeconomic, and health-related information, plus reasons for visiting and health-care utilization in the past year; this survey was similar to a previously used instrument.¹⁰ The postvisit survey included a modified version¹¹ of the American Board of Internal Medicine PSQ.⁴ It also asked for waiting time, visit length, and interpreter use. Lastly, the postvisit survey included an open-ended question: "What are the 1 or 2 things that are most important to you when you see a doctor?" Our analysis of the responses to this question forms the basis of this report.

Study Sample

During the study period, 2190 of 4663 scheduled patients (47%) showed up for their appointments. Residents saw 1240 of these patients. The form had been completed previously by 257 patients, and 102 spoke neither English nor Spanish, making 881 eligible. A total of 810 patients were approached (71 were not approached because the clinic was too busy). Two hundred fifty-nine of these patients were either unable (eg, they were hard of hearing, they were too sick) or refused to participate, 36 completed the previsit survey but either left out the resident's name or did not complete a PSQ, and 4 surveys were missing, leaving 511 completed PSQs (63% of the patients approached). Responses to the open-ended question were obtained from 355 patients (69% of the 511 who completed PSQs, or 43% of the 810 approached).

Data Analysis

We used the taxonomy shown in Table 1 to categorize the responses to the open-ended question.⁸ Methods similar to those used previously were used to categorize each response into a particular domain.¹² Respondents who indicated more than one domain in their responses were given credit for each domain they mentioned. Table 1 shows examples of responses given by patients in this study that were categorized into each of the domains.

We compared respondents to nonrespondents, using *t* tests for continuous variables, Pearson chi-squares for dichotomous variables, and Mann-Whitney *U* tests for ordinal categorical variables. The nonrespondents included the 36 patients who completed the previsit survey but either left out

Table 1. Previously Defined Domains of Care

Physician clinical skill	"That the doctor takes his time examining me and that he listens to what the problem is." "That I am examined and checked out very well."
Physician interpersonal skill	"They listen and take an interest in you." "Our relationship as doctor and patient."
Office support staff	"That everyone is pleasant and that they speak to me clearly."
Office environment	"There are always a lot of people in the waiting room."
Provision of health-related information	"Clear explanations." "Answer my questions, and explains everything I need to know."
Patient involvement in care and respect for patient values	"That they pay a lot of attention to what one has to say about one's own health." "He/she...respects me as a person."
Nonfinancial access	"Promptness, . . ." "Not having to wait a long time to be seen."
Finances	". . . I need my doctor to help me...with my insurance papers." "That they help me and give me Medicaid."
Coordination of care	"Test results." ". . . Ask other doctor when in doubt."

the resident's name or did not complete a PSQ, plus the 156 patients who completed both the previsit survey and the PSQ but did not answer the open-ended question.

... RESULTS ...

Patient Characteristics

The characteristics of the respondents are shown in Table 2. Those respondents who spoke neither English nor Spanish completed their surveys primarily through their interpreters. Some respondents spoke neither English nor Spanish, but were accompanied by an interpreter who completed the survey for the respondent. The respondents were more likely than the nonrespondents ($P < .05$) to report better health status, fewer reasons for visiting, longer visit lengths (29.4 vs 24.1 minutes), and higher socioeconomic status.

Categorization of Responses

The responses were read simultaneously to two of the coauthors (BDB and GER, whom we also will refer to as raters), who independently categorized each response into one or more of the domains in the taxonomy. They agreed on the categorizations of 189 of the 355 responses (53%). Cohen's κ statistic,¹³ the interrater reliability measure used when 2 raters are asked to rate nominal variables, was 0.57 (standard error = 0.02, $P < .001$). The remaining 166 responses were read to a third rater (MJF), who categorized the responses independently of the other 2; rater 3 agreed with rater 1 or 2 in another 65 cases. All raters had the option of categorizing a response into one or more domains that were not part of the taxonomy. When 2 of 3 raters agreed on the categorization of a response into a domain(s) that was part of the taxonomy, this was taken as the final domain(s) for that response. When all 3 raters disagreed with one another or at least one rater categorized a response into a domain that was not part of the taxonomy, the responses were categorized by group consensus. There were 101 (=166 - 65) responses in the former category and 44 in the latter.

As a result of this process, all of the responses could be classified into 1 of 4 categories. They were (a) reporting one or more domains that were part of the taxonomy; (b) reporting one or more domains that were not part of the taxonomy (ie, new domains); (c) reporting domains from both categories a and b; or (d)

not categorizable (ie, they were either too vague to categorize or suggested the respondent did not understand the question). Three hundred and three (85.4%) of the 355 responses were classified into category a, 9 (2.5%) into category b, and 11 (3.1%) into category c. Thirty-two (9.0%) were classified into category d; examples included "Everything," "My health," "OK," "Very good," "Excellent," and "All the doctors are very good."

Table 2. Characteristics of Respondents (n = 355)

Characteristic	Value
Age (n = 327)	
Mean \pm SD	47.0 \pm 14.7 y
Gender (n = 354)	
Male	125 (35.3%)
Female	229 (64.7%)
Socioeconomic status (SES) (n = 325)	
Low*	141 (43.4%)
Not low	184 (56.6%)
Self-rated health status (n = 352)	
Poor	72 (20.5%)
Fair	201 (57.1%)
Good	71 (20.2%)
Excellent	8 (2.3%)
Language (n = 355)	
English	236 (66.5%)
Spanish	110 (31.0%)
Seven other languages	9 (2.5%)
Country of origin (n = 344)	
Continental United States	207 (60.2%)
Dominican Republic	48 (14.0%)
Puerto Rico	51 (14.8%)
Nineteen other countries	38 (11.0%)
Number of reasons for visiting (n = 348)	
1	207 (59.5%)
2	96 (27.6%)
≥ 3	45 (12.9%)
Complexity (n = 334)	
Complex [†]	174 (52.1%)
Not complex	160 (47.9%)
Waiting time (n = 334)	
Mean \pm SD	40.1 \pm 31.1 min
Visit length (n = 309)	
Mean \pm SD	29.4 \pm 18.2 min

* Low SES = annual income less than \$10,000, education less than high school graduate, and not currently employed.

[†]Complex = more than 3 medications per day or more than 3 medical problems.

Table 3 shows the proportion of respondents who reported each domain, including new domains. The 3 most frequently mentioned domains were, in descending order, physician interpersonal skill, physician clinical skill, and provision of health-related information. At least 1 of these 3 domains was mentioned by 287 respondents (81%). We defined the 2 new domains as cultural sensitivity and physician honesty. Examples of responses that belonged to the former were "That Hispanics are treated better" and "Please try to hire Hispanic people to help your patients"; examples of responses that belonged to the latter were "To be honest" and "That they tell me the truth."

...DISCUSSION...

The purpose of this study, in which we asked an open-ended question as part of a patient satisfaction study, was twofold. Our first objective was to determine the proportion of responses to this question that could be categorized into 1 or more of the 9 pre-

viously defined domains of care in the taxonomy we used.⁸ We found that 85.4% of the responses (ie, the 303 responses) in category a could be categorized into one or more of these domains. Our second objective was to determine, from those responses that could not be entirely categorized into 1 or more of these 9 domains, whether we could identify any previously unidentified domains. Although not mentioned frequently, 2 such domains could be identified; they were cultural sensitivity and physician honesty.

Our first objective addressed the issue of whether using the taxonomy made it more feasible to analyze these open-ended responses. One way to use the taxonomy is to classify the responses more generally as being either in category a or "other" (ie, combining categories b, c, and d) and then focusing only on those that are classified into the "other" category, because they would form the basis for the new domains and therefore would require more detailed analysis. This use of the taxonomy acknowledges the fact that the responses in category a would be unlikely to provide information that is not already obtainable from a more conventional PSQ. Thus, the smaller the proportion of responses in the "other" category, the less complex the task of analyzing those responses. The fact that only 15% of our responses were classified into this latter category (and more than half of these were further classified into category d) suggests that using the taxonomy in this way can address some of the complexity issues. Even though identifying new domains is the most important reason to ask this type of question, it is important to acknowledge that making the determination of "new" or "not new" is likely to be the most time-consuming part of the analysis, even if a rater focuses on those responses that may describe new domains.

Therefore, an open-ended question can be used most practically in either of 2 ways. The first option is to ask such a question once, over a defined period of time, to identify domains that are unique to a particular setting; one can then revise the PSQ to include closed-ended items that address those areas. Cultural sensitivity is an example of a domain that we identified in this way. The other option is to ask for open-ended responses only if patients give low ratings on closed-ended items that are already in the PSQ (ie, they address domains that are not new); this would give them the opportunity to describe their reasons for such ratings.

Table 3. Number and Proportion of Respondents Who Reported Each Domain, Including Newly Defined Domains (n = 355)

Domain	No. (%) of Respondents*
Original	
Physician interpersonal skill	189 (53.2)
Physician clinical skill	90 (25.3)
Provision of health-related information	84 (23.7)
Nonfinancial access	33 (9.3)
Patient involvement in care and respect for patient values	18 (5.1)
Coordination of care	7 (2.0)
Finances	2 (0.6)
Office environment	1 (0.3)
Office support staff	1 (0.3)
Newly defined	
Cultural sensitivity	4 (1.1)
Physician honesty	16 (4.5)

* The sum of the percentages is greater than 100% because respondents were given credit for each domain they mentioned.

Physician honesty is an example of a domain in our setting that was identified in this way. This approach can be used in an ongoing fashion, whether or not one uses an open-ended question to identify unique domains. The other important advantage of this approach is that it does not require the same investment of time and effort as the first option.

The 2 domains we identified may be somewhat more specific to our setting. Approximately one third of the respondents were non-English-speaking, so it is not surprising that cultural issues are important for these patients when they see a physician. Conversely, such patients were not included in the studies in which the taxonomies were developed,^{8,9} so it is not surprising that this aspect of care was not mentioned by those respondents. Language and cultural barriers have been cited by many authors as having a negative impact on both the quality of care and access to care among patients with limited English proficiency.^{14,15} The open-ended question gave our patients the opportunity to raise their concerns about this aspect of their care; had we relied solely on a closed-ended questionnaire, this issue might have gone unnoticed. In addition to increasing a managed care organization's awareness of aspects of care that may be more specific to a certain setting, open-ended questions create an opportunity to revise a PSQ if a certain issue is raised frequently enough. The fact that relatively few respondents actually mentioned the issue of cultural sensitivity in their responses should not minimize its importance; this may have occurred because the PSQ that preceded the open-ended question did not mention this issue at all. Furthermore, the 4 (1.1%) respondents that mentioned cultural sensitivity actually represented 3.4% of the 119 non-English-speaking respondents, for whom this issue is most relevant.

There are many possible reasons why patients mentioned honesty as an important issue for them when they see a physician. Obviously this is an attribute that is desirable in all physicians; it is mentioned in our questionnaire,¹¹ and physician candor is included in the taxonomy developed in the veterans' study.⁷ The fact that some patients mentioned this attribute in their responses, however, suggests that they had concerns about whether or not their physicians were being honest with them. One can hypothesize that these concerns would arise more often in our type of setting than in a community-based setting, because patient-physician relationships in the former are not as longstanding as they are in the latter. Another reason is that these

patients, many of whom are medically underserved, may feel less trusting of physicians in general than other patient populations are.

The fact that physician honesty was mentioned at all in the open-ended responses emphasizes an important advantage of the open-ended approach. A closed-ended scale generates a level of satisfaction (or dissatisfaction) with a certain physician behavior or aspect of care. On the other hand, an open-ended response can provide a description of what caused dissatisfaction, as well as suggestions as to how a certain behavior or aspect of care can be improved, even one that is already known to be important. Because quality improvement is a high priority among providers, managed care organizations, and other healthcare delivery systems, this kind of information actually might be more helpful in bringing about improvements than the information available from a closed-ended questionnaire. Even though physician honesty was a previously identified and obviously important aspect of care, the open-ended responses provided information about that aspect of care in our setting that would not have been available from the PSQ alone. This information could then form the basis for interventions to improve the quality of care. Responses related to physician honesty could lead to additional training in interpersonal skills for the physicians involved, especially as it relates to building trust and breaking down barriers that result in mistrust. Similarly, the cultural issues raised could lead to the hiring of more interpreters, as well as education of providers and employees about cross-cultural aspects of care.

Several limitations need to be kept in mind when interpreting these results. The first, and perhaps most important, issue is the practical one already discussed. Obviously, significant costs in time and money are incurred when responses to an open-ended question are analyzed, which might make it prohibitive for a site to obtain this kind of information.

Another issue relates to the generalizability of our findings. These data were collected in a specific setting. Therefore, the potential usefulness of this taxonomy for the purpose of analyzing responses to an open-ended question should be assessed in other academic and community-based settings, which include diverse socioeconomic and ethnic groups, to determine the reproducibility of our findings.

One also needs to be aware of the potential for bias related to selection, nonresponse, and the distribution of responses. We have data from patients who came for scheduled appointments, agreed to

complete pre- and postvisit surveys, and answered the open-ended question. The participants who did not answer this question tended to report poorer health status, lower socioeconomic status, more reasons for visiting (ie, they had greater expectations for their visits), and shorter visit lengths. In an earlier study, we found these factors to be associated with lower levels of patient satisfaction.¹¹ The data might have looked different if we had been able to collect information from the nonrespondents as well as those who were otherwise excluded (eg, patients who refused, were too ill to participate, or were missed because the clinic was too busy). The distribution of responses in Table 3 shows that relatively few respondents mentioned domains other than the 3 most frequently mentioned ones. This potential bias in the responses may have resulted from having the open-ended question follow a PSQ that emphasized physicians' interpersonal skills. The distribution might have been different if the questionnaire had asked about other aspects of care, or if patients were asked the open-ended question before completing the questionnaire.⁷

A final limitation deals with the subjectivity of the method of analysis. This is an important issue to address in analyzing qualitative data; it is possible that these results could have turned out differently if the responses had been analyzed by other investigators. We tried to minimize this effect by using multiple raters and group consensus for all decision points in the analysis of the responses, but the issue of subjectivity remains a potential concern. Recently developed software packages that analyze qualitative data are available to analyze these types of responses.¹⁶ This software also can be used to determine the reproducibility of our findings.

Nonetheless, we can draw 2 main conclusions. First, the taxonomy of 9 domains can be used to categorize the large majority of responses to an open-ended question that asked patients for their opinions about what is important to them when they see a physician. Second, the small proportion of responses that cannot be categorized into the taxonomy can identify certain aspects of care that are important to patients. They may be previously unidentified aspects of care that are unique to a particular setting (eg, cultural sensitivity) or aspects of care already known to be important, but that were given low ratings (eg, physician honesty). These responses can help both providers and managed care organizations improve the quality of care they

provide. Therefore, we recommend that an open-ended question be used more widely in satisfaction surveys, whether it is being used for one or both of the above purposes. Further research is needed to address the reproducibility of these findings in other settings, to develop alternative methods of analyzing these responses, and to determine whether these domains add anything analytically to existing measures of patient satisfaction, thereby improving their accuracy.

Acknowledgments

This study was supported by the Department of Medicine, Rhode Island Hospital, Providence, RI. It was done at Rhode Island Hospital, Providence, RI, when Dr. Bialor was a fellow in the Division of General Internal Medicine.

The authors wish to thank Lori Rothenberg, PhD, and Elaine Hockman, PhD, for their assistance with the data analysis.

... REFERENCES ...

1. Epstein A. Sounding board: Performance reports on quality—prototypes, problems, and prospects. *N Engl J Med* 1995;333:57-61.
2. Rubin HR, Gandek B, Rogers WH, et al. Patients' ratings of outpatient visits in different practice settings: Results from the Medical Outcomes Study. *JAMA* 1993;270:835-840.
3. Doyle E. Can patient satisfaction surveys help diagnose (and improve) your practice? *ACP Observer* 1996;16(2):1,14,17.
4. Webster GD. Final report on the Patient Satisfaction Questionnaire project [executive summary]. Philadelphia, PA: American Board of Internal Medicine; 1989.
5. Linder-Pelz S, Struening EL. The multidimensionality of patient satisfaction with a clinic visit. *J Community Health* 1985;10:42-54.
6. DiMatteo MR, Hays R. The significance of patients' perceptions of physician conduct: A study of patient satisfaction in a family practice center. *J Community Health* 1980;6:18-34.
7. Hsieh M, Kagle JD. Understanding patient satisfaction and dissatisfaction with health care. *Health Soc Work* 1991;16:281-290.
8. Laine C, Davidoff F, Lewis CE, et al. Important elements of outpatient care: A comparison of patients' and physicians' opinions. *Ann Intern Med* 1996;125:640-645.
9. Concato J, Feinstein AR. Asking patients what they like: Overlooked attributes of patient satisfaction with primary care. *Am J Med* 1997;102:399-406.
10. Tamblyn R, Benaroya S, Snell L, et al. The feasibility and value of using patient satisfaction ratings to evaluate internal medicine residents. *J Gen Intern Med* 1994;9:146-152.

- 11.** Bialor BD, Gimotty PA, Poses RM, Fagan MJ. The effect of primary care training on patient satisfaction ratings in a medical residents' clinic. *J Gen Intern Med* 1997;12:776-780.
- 12.** Clements AD, Rothenberg L. Testing at higher taxonomic levels: Are we jeopardizing reliability by increasing the emphasis on complexity? *Res Schools* 1996;3:45-50.
- 13.** Wilcox RR. *New Statistical Procedures for the Social Sciences: Modern Solutions to Basic Problems*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1987:309-311.
- 14.** Woloshin S, Bickell NA, Schwartz LM, Gany F, Welch HG. Language barriers in medicine in the United States. *JAMA* 1995;273:724-728.
- 15.** Baker DW, Parker RM, Williams MV, Coates WC, Pitkin K. Use and effectiveness of interpreters in an emergency department. *JAMA* 1996;275:783-788.
- 16.** Hannibal M, Gahan C. *Doing Qualitative Research Using QSR NUD*1ST*. Thousand Oaks, CA: Sage Publications; 1998.