FREQUENCY OF PATIENT FEEDBACK TO PHYSICIANS

A GRADUATE MANAGEMENT PROJECT SUBMITTED TO
THE RESIDENCY COMMITTEE
IN CANDIDACY FOR THE DEGREE OF
MASTER OF HEALTH ADMINISTRATION

GRADUATE PROGRAM IN HEALTH CARE ADMINISTRATION

BY

CAPTAIN Wm. CHRISTOPHER CHAMBERS

WILLIAM BEAUMONT ARMY MEDICAL CENTER, EL PASO, TEXAS
JUNE 1994
FREQUENCY OF PATIENT FEEDBACK TO PHYSICIANS

CPT Wm. Christopher Chambers, MS

WILLIAM BEAUMONT ARMY MEDICAL CENTER
EL PASO, TEXAS

US ARMY MEDICAL DEPARTMENT CENTER AND SCHOOL
BLDG 2841 MCCS HRA US ARMY BAYLOR PGH IN HCA
3151 SCOTT ROAD
FORT SAM HOUSTON TEXAS 78234-6135

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................. ii
LIST OF ILLUSTRATIONS ........................................... iii
LIST OF TABLES ......................................................... iv

Chapter                                                                 | Page
1. INTRODUCTION .................................................. 1
   Conditions Which Prompted the Study .......................... 2
   Statement of the Problem or Question ....................... 5
   Literature Review ............................................. 6
   Purpose ......................................................... 10
2. METHODS AND PROCEDURES ....................................... 11
   Design .................................................................. 11
   Problem .................................................................. 14
   Solution .................................................................. 17
3. RESULTS .............................................................. 19
   Touch Screen Reliability ....................................... 19
   Descriptive Statistics ......................................... 20
   Relationship of Perception to Feedback .................... 23
   Direct Comparison of Feedback Frequencies .............. 26
4. DISCUSSION .......................................................... 28
5. CONCLUSIONS AND RECOMMENDATIONS .......................... 31

REFERENCE LIST ....................................................... 34

Appendix

1. Patient Perceptions of Humanism in Physicians
   Abbreviated 8 Item Questionnaire ............................ 39
2. Sample Feedback Form ........................................... 40
3. Exempt Protocol Statement .................................... 41
ACKNOWLEDGEMENTS

Express appreciation is necessary for the following individuals. Colonel Robert F. Bories, Jr. patiently guided and assisted in ways that made this project possible. I could never have completed this project without his support and belief. Captain Charlotte Nielsen was a friend and sounding board throughout the residency period and an experimental subject during the development of the computer program. More than words can convey I depend on Chong-ah Chambers, my wife, for her endurance of my quest and my daughter, Rachel Chambers, for her giving more brightness to my days. Finally, thank you to Lieutenant Colonel "Doc" Hawley-Bowland for her permission to conduct my study in her clinic with her physicians and patients.

The opinions expressed herein are strictly those of the author and do not reflect the official policy or position of the Department of the Army, the Department of Defense, or the U.S. Government.
LIST OF ILLUSTRATIONS

Figure                                                                 Page

Average Score by Feedback Group by Trial Period ...... 23
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Descriptive Statistics</td>
<td>21</td>
</tr>
<tr>
<td>2. Tests of Effects of Feedback Frequency to Physicians on Patient Perception of Physician Humanism</td>
<td>25</td>
</tr>
<tr>
<td>3. Tests of Effects of Feedback Frequency to Physicians on Patient Perception of Physician Humanism with Test Data Model</td>
<td>27</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION


The control of complex, higher-order, large-scale systems requires propinquity between feedback and processes (Lunze 1992). Health care is one of the most complex, higher-order, large-scale systems that exists. It is certainly a system in great social and economic turbulence. To competitively evolve the delivery of health care requires immediate information and feedback to enable flexible decision making at the lowest possible level: providers.
Conditions Which Prompted the Study

Contemporary health care is oriented towards the delivery of primary care medicine in an environment of ready access with appropriate cost. The customer's satisfaction is the key to the perception of quality and appropriateness of care that is delivered. If the customer perceives that the health care delivered in the primary care environment is high quality, then it is a prima facie certainty that the customer will be less inclined to seek second opinions, direct access to specialty care, or otherwise will circumvent the gatekeeping function inherent in a primary medicine, managed care environment. Therefore, the satisfaction of the customer is more important than at any other time in the delivery of health care.

Eligible beneficiaries will soon be able to elect to receive health care from our medical treatment facilities or newly contracted managed care sources in addition to traditional options. Our military hospitals will compete directly with other health plans to enroll our eligible beneficiaries. Virtually all health care delivery systems will manage care through the use of patient regulation in a primary medicine environment.

Outpatient measurements of customer satisfaction are less developed than the wealth of instruments developed to measure inpatient customer satisfaction. Customer surveys are typically administered yearly in military hospitals in compliance with military regulation and accreditation requirements. Military
hospitals do have feedback from patient affairs officers to the pertinent providers on an exception basis; review and analysis reports are routinely provided to the hospital headquarters. Still, there is a vast, untapped ability for today's information systems to supply a real-time link from the customer to the provider with valid and reliable measurements of the customers' satisfaction. Such real-time feedback to the providers of health care would enable the continuous improvement of the delivery of quality care as perceived by the patient. The cumulative reports would impel the primary medicine, managed care delivery team to satisfy, gatekeep, and retain those customers.

Few specific questionnaires exist that measure outpatient perception of physicians. One questionnaire measures perceptions of physician humanism (Appendix 1). It is not clouded with questions of patient parking, technology of the health care delivered, or quality of the receptionist (for example). While these are important questions, the physician has no control over these variables. Many physicians do not control drug formularies or provider networks. Physicians do control how they present themselves to the patients. Patient assessment of physician humanism correlates with patient satisfaction and compliance (Hauck et al. 1990). This is good information for feedback.

The frequency of feedback of customer perceptions may be as high or low a rate as would be beneficial to the health care providers. Ostensibly, a hypothetical rate of feedback that is
much too frequent could be ignored by physicians. At worst, too much feedback could produce overcorrection in provider behavior (stress) or even physician anger. Feedback that is too frequent may become noise and produce a turbulent system that is chaotically responsive. At the other extreme, the current yearly rote surveys with questionable feedback mechanisms are the lowest rate that regulatory standards permit. Insufficient feedback may produce a static, unresponsive system (Arthur 1987; Lunze 1992).

Interactive information systems are prolific. Corporate strategies abound to deliver interactive services via cable, telephone, or wireless transmission. Information kiosks are ubiquitous. Such systems usually consist of a touch screen computer terminal that is extremely simple and customer friendly. These systems include airport kiosks that access hotel systems, Social Security inquiry systems, bank teller machines, and even questionnaires. One company that I found offered an inpatient questionnaire on a touch screen system.

An information kiosk placed in a clinic to ask outpatients simple questions about their perceptions of physician humanism would create an ideal imbedded feedback system that could be given to the health care provider as frequently as warranted. This simple, modular information system could then be integrated later with other touch screen appointment systems, information systems, or various hospital information systems.
Statement of the Problem or Question

"Does customer perception of physician humanism change over time in relationship to the frequency of feedback of customer perception of physician humanism?" "What frequency of feedback is positively associated with customer perception of humanism?"

The creation of a patient questionnaire information kiosk would enable specific identification of that physician who delivered health care to the responding customer. Placement of the information kiosk in an outpatient area would permit greater response rates and feedback rates than manual survey techniques.

Feedback can be given to separate groups of physicians at different frequencies. Testing different frequencies of feedback would be done by comparing any changes in patient perceptions of physician humanism of different groups over time. The group of physicians that has the most improvement (i.e., statistically significant) in patient perceptions of physician humanism would indicate a more effective feedback rate for that population.

Formulation of the definition of humanism was based on the values that are established in relationships between people such as dignity, justice, and respect. With physicians, this implies a "sensitive, non-humiliating, and empathetic way of helping [a patient] deal with some problem or need" and with a "double obligation to protect the values of the patient and to be faithful to his (or her) own" (Hauck et al. 1990).
Literature Review

The selection of an outpatient survey to implement automated feedback via an information kiosk was the primary focus of the initial literature review. The selection criteria of the instrument that measures patient perceptions of humanism in physicians are detailed in the Methods and Procedures Chapter, Problem Section that follows. Evidence indicates that extensive debate and credibility issues surround patient evaluation of care quality (Rubin 1990). Due to this extensive debate, a patient's evaluation of physician humanism seemed much more appropriate.

Patient Surveys

Practical problems abound in the conduct of patient surveys to include patient access, patient selection, housestaff selection (which physicians to evaluate), selection of survey instruments, confidentiality, patient characteristics, and procedural issues (Parker and Kroboth 1991). To keep these complexities out of the test of frequency of feedback, the least complicated questionnaire again seemed most appropriate. Absolutely no identifying patient information should be collected. This avoids lengthy consent procedures that would obviously defeat the purpose of an information kiosk setting.

Inpatient short form surveys are more developed. The most developed Patient Comment Card (PCC) is based on the extensively used Davies and Ware Patient Judgment System (PJS). The validity and reliability data for this instrument are very sophisticated
(Nelson et al. 1991). Regretfully, this inpatient questionnaire is not amenable to feedback of physician-centered inquiries.

An outpatient questionnaire of satisfaction was developed for emergency room patients (Bursch, Beezy and Shaw 1993). The extensive instrument required follow-up interviews (post emergency) and supported the perception of caring as a determinant of satisfaction. Clearly, patient satisfaction surveys are a broadly recognized opportunity for total quality improvement (Nelson and Niederberger 1990). Currently, the major emphasis in the literature is on inpatient surveys.

Information Kiosks

Touch screen information systems have been eagerly accessed when made available to the health care consumer. Healthpoint is a community-based touch screen health information system that makes health information available to consumers in various public areas. Population samples indicate that 25% of consumers initially exposed to Healthpoint will use the system (Jones, Navin and Murray 1993).

Feedback to Physicians

Recently, researchers have provided feedback to physicians in efforts to control costs (Curtis et al. 1993). The literature is conclusive that physicians will respond to specific feedback by altering practice patterns. One study provided feedback at specific time intervals of one day and seven day time packets
(Studnicki et al. 1993). Regrettably, this feedback was provided simultaneously without evaluation of preferred rate. Another physician feedback study allowed physicians to compare their individual performance against peers. Physicians responded positively to this information (Buntinx et al. 1992).

The best comparative research on the optimal frequency of feedback examines health care workers' (nursing aides) response to continuous, intermittent or no feedback schedules. The study demonstrated that increased feedback density (frequency) produced correlated increases in the rate of the subjects' acquisition of basic skills (Alavosius and Sulzer-Azaroff 1990). This exemplar work in health care is fundamental. The results are consistent with virtually all established systems theory and sensory-motor skill acquisition theory (Berrien 1968; Royce and Powell 1983). Key differences exist between the cited research and the research proposal. Physicians perform exceedingly more complex tasks (to include humanistic patient consultation). Physician status is perceived to be past the acquisition of basic skills. Feedback mechanisms did not directly link customers with providers in the cited research and did not institute a cybernetic model (feedback was temporary until basic skills were acquired; the "learning" trials were then complete). Most importantly, the control group with no feedback was not obtrusively measured in their workplace.

In health professions education and training, self assessment is compared to external measurements for feedback
(Gordon 1991) and clinical outcomes are widely used for provider feedback (Borbas et al. 1990; Magoffin 1990; Mugford, Banfield and O'Hanlon 1991). Feedback is used to educate providers on laboratory test request appropriateness and drug interactions effects (Gama et al. 1992; Kroenke and Pinholt 1990). Medical research results are compared with physician practice patterns to gauge the effectiveness of journal publications (McCormick 1990). Simulated case feedback is used to improve provider diagnostic ability (Wigton et al. 1990). Resuscitation training requires feedback to improve basic skills to augment experience (Marteau et al. 1990). The use of feedback in education and training is a natural sequitur to the clinician use of feedback in patient behavior modification as a treatment (Pine and Jacobs 1991).

In delivery of health care, practice guidelines and feedback are used to alter lengths of stay in intensive care units (Eagle et al. 1990). Patient satisfaction and outcomes are monitored in ambulatory surgery settings (Fallow 1991). Computer generated feedback is used in health maintenance organizations to provide preventive medicine vaccination reminders and peer comparisons to physicians (Barton and Schoenbaum 1990). Another preventive medicine application includes the use of community feedback to improve health surveys (Clarke et al. 1990).

Generally, patient expectations are measured and managed through feedback (Kuykendall 1992). In the health care industry, customer feedback is used to market services (McKenna 1991), to
define quality (Miller 1992), and to eliminate defects (Reichheld and Sasser 1990). Feedback is a key management tool between health care supervisors and their organization (MacStravic 1990).

Throughout the literature, feedback is a traditional tool of physicians. Physician education, training, and practice patterns are modified by feedback. Current quality, service and business initiatives are replete with the implementation of feedback in conformity with modern systems theory and management practice.

**Purpose**

The Joint Commission on Accreditation of Healthcare Organizations requires that measurements be in place for patient satisfaction and continuous quality improvement. An information kiosk questionnaire with tested feedback effects and outcomes transcends that requirement. An information kiosk system will enable the most rapid interface between customers and providers in an objectively measured and flexibly queried cybernetic feedback system. The managed care component for beneficiary enrollment will be supplied strategic, real-time information. Physicians will be impelled to present their services to patients in a manner perceived as humanistic by patients. Physicians will be enabled to perform their gatekeeping role while providing appropriate care in a primary medicine environment.
CHAPTER 2

METHODS AND PROCEDURES

The design, problem and solution of the research project is addressed in this chapter to include: the events measured (who); the variables, design, and data types (what); the type of analysis (how); and the schedule of procedures (when and where).

Design

This project was a quantitative project. The design was a modified Pretest-Posttest Control Group Design (Campbell & Stanley 1963). I initially considered the use of a control group that received zero frequency feedback. That was when I thought there were about 30 physicians available in a primary care clinic at William Beaumont Army Medical Center. The number of primary care physicians was stated in Full Time Equivalents (FTEs) that were contracted. In effect, any number of physicians may arrive under contract to fulfill the requirement. These contract FTEs were not statistically amenable to my investigation (no unit equivalence). A sufficient number of physicians were available for my study (24) in the Obstetrics and Gynecology Clinic. This clinic allowed primary access to care for appropriate patient needs as well as specialty care patients.
The patients' perceptions of humanism in their physicians were measured in a direct access clinic. Each physician was provided summary reports of patient perceptions of that physician's humanism. The physicians were also provided summary information of the other physicians' statistics so each physician could compare their own patients' perceptions with those of the clinic overall. The reports identified physicians by code. This permitted the physicians to anonymously compare their feedback within their clinic or to openly compare their data.

The physicians were randomly assigned to two groups. One group received weekly feedback. The other group received daily feedback. The daily feedback group was only able to compare daily reports with other physicians in the daily feedback group. All physicians received weekly feedback and comparisons were possible among all physicians in the clinic (both daily and weekly feedback groups). Given successful random assignment and sufficient sample size, the first measure of patient perceptions was expected to be the same between groups (pretest). During and after the experimental trials, comparisons of the two groups were made again to detect differences between each group's change in their patient perceptions (test and posttest).

The null and alternate statements of the hypotheses follow:

**Hypothesis 1**

1. **H0**: Continued survey respondent scores of patient perceptions of physician humanism would not significantly vary in relationship to feedback rates of those responses to physicians.
**Hypothesis 1_{alt}** = Continued survey respondent scores of patient perceptions of physician humanism would significantly vary in relationship to feedback rates of those responses to physicians.

**Hypothesis 2_{null}** = The Daily Feedback Group versus Weekly Feedback Group respondent scores would not be significantly different when provided different feedback rates of continued survey responses.

**Hypothesis 2_{alt}** = The Daily Feedback Group versus Weekly Feedback Group respondent scores would be significantly different when provided different feedback rates of continued survey responses.

The independent variable was the frequency of feedback given to each group (daily or weekly). The dependent variable was the measurement of patient perceptions before, during and after the feedback frequency treatment trials. Feedback was operationally defined in this study as the reporting of patient measurement of their perceptions of a physician's humanism to that physician. Specific feedback was accompanied by group data for comparison. Feedback frequency was operationally defined as how often feedback was provided over a block of time; daily feedback was provided four to five times (working weekdays) per week and weekly feedback was provided one time per week. Patient perception of physician humanism was operationally defined as the characteristic measured by the The Humanism Scale Questionnaire (Hauck et al. 1990) by information kiosk survey of outpatients.
Problem

The original problem statements, "Does customer perception of physician humanism change over time in relationship to the frequency of feedback of customer perception of physician humanism?" and "What frequency of feedback is positively associated with customer perception of humanism?" were resolved by the study's simple design which tested these questions.

Validity of the study was primarily derived from clear face validity of the patient questionnaire. The questions were blunt and specific about patient impressions of their physicians (Appendix 1). The amassed weight of patients expressing their opinion through automatic feedback was given to physicians.

The scale was reviewed by an expert panel of physicians and social scientists that determined content validity for each item as an appropriate measure of humanism (Hauck et al. 1990).

Humanism explained 60.5% of the variance in patient satisfaction with physician-related aspects of care after removing effects of Insurance, Support, and Race ($R^2 = .675$, $R^2\Delta = .605$, $F = 286.24, p < .001, N = 185$). Physician humanism even explained 25.3% of the variance in patient satisfaction with the non-physician aspects of care after removing effects of Occupation, Support, Race, Insurance and Employment ($R^2 = .300$, $R^2\Delta = .253$, $F = 46.00, p < .001, N = 185$) (Hauck et al. 1990).
The test instrument obtained a reliability coefficient (Cronbach's alpha) of .95. The abbreviated eight item scale had a reliability coefficient of .93. Validity was determined by measuring patient satisfaction (6 item scale) and using hierarchical multiple regression analysis to measure relationships with humanism scores (Hauck et al. 1990).

There were many possible motivations of physician responses to patient feedback. Each physician may have responded to the desire to deliver good quality care, healthy competition with oneself and one's peers, increased self awareness, the knowledge of institutional review, or the desire to impress the researcher (seeminal works are legion which cite the perennial attention awareness response, a.k.a. the "Hawthorne Effect").

It was beyond the scope of this study to determine physician motivation for several reasons. The complexity of personality and individual differences exceeded this study's scope. Feedback may have animated myriad motivations and caused responses; the goal of the feedback was precisely to trigger such responses.

The only way to have controlled for the effects of feedback was to have a control group of physicians. Their patients would be surveyed, but control group physicians would have been deprived of the feedback. That physician control group would be eminently aware that an information kiosk was measuring their patient's opinions. If feedback were withheld from one group while given to other groups there could have been a potentially
disastrous Hawthorne Effect. The other reason to not have had a control group was the limited number of physicians.

But the raison d'être to develop an embedded feedback system was the immanent capacity of cybernetic feedback in systems to perpetually monitor, guide and control any process (e.g., physician delivery of health care to patients). Control of a process, given continual feedback, is valid regardless of the motivation. In short, if a Hawthorne Effect or any motivator improved a patient's perception of a physician's humanism, such an effect or motivator would be a continually elicited response via the imbedded feedback system. This study was designed to determine which rate of feedback is better, daily or weekly, and then keep up that rate of feedback.

Also, I was not replicating the plethora of works that demonstrate that the presence of feedback is significantly different than not having feedback. The fundamental question was to determine if different feedback rates have different results.

Finally, I was not designing a questionnaire. For that reason I have reviewed the literature broadly. While I was surprised at the small number of outpatient questionnaires, I did discover an instrument of sound experimental design with complete information of reliability and validity. The short form of the survey was investigated after reduction by factor analysis and functioned as an alternate form in the investigation of their research validly (Hauck et al. 1990).
Solution

I wrote a computer program to present survey respondents with a "Likert-Type" choice on a touch screen computer monitor. I used a graphic user interface database (Borland International, Inc. 1992a, 1992b, 1992c, 1992d; Biow 1993) for both the questionnaire and feedback reports. The development of an eight question database was very simple. The development of an interface was very difficult. Complex control was available over the graphical user interface. The program was made sufficiently rugged to leave it in the clinic unattended.

The feedback report to the physicians required the same careful development. To prevent an obstacle over the quality of the feedback report format, I developed this report with great care towards the data being both sufficient for intelligent interest and in summary format (graphic) for busy physicians. Within the capability of the software, a Patient-Physician Feedback Summary was developed with physician assistance (including physician executives), administrators of clinics, and the program preceptor (Appendix 3). The integrated program was completed and installed after the arrival of a touch screen unit as Government Furnished Equipment. The information kiosk was in place by April and collected data for two months.

Limited time for the study prohibited evaluation of monthly feedback. Weekly and daily reports are common and were the next most practical feedback frequencies to investigate. The
selection of the frequencies of feedback was dependent on the time allotted.

This study was considered an Exempt Protocol in accordance with Army Regulation 40-38. Exempted Protocol status was due to the following categories of "no more than minimal risk" and survey responses were "recorded in such a way that subjects cannot be identified directly or indirectly." There was no requirement for submission to the Institutional Review Board (Appendix 3).
CHAPTER 3
RESULTS

The eight questionnaire items were assigned a score of 1 for "Strongly Disagree" through 5 "Strongly Agree" for "Likert-Type" responses by the subjects. The possible total score ranged from 8 to 40. Actual humanism scores throughout the pretest, test, and posttest periods ranged from 8 through 40 with 88 responses.

Random assignment of the physicians to the daily or the weekly feedback groups was verified by Student's t-test for independent samples ($t = -.07, \text{df} = 20, p = .94$) at the end of the pretest period. The results were not significant and the 24 physicians (and 1 nurse practitioner) remained in the initially assigned groups throughout the test and posttest periods.

Touch Screen Reliability

The responses from all test periods were analyzed to compare the touch screen presentation of the questionnaire with the source questionnaire's data analysis of reliability (Hauck et al. 1990). The reliability coefficient (Cronbach's alpha) of the 8 humanism items was .94; compared with .93 for the original questionnaire. Average inter-item correlation was .71 for the touch screen media presentation of the abbreviated 8 item questionnaire.
Descriptive Statistics

As previously stated, there were 88 responses over 34 days of data collection. There was an average of 2.59 responses per day. There was no possible ability to reject either the first or the second null hypotheses. Feedback to the 12 health care providers in the daily group omitted individual feedback to 11 providers on an average daily basis. Feedback to the 13 health care providers in the weekly group omitted individual feedback to 8 providers on an average weekly basis. Hypothesis $H_{0,1}$ is accepted; continued survey respondent scores of patient perceptions of physician humanism did not significantly vary in relationship to feedback rates of those responses to physicians. Hypothesis $H_{0,2}$ is accepted; the Daily Feedback Group versus Weekly Feedback Group respondent scores were not significantly different when provided different feedback rates of continued survey responses. The basis of accepting the first and second null hypotheses was the insignificant respondent rate. The density of respondent rate in ratio to the physician population was insufficient for significant individual feedback. The survey respondents represented an insufficient sample of the patient population to comment on the clinic's overall patient population.

An analysis of the descriptive statistics is included to provide statistical models and permit comparison with replication of this research. It would be of interest to compare the scores of increased response rates with the results of this research.
A 95% confidence interval of the mean of the responses (31.67) was calculated (29.97 to 33.25). Data that were significantly outside the mean (1.96 * σ) were present. Those data were analyzed by comparing a standard residual plot of the outliers (> 2 sigma from the regression line in the following hypothesis testing model). This analysis revealed 7 extreme responses that would significantly influence the model given the extremely small number of responses. Those responses were discarded and not included in the descriptive statistics. These discarded responses were not from the pretest period. There were 81 responses included in the analysis with a mean of 33.30 and a range of scores from 20 through 40. The standard deviation for all responses was 5.16 and there was a decrease in the range of the 95% confidence interval (32.17 – 34.45). The descriptive statistics for each of the periods and groups are in Table 1.

<table>
<thead>
<tr>
<th>Feedback Group</th>
<th>Pretest Mean</th>
<th>Std Dev</th>
<th>Test Mean</th>
<th>Std Dev</th>
<th>Posttest Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>33.57</td>
<td>4.72</td>
<td>35.40</td>
<td>4.26</td>
<td>35.75</td>
<td>2.96</td>
</tr>
<tr>
<td>Weekly</td>
<td>33.75</td>
<td>6.18</td>
<td>31.67</td>
<td>5.77</td>
<td>33.25</td>
<td>3.01</td>
</tr>
<tr>
<td>No ID</td>
<td>30.00</td>
<td>6.68</td>
<td>28.80</td>
<td>6.65</td>
<td>27.00</td>
<td>5.66</td>
</tr>
<tr>
<td>Aggregate</td>
<td>33.07</td>
<td>5.42</td>
<td>33.29</td>
<td>5.55</td>
<td>33.67</td>
<td>4.07</td>
</tr>
<tr>
<td>Grand Total</td>
<td>33.30</td>
<td>5.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 1
DESCRIPTIVE STATISTICS

21
The "n" reported in Table 1 indicates the number of patient responses recorded for a specific health care provider feedback group in a particular period of the experiment. For example, there were 12 patient responses during the feedback test period provided the physicians assigned to the weekly feedback group. As a further example, during the posttest period, the 8 patient responses identifying physicians assigned to the daily feedback group had an average of 35.75 and a standard deviation of 2.96.

The 8 items of the abbreviated scale were averaged with scores 1 through 5 to better represent the data graphically in the following figure. The scale of the Y axis was truncated to indicate values of 3 through 5 to represent the concentration of the range of values.

Replications of this research should expect to see average scores with no significant differences in the pre-feedback (pretest) period. Graphic differences in the feedback test period would be apparent and examined by regression analysis as applied to the data in the following two sections to perform statistical hypothesis testing. In the present data set, there are clear graphic differences between the group average of the daily feedback group and the weekly feedback group during the feedback test and posttest periods. These differences have no significant relationship to the hypothesis statements because of the insufficient response rate sample size or feedback density. Discussion of significance is for replication comparisons only.
Physician Humanism Scale
Average Score by Feedback Group by Trial Period

Figure. Average Score by Feedback Group by Trial Period

Relationship of Perception to Feedback

The variance apparent in the figure (above) was analyzed to provide a model for testing the hypotheses in this research. This analysis was performed to provide a replication guideline and to compare replications with the present results.

A hierarchical multiple regression analysis model was used to analyze the first hypothesis to test whether the respondents' perception of physician humanism changed over time in relation to the frequency of feedback to physicians. There was 17% of the scores' variance explained ($R^2 = .17$, $R^2$ adjusted = .11). Each
of the independent variables were separately tested against the
full regression model of all independent variables to determine
the unique contribution of each independent variable to that 17%
variance. The first hypothesis test of frequency of feedback
results are in Table 2.

The overall model to analyze each separate effect of
feedback during each period was \( R^2_{\text{Full}} = c_{\text{Intercept}} + b_{\text{Pretest Daily}}(X_1) +
\)
\( b_{\text{Pretest Weekly}}(X_2) + b_{\text{Test Daily}}(X_3) + b_{\text{Test Weekly}}(X_4) + b_{\text{Posttest Daily}}(X_5) +
\)
\( b_{\text{Posttest}}(X_6) \). The effects of respondents' scores who did not
identify physicians were retained in the model. Those responses
were dichotomously coded throughout the model with the period
identified. Daily and weekly feedback variables both indicated
no feedback (null condition) when no physician was identified.
Additionally, the variable of "no physician identification"
(i.e., no feedback) prevented having an ill-conditioned matrix
while retaining each of the primary conditions explicitly stated
in the model. There was no need for dummy coded variables.

The individual effect of the period of each group was tested
by iteratively removing the effect of each group from the model
and restricting the contribution of that effect from the full
model. To test for the effect of the daily feedback group during
the pretest period the regression model was \( R^2_{\text{Restricted}} = c_{\text{Intercept}} +
\)
\( b_{\text{Pretest Weekly}}(X_2) + b_{\text{Test Daily}}(X_3) + b_{\text{Test Weekly}}(X_4) + b_{\text{Posttest Daily}}(X_5) +
\)
\( b_{\text{Posttest}}(X_6) \). The difference between the \( R^2_{\text{Full}} \) and the \( R^2_{\text{Restricted}} \) then
indicated the unique variance contributed by that restricted
variable (Pretest Daily) to the full regression model. The F ratio was then computed for each difference value and tested for significance.

The collective contribution of each period of all feedback conditions was tested by restricting each period's variance. Pretest variance was restricted by the model \( R^2_{\text{Restricted}} = c_{\text{Intercept}} + b_{\text{Test Daily}}(X_3) + b_{\text{Test Weekly}}(X_4) + b_{\text{Posttest Daily}}(X_5) + b_{\text{Posttest}}(X_6) \) and the difference from the full regression model tested.

### TABLE 2

TESTS OF EFFECTS OF FEEDBACK FREQUENCY TO PHYSICIANS ON PATIENT PERCEPTION OF PHYSICIAN HUMANISM (N = 81)

<table>
<thead>
<tr>
<th>Variable</th>
<th>( R^2 ) Full</th>
<th>( R^2 ) Reduced</th>
<th>Unique Variance</th>
<th>df1</th>
<th>df2</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest Daily</td>
<td>.1749</td>
<td>.1022</td>
<td>.0727</td>
<td>4</td>
<td>74</td>
<td>3.30</td>
<td>.0152</td>
</tr>
<tr>
<td>Daily Pretest</td>
<td>.1749</td>
<td>.1172</td>
<td>.0577</td>
<td>5</td>
<td>74</td>
<td>5.25</td>
<td>.0004</td>
</tr>
<tr>
<td>Weekly Pretest</td>
<td>.1749</td>
<td>.1292</td>
<td>.0457</td>
<td>5</td>
<td>74</td>
<td>4.15</td>
<td>.0022</td>
</tr>
<tr>
<td>Test Pretest</td>
<td>.1749</td>
<td>.0311</td>
<td>.1438</td>
<td>4</td>
<td>74</td>
<td>6.54</td>
<td>.0001</td>
</tr>
<tr>
<td>Daily Posttest</td>
<td>.1749</td>
<td>.0316</td>
<td>.1433</td>
<td>5</td>
<td>74</td>
<td>13.03</td>
<td>.0000</td>
</tr>
<tr>
<td>Weekly Posttest</td>
<td>.1749</td>
<td>.1592</td>
<td>.0157</td>
<td>5</td>
<td>74</td>
<td>1.43</td>
<td>.2270</td>
</tr>
<tr>
<td>Posttest Daily</td>
<td>.1749</td>
<td>.0714</td>
<td>.1038</td>
<td>4</td>
<td>74</td>
<td>4.72</td>
<td>.0019</td>
</tr>
<tr>
<td>Posttest Weekly</td>
<td>.1749</td>
<td>.0780</td>
<td>.0969</td>
<td>5</td>
<td>74</td>
<td>8.81</td>
<td>.0000</td>
</tr>
<tr>
<td>Posttest Weekly</td>
<td>.1749</td>
<td>.1391</td>
<td>.0358</td>
<td>5</td>
<td>74</td>
<td>3.25</td>
<td>.0104</td>
</tr>
</tbody>
</table>

The hypothesis that respondents' perception of physician humanism changed over time in relation to the frequency of feedback to physicians was tested. The pretest period was found to contribute significant variance to the full regression model and indicated that pretest condition uniquely varied from the
remaining explained variance in the model of the respondents' perception of physician humanism. The posttest period also contributed significant variance to the full regression model and indicated that pretest condition varies from the remaining explained variance in the model. The test period contributed significant variance; the weekly feedback test period did not contribute significant variance to the explained variance in patient perceptions of physician humanism.

Hypothesis 1 was tested. The respondents' scores changed over time in relationship with daily feedback to physicians. The respondents' scores did not change over time in relationship with weekly feedback to physicians. Overall, Hypothesis 1_{null} prevails over Hypothesis 1_{alt} due to the insignificant total response rate.

Direct Comparison of Feedback Frequencies

Again, the test of the second hypothesis was conducted only to provide a hypotheses testing model for this research. This analysis was performed to provide a guideline for replication and to compare results of replication with the present research. To compare daily to weekly feedback, the data from periods with no feedback (the pretest and posttest periods) were deleted. Daily and weekly feedback rates were the primary independent variables of interest. The respondents' scores during the pretest and the posttest periods were used to test the first hypothesis. After
the test of the first hypothesis, periods without feedback were not consequential to the testing of the second hypothesis.

The full regression model to analyze the effects of daily and weekly feedback was \( R^2_{full} = C_{intercept} + b_{Test\ Daily}(X_1) + b_{Test\ Weekly}(X_2) \).

The effects of respondents' scores who did not identify physicians during the test period were again retained in the model. The effects of daily feedback was tested by the restricted model of \( R^2_{Restricted} = C_{intercept} + b_{Test\ Daily}(X_1) \). The model for testing weekly feedback was \( R^2_{Restricted} = C_{intercept} + b_{Test\ Weekly}(X_2) \).

The full model produced a \( R^2 \) of .20 (adjusted \( R^2 = .15 \)).

\[ \begin{array}{ccccccccc}
\hline
\text{Variable} & R^2 \text{ Full} & R^2 \text{ Reduced} & \text{Unique Variance} & df1 & df2 & F & p \\
\hline
\text{Daily} & .1996 & .0426 & .1570 & 1 & 34 & 6.87 & .0130 \\
\text{Weekly} & .1996 & .1734 & .0262 & 1 & 34 & 1.15 & .2911 \\
\hline
\end{array} \]

Respondent scores were significantly different between the daily and weekly feedback groups \( (t = 2.10, df = 30, p = .04) \).

Daily respondent scores contributed to the explained variance and varied in a positive direction. \textbf{Hypothesis 2}_{\text{null}} still prevails over \textbf{Hypothesis 2}_{\text{alt}} due to the insignificant total response rate.

27
CHAPTER 4

DISCUSSION

There were no statistically significant overall results. The paucity of the response rate eliminated statements about the other patients of the same physicians. The analysis was limited to those responses measured by a touch screen computer survey. The respondent sample size was not significant. Furthermore, there were not enough responses to regularly provide each physician with respondent scores at most intervals of feedback.

An inference that physicians would change their behavior toward patients as a result of their rate of feedback was dominant in both hypotheses. This inference was only valid if there was a significant change in respondents' perception of physician humanism explained by the feedback frequency. The low number of responses has made the alternate hypotheses untestable.

The data model analysis suggests that the physicians in the daily feedback group may have changed their behavior towards the respondents. Physicians could not selectively change their behavior only toward patients that responded to a touch screen computer survey. The physicians did not know which patients would respond to the survey. Replication of this study can test if the results of the analysis are significant or consistent.
The time limitations on this research prohibited longer test periods. The procurement time required for government furnished equipment limited this research to 7 days of pretest data, 20 days of feedback test data, and 7 days of posttest data. Time limitations limited comparisons across clinics. Selection of the clinic was based on a large, stable physician population in a clinic that treated primary care patients. An additional criterion was to select a clinic without confounding variables of access or other issues that would influence the hypothesis tests. Therefore, a clinic with a stable population with good access had a lower patient complaint rate in the hospital. This possibly limited the patient response rate. Certainly, the limited time for the experiment would have been negatively influenced by a problematic clinic with the confounding variables overwhelming the variables of interest even if there were more responses.

The test instrument was initially validated with the same number of physicians with a remarkably similar Cronbach's alpha. There is no ability to compare response rates between the original instrument (a stratified sample mailed out with a 45% response rate). The selected clinic for the touch screen survey accessed patients through a walk-in waiting room and a waiting room for patients with appointments. Patients with appointments often had no waiting time and went directly to the physician's office or a hallway waiting area directly outside the office. Patients often left through traffic areas separate from the
appointment waiting room. The patients frequently had follow-on medical requirements for laboratory work, prescription requests, or specialty consultation visits. Many patients had children to pick up from day care centers or schools after appointments. Spot observations of the patients in the waiting room and exiting by that waiting room still indicated low response rates.

The original questionnaire research indicated that 60% of measured patient satisfaction is uniquely explained by patient perception of physician humanism. Daily feedback to physicians of respondents' perception of physician humanism accounted for 15% unique variance in respondents' perceptions of physician humanism. There is potential to improve patient satisfaction with physician aspects of care given this commonality. Patient perception of physician humanism also explained 25% of the variance in patient satisfaction with nonphysician aspects of care in the same original study (Hauck et al. 1990).

Finally, this research was limited due to the type of clinic. The clinic is for female patients and no inference may be drawn to the male population. The clinic is arguably a primary care clinic. Obstetric and gynecology residency education programs have resisted proliferation of subspecialties and have emphasized preventive and primary care (Pearse 1993; Herold et al. 1993). Multidisciplinary trends exist to blend obstetrics and gynecology with primary and ambulatory care (Johnson and Dawson 1990).
CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

This was exploratory research. The body of introductory theory and the literature review indicate that more feedback is better, but logically too much feedback is possible (every minute, hour, etc.). The data model analysis suggests that daily feedback from the respondents to their physicians was not harmful and was of benefit (albeit statistically insignificant overall).

The utility of this study was to offer physicians feedback. Replication of this research is warranted to further compare the feedback frequencies in varieties of clinics with other valid patient measurements. Systemic benefits must be measured after full implementation of the feedback system over an extended period of time. Network implementation of the database that results from this measurement of patient satisfaction allows comparison with a vast array of internal and external factors.

Direct feedback allows for physician autonomy and permits self correction without recourse to the patient representative to file complaints and even litigious choices by unhappy patients. Institutional needs for control, quality improvement and other macro level analyses are met by the ability to integrate and analyze information continually measuring processes and outcomes.
The reliability of the survey instrument administered by a touch screen was equal to the written instrument. Responses to the feedback were consistent with the literature. Technology can enable cybernetic application of diverse measurement instruments throughout the healthcare system.

The low response rate indicates the need for traditional survey methods. Supplementary application of feedback systems allows for increased timeliness of information and correction. The density of feedback directed at particular physicians was correspondingly low with the low response rate (12 physicians received 20 days feedback with 20 identified and 5 unidentified responses). The tentative positive increase in measurements of patient perception of physician humanism was remarkable even if not significant. Constant feedback of information considered as valid by the physicians may be effective even when it is not specifically identified. The awareness of the measurement process is triggered with each feedback iteration and the response (the perennial Hawthorne effect) can be elicited over time. Cybernetic feedback systems imbedded in the health care environment are potentially effective.

The prime purpose of this research was to test a system that could improve patient perceptions of the quality of health care in an outpatient setting. Since the research project, primary care physicians and key physician executives have identified other key locations to place this system. The local staff
physician demand for this valid, timely, and crucial information indicates the value physicians perceive in patient feedback.

Most importantly, evaluation of performance based on external customer feedback will reduce the singular emphasis on management driven feedback (Nevling 1992).


APPENDIX 1

Patient Perceptions of Humanism in Physicians
Abbreviated 8 Item Questionnaire (Hauck et al. 1990):

Instructions: After reading each statement, please select the response that most correctly describes your response to the statement.

1) My doctor seems to take a personal interest in me.
2) Even when my problem is small, my doctor is concerned.
3) I have confidence in my doctor's decisions.
4) My doctor respects my beliefs.
5) I would talk to my doctor if something were troubling me.
6) My doctor takes an interest in my home life.
7) My doctor is easy to talk to.
8) My doctor seems to know what I am going through when I tell him/her about a problem

Note: Reliability coefficient (Cronbach's alpha) = .93
APPENDIX 2

Patient - Physician Feedback Summary: 31 JAN 94

Average Response: Day Prior

Respondent Number: Day Prior

Average Response: 7 Day Cumulative

Respondent Number: 7 Day Cumulative

Average Response: range is from 5 [Strongly Agree] to 1 [Strongly Disagree]; 5 is better.
Respondent Number: states the total respondents for the time period covered in the average
Physician Identity Codes: are listed directly under the respective depicted data bar.
MEMORANDUM FOR Commandant, Academy of Health Sciences,
ATTN: HSHA-MH (Residency Committee),
Fort Sam Houston, TX 78234-6100

SUBJECT: Institutional Review Board Exempted Protocol

1. In accordance with the AR 40-38 and instructions contained in the Clinical Investigator's Handbook (WBAMC), the proposed research project by CPT Wm. Christopher Chambers, "Frequency of Patient Feedback to Physicians," is an Exempt Protocol.

2. CPT Chambers is directed to implement the research in a manner to pose no risk to survey subjects and to record survey responses in such a way that subjects cannot be identified directly or indirectly.

Robert F. Borries Jr.
ROBERT F. BORIES, JR.
Colonel, Medical Services Corps
Chief of Staff