Factors That Influence Line Managers' Perceptions of Hospital Performance Data

Liane Soberman Ginsburg

Objective. To design and test a model of the factors that influence frontline and midlevel managers' perceptions of usefulness of comparative reports of hospital performance.

Study Setting. A total of 344 frontline and midlevel managers with responsibility for stroke and medical cardiac patients in 89 acute care hospitals in the Canadian province of Ontario.

Study Design. Fifty-nine percent of managers responded to a mail survey regarding managers' familiarity with a comparative report of hospital performance, ratings of the report's data quality, relevance and complexity, improvement culture of the organization, and perceptions of usefulness of the report.

Extraction Methods. Exploratory factor analysis was performed to assess the dimensionality of performance data characteristics and improvement culture. Antecedents of perceived usefulness and the role of improvement culture as a moderator were tested using hierarchical regression analyses.

Principal Findings. Both data characteristics variables including data quality, relevance, and report complexity, as well as organizational factors including dissemination intensity and improvement culture, explain significant amounts of variance in perceptions of usefulness of comparative reports of hospital performance. The total R^2 for the full hierarchical regression model = .691. Improvement culture moderates the relationship between data relevance and perceived usefulness.

Conclusions. Organizations and those who fund and design performance reports need to recognize that both report characteristics and organizational context play an important role in determining line managers' response to and ability to use these types of data.

Key Words. Performance reporting, report cards, feedback, quality improvement

Performance measurement has emerged as an important public policy issue in health care. Data gathering on the performance of health systems, organizations, and individual care providers is being undertaken widely. These efforts measure clinical performance, financial performance, and performance in the area of patient satisfaction. A variety of stated purposes for performance measurement include helping consumers decide where to seek

care, holding providers accountable for outcomes of care, identifying areas for improvement, and identifying benchmark sources (Nelson et al. 1995; Miller and Leatherman 1999). The idea is that measurement is crucial if quality is to be improved (Nash 1998), and consumers should be informed when they make decisions about where to seek care (Wicks and Meyer 1999). The reality, however, is that despite increased efforts to measure performance, data are often not used (Marshall et al. 2000).

In response to calls to pay more attention to the impact of performance measurement (Bentley and Nash 1998; Eddy 1998; Turpin et al. 1996), this study draws on the feedback literature from organizational behavior and existing improvement literature to design and test a conceptual model of the factors that influence frontline and midlevel managers' perceptions of usefulness of comparative reports of hospital performance[1]. Frontline and midlevel hospital managers have been chosen as the focus for this study because although they are key actors in the quality improvement process, there were no studies found in the literature that systematically look at the impact of comparative reports of hospital performance on this group.

LITERATURE

Few studies have systematically looked at manager or provider perceptions of usefulness of health care performance reports. Maxwell (1998) conducted a survey to solicit senior executives' views of the usefulness and importance of cost and clinical outcomes data gathered for 59 diagnosis-related groups in all acute care hospitals in Pennsylvania as part of that state's Health Care Cost Containment Council's (PHC4) mandate. Respondents reported that the performance data were most important for the two quality purposes (improving clinical quality and encouraging competition based on quality) rather than cost purposes (e.g., cost containment and encouraging competition

Address correspondence to Liane Soberman Ginsburg, Ph.D., Assistant Professor, School of Health Policy & Management, York University, 4700 Keele Street, Atkinson 202A, Toronto, ON M3J 1P3, Canada.

based on price). Romano et al. (1999) asked California and New York administrators and quality managers about the perceived usefulness of risk-adjusted acute myocardial infarction (AMI) mortality reports for several different purposes including improving quality of care, improving quality of ICD-9 coding, negotiating with health plans, and marketing and public relations. A higher percentage of respondents reported that the performance information is more useful for quality improvement than for the other three purposes. These data shed some empirical light on Nelson et al.'s (1995) discussion of the potential uses of performance measurement data and suggest that hospital executives feel the data are of greater value for clinical improvement purposes than for financial purposes.

In terms of literature looking at evaluations of usefulness of large-scale comparative performance reports, the data are not very favorable. Berwick and Wald (1990) asked hospital leaders their opinions about some of the first hospital-specific mortality data released by the Health Care Financing Administration (HCFA) in 1987. A national sample was surveyed and hospital leaders had consistently negative opinions about the accuracy, interpretability, and usefulness of the data (e.g., 70 percent of respondents rated the data as poor in terms of overall usefulness to the hospital). Berwick and Wald concluded that: "publication of outcome data to encourage quality improvement may face severe and pervasive barriers in the attitudes and reactions of hospital leaders who are potential clients for such data" (1990, p. 247). It is true that these early attempts to measure performance using administrative data have since been heavily criticized for their inaccuracy; however, some of the most sophisticated risk-adjusted clinical outcomes data currently available, such as data collected and reported in New York and Pennsylvania, continue to receive negative reactions from providers and hospital executives. For instance, using the same measure as Berwick and Wald (1990), administrators in New York and California rated the overall quality of their statewide mortality reports as fair to good on a poor to excellent scale (Romano, Rainwater, and Antonius 1999). In other surveys of cardiovascular specialists, coronary artery bypass graft (CABG) mortality data were perceived to be of limited value and were reported to have had a limited effect on referral recommendations of cardiologists in Pennsylvania (Schneider and Epstein 1996) and New York (Hannan et al. 1997).

Two additional studies have looked at whether actual performance level achieved moderates ratings of usefulness. Romano et al. (1999) found that the chief executives of New York hospitals who were low-CABG-mortality outliers perceived the data to be significantly more useful than nonoutliers and

high-mortality outliers. In addition, nonoutliers rated the quality of the report significantly higher than high-mortality outliers and significantly lower than low-mortality outliers. Researchers in Quebec found that managers in hospitals that achieved poor performance ratings for complications and readmissions following prostatectomy, hysterectomy, and cholecystectomy provided ratings of data usefulness that were significantly lower than ratings provided by managers achieving average and strong performance ratings on those indicators (Blais et al. 2000). These data are consistent with ego-defensive reactions described in the feedback literature (Taylor, Fisher, and Ilgen 1984).

Simply generating performance data will not, in and of itself, lead to improvement (O'Connor et al. 1996; Eddy 1998). Moreover, there are elements of comparative performance reporting that may cause those who are the subject of measurement to question the validity of the data and react against the feedback source (Taylor, Fisher, and Ilgen 1984). These facts may help to explain the low ratings of perceived usefulness of cardiac outcomes data given by New York, California, and Pennsylvania executives and physicians.

Fortunately literature on initiatives such as the Massachusetts Health Quality Partnership (MHQP) (Rogers and Smith 1999) and the Northern New England Cardiovascular Disease Study Group (NNECVDSG) (O'Connor et al. 1996) suggests that, under certain conditions, regional comparison data can be perceived as being quite useful to providers and managers. Although these two initiatives did not provide explicit ratings of perceived usefulness the way that evaluations of the New York, Pennsylvania, and California initiatives have, it has been suggested that the attention that was paid to data reporting and improvement planning in both the MHQP and the NNECVDSG initiatives facilitated acceptance and use of the data by their provider/manager audiences. In the case of the Northern New England group, success may also be attributable to the nature of the sponsorship of this network—unlike most performance measurement initiatives, health care professionals led this initiative.

CONCEPTUAL MODEL AND STUDY HYPOTHESES

Based on the feedback literature as well as existing health care performance and improvement literature, Figure 1 shows a model of the factors hypothesized to influence perceived usefulness of performance data conceptualized in this study. Several data characteristics variables have been included in the model because, to be used effectively, the literature suggests

Figure 1:   Antecedents of Perceived Usefulness of Hospital Performance Data
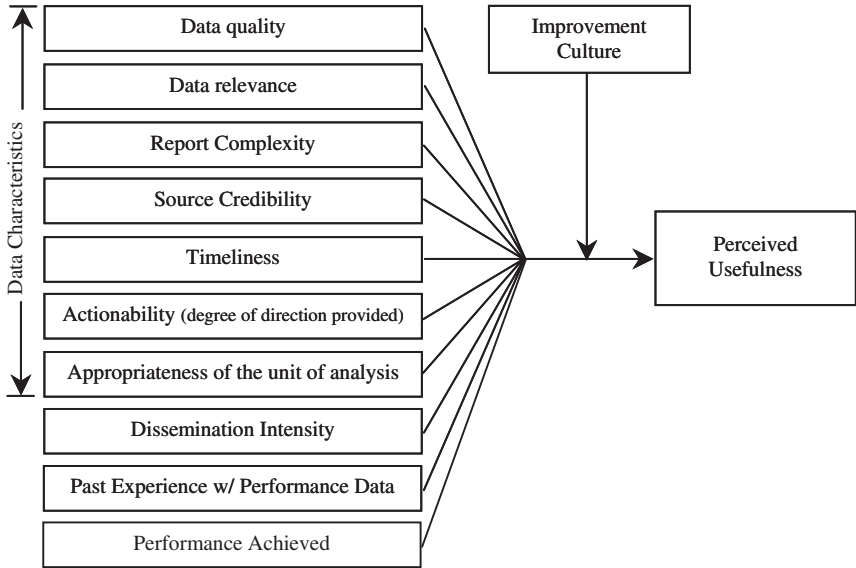
that report cards need to meet certain requirements: the information needs to be reliable (Nelson et al. 1995) and accurate (Schneider et al. 1999; Jennings and Staggers 1999), the data need to be relevant (Donaldson and Nolan 1997; Turpin et al. 1996), timely (Schneider et al. 1999; Nash 1998; Rainwater, Romano, and Antonius 1998), appropriately complex (Hibbard 1998; Solberg, Mosser, and McDonald 1997), and actionable (Eddy 1998; Thomas 1998; Palmer 1997; Bentley and Nash 1998; Nash 1998; Schwartz et al. 1999). Consistent with institutional theory (Meyer and Rowan 1977; DiMaggio and Powell 1983) and Huberman's (1994) model of the dissemination of research findings, managers' past experience with performance data, and the intensity of the dissemination effort, respectively, are also hypothesized to be important antecedents of perceived usefulness. The effect of actual performance achieved on perceived usefulness, discussed in the literature above, is also reflected in Figure 1. Finally, Figure 1 suggests that the improvement culture of an organization will moderate each of these relationships. This hypothesized moderator effect is based on work by Schneider (1983) and other interactionists, which would suggest that in addition to the data and reporting variables, there are likely to be contextual variables that influence perceptions of usefulness of hospital performance data. Although work by Romano and colleagues (1999) did not find a significant relationship between structural

contextual variables such as hospital ownership, size, occupancy, and AMI volume and perceived usefulness, it is argued here that a contextual variable reflecting the improvement culture of the hospital will be positively related to perceived usefulness. Corresponding to the relationships outlined in Figure 1, several hypotheses were developed (e.g., perceived data quality will be positively and significantly related to perceived usefulness). Improvement culture was hypothesized to moderate the relationship between each of the other variables outlined in Figure 1 and the dependent variable, perceived usefulness of performance data.

METHODS

Hospital Report '99 (*HR99*) is a multidimensional report of acute care hospital performance that was publicly released in the Canadian province of Ontario in December 1999 (Baker et al. 2000). The report was published as an independent research report by a team of researchers at the University of Toronto and it contained provincial-level as well as hospital-specific data for 38 indicators of hospital performance in four quadrants: patient satisfaction, clinical utilization and outcomes, financial performance and condition, and system integration and change.[2] In order to test the conceptual model outlined in Figure 1, a cross-sectional survey, with questions framed around the clinical utilization and outcomes data in *HR99,* was used. Survey data were collected between November 2000 and February 2001.

Sample and Questionnaire Administration

Eighty-nine acute care hospital corporations that participated in *Hospital Report '99* were used to define the sample. These 89 organizations provide more than 90 percent of acute care services in the province of Ontario. In the fall of 2000 a contact person at each hospital (identified previously during the data collection phase of *HR99*) received a letter describing this study and asking him or her to identify all line and midlevel managers with responsibility for in-patient stroke and medical cardiac patients (based on detailed role descriptions provided by the researcher). This process led to the identification of 344 line and midlevel managers. Stroke and cardiac care were chosen as the two clinical areas because both areas have a high volume of cases, thus ensuring that even the smallest and most remote hospitals in the province would have managers who met the sample eligibility criteria. Questionnaires were mailed along with a cover letter. These were followed up by reminder

cards two weeks later and a second mailing to all nonrespondents four weeks after that. Two hundred and two managers returned a study questionnaire for a response rate of 59 percent. Nonrespondents did not differ from respondents with respect to role or organizational tenure; however, physician managers and managers from teaching hospitals were underrepresented in the respondent group.

Questionnaire Content/Study Measures

The study questionnaire contained several sections.[3] A series of questions were included to determine the extent to which frontline and midlevel managers are familiar with the results of *HR99*. This information was important for understanding the extent to which these performance reports find their way to line-level managers and for determining eligibility for model testing. In terms of the variables in the conceptual model, dissemination intensity was computed as a summative score based on responses to nine binary items in the questionnaire (e.g., "I attended a presentation of the results in the hospital," "I went to the Internet and reviewed some of the report"). Respondents received one point for each item to which they provided an affirmative response. Respondents were given two points for two of the items that reflected more proactive behaviors on the part of the respondent: "I shared some of the results with staff or other managers in my organization" and "I am involved in ongoing initiatives that have resulted primarily from *Hospital Report '99.*" The maximum possible score for this variable is 11[4].

Four items were initially created to measure report complexity of *HR99*; however, they were not found to have an acceptable coefficient alpha (.64). Accordingly, one of the items ("*Hospital Report '99* is excessively complex") was retained so that a measure of report complexity could be included in the regression analysis. This item was measured using a seven-point agree–disagree Likert-type response scale. Fifteen items were created to measure the remaining data characteristics variables in Figure 1. Exploratory Factor Analysis (EFA) was performed on these 15 items. After the removal of two items with complex loadings, EFA using principal axis factoring and oblique rotation revealed the presence of only two nontrivial data characteristics factors. These factors have been labeled data quality and data relevance. The factor-loading matrix is provided in Table 1. Table 1 also includes a four-factor matrix, based on more liberal decision rules regarding the number of factors to extract[5]. The four-factor model is included simply to suggest that, with the creation of additional items and the collection of more data, a future study may

reveal that the data quality factor is comprised of separate timeliness, believability, and actionability factors.

   Based on the results of the factor analysis, a data relevance variable was computed as the mean of the first three items in Table 1 (e.g., "The clinical data in *HR '99* are applicable to me and my day to day work"), all measured using a

Table 1:   Factor Loadings for Two and Four-Factor Data Characteristics Models

Questionnaire Item*	Two-factor solution		Four-factor solution			
	1[†]	*2[††]*	*1[†]*	*2[‡]*	*3[‡‡]*	*4[e]*
The clinical data in *HR '99* apply to the level in the organization at which I work.	− **.911**	− .124	− **.865**	− .038	− .089	.020
The clinical data in HR '99 are relevant to my work.	− **.835**	.115	− **.810**	− .031	.066	.173
The clinical data in HR '99 are applicable to me and my day to day work.	− **.631**	.268	− **.696**	.220	.154	− .102
It is clear from the clinical data in HR '99 where we have opportunities for improvement.	− .060	**.743**	− .079	**.660**	.039	.123
The clinical data in HR '99 provide insufficient direction for change. (reversed)	.128	**.615**	.116	**.663**	− .046	.030
The clinical data in HR '99 provide meaningful hospital peer group comparison information.	− .081	**.685**	− .108	**.656**	.043	.045
The clinical data in HR '99 provide direction about actions we can take in order to bring about improvement.	− .064	**.695**	− .117	**.627**	.165	− .029
The clinical data in HR '99 yielded results at a level that is impractical for me. (reversed)	− .202	**.567**	− .214	**.488**	.022	.121
The Clinical Utilization & Outcomes (clinical) data in HR '99 were collected and reported in a timely fashion.	.017	**.521**	− .013	− .095	**.866**	.118
By the time we received the clinical data in HR '99, they were outdated. (reversed)	.052	**.392**	− .007	.105	**.468**	− .046

Table 1:   Continued

Questionnaire Item*	Two-factor solution		Four-factor solution			
	1[†]	2[††]	1[†]	2[‡]	3[‡‡]	4[e]
The clinical data in HR '99 were provided by a credible source.	− .256	**.374**	− .167	.024	− .113	**.657**
The clinical data in HR '99 are believable.	− .022	**.688**	.063	.265	.048	**.628**
The clinical data in HR '99 accurately describe one aspect of hospital performance.	− .033	**.532**	.029	.023	.201	**.562**

*Note that 2 of the 15 items had complex loadings and were therefore removed from the factor analysis, leaving 13 items in this table. In order to have a sufficient sample, EFA (Explanatory Factor Analysis) was performed using a second sample of hospital managers who responded to the same items on a similar questionnaire. Accordingly, $n = 218$ for the EFA.

[†]Items relate to data relevance.

[††]Items relate to overall data quality issues (including actionability, timeliness, and believability).

[‡]Items relate to actionability or, the degree to which clinical data provide direction for how to bring about improvement.

[‡‡]Items relate to data timeliness.

[e]Items relate to believability/credibility of the data.

seven-point agree–disagree Likert-type scale. The coefficient alpha for this three-item scale is .88. Using the same response scale, the data quality variable was calculated as the mean of the remaining 10 items shown in Table 1 (e.g., "The clinical data in *HR '99* are believable"). The coefficient alpha for this 10-item scale is .86.

Past experience with performance data was computed as a mean of five items created to measure respondents' experience with performance indicators and indicator data other than *HR99*. Sample items in this scale include: "not counting *HR99,* I have been involved in work where indicators have been identified and/or collected" and "not counting *HR99*, I have received performance indicator data collected by others in my organization." "Indicators" were defined in the question stem and a four-point response scale was used ("never" "on 1 occasion," "on 2–4 occasions," and "on >4 occasions"). The coefficient alpha for these five items was .77. The perceived improvement culture measure reflects the extent to which a respondent feels his or her hospital values performance data and supports using the data to bring about improvement. The variable is calculated as the mean score of nine items created to measure improvement culture. Using a seven-point

agree–disagree response scale, the variable is based on two items borrowed from the Quality in Action (QiA) questionnaire (Baker, Murray, and Tasa 1998), two items borrowed and adapted from Caron and Neuhauser's (1999) work on the organizational environment for quality improvement, and five items created to measure improvement culture (e.g., "This organization devotes resources to measurement initiatives, but the results often end up sitting on a shelf," "A leader(s) in this organization truly promotes/champions performance measurement and improvement activities"). Exploratory Factor Analysis of these nine items supported one improvement culture factor, with coefficient alpha equal to .84.

The dependent variable, perceived usefulness of performance data, was calculated as the mean of six items adapted from Davis's (1989) measure of perceived usefulness in the context of technology assessment. Davis (1989) defined perceived usefulness as "the tendency to use or not use an application to the extent that a person believes it will help improve job performance" (Dansky 1999, p. 442). A seven-point agree–disagree Likert-type response scale was again used. Examples of the adapted items include: "The clinical data in *HR99* will contribute to work productivity" and "The clinical data in *HR99* will enhance effectiveness in my work." The coefficient alpha for this six-item scale is .95. Items related to managers' role and organizational tenure, functional and educational backgrounds were also incorporated into the questionnaire. Responses to all negatively phrased questionnaire items were recoded so that higher scores indicate more positive ratings for all study variables. Means and standard deviations for each variable can be found in Table 2. Where applicable, scale alphas have also been included in the diagonal of Table 2.

Analysis

As described above, Exploratory Factor Analysis (EFA) was performed where it was necessary to establish the dimensionality of a study construct. Because factor analysis did not support the presence of unique timeliness, actionability, or appropriate unit of analysis factors, some of the hypothesized relationships outlined in Figure 1 could not be tested in the regression analysis. To test the remaining study hypotheses, five hierarchical (moderated) regression analyses were carried out to test whether improvement culture moderates the relationship between the independent variables (data relevance, data quality, report complexity, dissemination intensity, and past experience with performance data) and the dependent variable, perceived usefulness of

Table 2: Means, Standard Deviations, Alphas‡‡, and Pearson Correlations†

	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
1. Past experience	3.39	0.61	.77										
2. Physician dummy	‡19.6%		–.22*										
3. Organization tenure	14.77	9.82	.08	.24*									
4. Dissemination	3.71	2.44	.38**	–.21*	.01								
5. Teaching hospital	‡15.9%		–.05	.13	–.04	–.04							
6. Hospital revenue (millions of $)	67.7	65.8	.03	.03	–.15	.05	.77**						
7. Improvement culture	4.53	0.97	.15	–.10	.04	.26**	–.17	–.07	.84				
8. Report complexity	4.36	1.28	.00	.03	.18	.25*	.00	.04	–.01				
9. Data relevance	5.13	0.97	.24*	–.18	.01	.42**	.01	.07	.27**	.37**	.88		
10. Data quality	4.53	0.81	.14	.01	.12	.39**	–.04	–.03	.25*	.46**	.53**	.86	
11. Perceived usefulness	4.44	1.12	.11	–.05	.11	.40**	–.14	–.10	.35**	.40**	.48**	.79**	.95
12. Low AMI performance	‡24%††		–.31*	–.03	.07	–.05	–.07	–.05	–.13	–.15	–.10	–.18	–.21*

† $n = 107$. Although 202 subjects provided a questionnaire response, as reported in the results, 35% (70) were not familiar with HR99, leaving 132 subjects for model testing. However, only 107 subjects provided complete data for model testing; 25 subjects were lost due to missing data on the three data characteristics variables (report complexity, data relevance, data quality). Values could not be imputed because these 25 subjects responded to only 1 or 2 of the 19 items used to measure these three variables.

‡‡Coefficient alphas are italicized and reported in the diagonal where applicable.

‡Dummy variables (percent positive reported).

††Performing low on at least one of four AMI or stroke clinical utilization and outcomes indicators.

*Correlation is significant at the 0.05 level (two-tailed).

**Correlation is significant at the 0.01 level (two-tailed).

performance data. Moderation would be supported by a significant change in the multiple squared correlation coefficient (R^2) when an interaction term is added in step three in each hierarchical regression model.[6]

Hierarchical regression was also used to test the unique effect of sets of individual versus organizational versus data characteristics variables (including several control variables) on perceived usefulness of performance data. Individual level variables including a dummy variable for physician status, tenure in the organization, and experience with performance data are entered in block (step) one. In block two, organizational level variables are added. These include a dummy variable for low performance on the AMI and stroke indicators in the respondents' organization, a dummy for teaching status of the respondents' organization, total revenue (a proxy for hospital size), perceived improvement culture, and dissemination intensity. In blocks three, four, and five each of the three data characteristics variables are added separately so their effects can be tested in the presence of individual and organizational control variables. In block six all of the interaction terms are added in order to rule out the possibility that any of their effects were suppressed in the earlier analysis where the control variables were absent.

RESULTS

For the 202 respondents, the average tenure in the organization was 15.8 years (SD = 9.9), with average role tenure equal to 5.9 years. Nineteen percent of respondents are physician managers, 66 percent of respondents have a nursing background, with two-thirds of that group prepared at the Baccalaureate level or higher. Twelve percent of respondents have a Masters of Health Administration degree (M.H.Sc./M.H.A.) or an M.B.A.

More than one-third of respondents (35 percent) indicated that they were "not at all familiar with any of the *HR99* related reports or results." This seemingly high level of unfamiliarity, low levels of knowledge about *HR99* reported by respondents, and mediocre evaluations of the report's content (evident from the mean scores on several study variables reported in Table 2), all suggest that *HR99* was not highly valued by frontline and midlevel managers in hospitals who were the subject of the report. This conclusion is quite consistent with open-ended comments provided by study respondents indicating that (1) poor access to the report from within the organization, (2) lack of resources, (3) competing priorities, (4) lack of time and support to understand the data, and (5) lack of specificity in many of the results, together,

posed significant barriers that prevented people from being able to make use of the data in their organization.

Correlation and Moderated Regression Analyses

Table 2 provides descriptive statistics and zero-order correlation coefficients for all variables in the conceptual model as well as the control variables listed previously. Physician managers reported having less experience with performance indicator data ($r = -.22$, $p<.05$) and longer tenure in the organization ($r = .24$, $p<.05$). Low performance on the AMI and stroke clinical indicators in *HR99* was negatively correlated with perceptions of usefulness of *HR99* ($r = -.21$, $p<.05$) and a respondent's past experience with indicator data ($r = -.31$, $p<.05$). As expected, dissemination intensity, improvement culture, and the three data characteristics variables were all positively correlated with perceived usefulness ($r = .35 - .79$, $p<.01$). The data characteristics variables were also significantly interrelated ($r = .37 - .53$, $p<.01$).

Hypothesis Testing. Table 3 shows the results of each of the separate hierarchical regression analyses. Data quality, data relevance, report complexity, and dissemination intensity each explained a significant amount of variance in perceived usefulness over and above variance explained by improvement culture. The relationship between past experience with performance data and perceived usefulness was not supported ($R^2 = .004$, n.s.). The role of improvement culture as a moderator was only supported for data relevance ($R^2 = .03$, $p<.05$), indicating that the effect of data relevance on perceived usefulness is conditional on the level of perceived improvement culture.

Post hoc testing and plotting were carried out to investigate the nature of the interaction between improvement culture and data relevance. Figure 2 shows the relationship between data relevance and perceived usefulness at three different levels of improvement culture. The graph shows that at very high levels of perceived improvement culture, data relevance is no longer significantly related to perceived usefulness—with a strong improvement culture people perceive performance data to be useful even if the data *are not* seen as having much relevance. The effect size for this interaction (strength of the interaction over and above the main effects) is small ($f^2 = .05$) (Cohen and Cohen 1983).

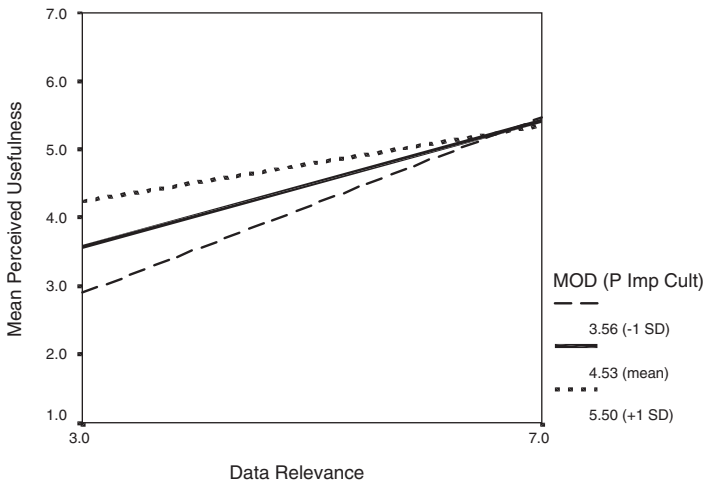### Full Hierarchical Regression Model with Control Variables

The unique effect of sets of individual versus organizational versus data characteristic variables on the dependent variable, perceived usefulness was

Table 3:   Results of Moderated Regression Analyses[†]

Step	Independent Variable	Total R^2	R^2
1	Improvement culture	.13**	.13**
2	Data quality	.67**	.54**
3	Improvement culture X data quality	.67**	.00
1	Improvement culture	.13**	.13**
2	Data relevance	.30**	.18**
3	Improvement culture X data relevance	.33**	.03*
1	Improvement culture	.13**	.13**
2	Report complexity	.29**	.16**
3	Improvement culture X report complexity	.31**	.02
1[††]	Improvement culture	.12**	.12**
2[††]	Dissemination intensity	.23**	.11**
3[††]	Improvement culture X dissemination	.23**	.00
1	Improvement culture	.13**	.13**
2	Past experience	.13**	.00
3	Improvement culture X experience	.13**	.00

[†]$n = 107$, [††]$n = 106$, *$p < .05$, **$p < .01$

Figure 2:   Interaction between Improvement Culture and Data Relevance

Table 4:   Results of Full Hierarchical Regression Analysis

	Model 1 Beta	Model 2 Beta	Model 3 Beta	Model 4 Beta	Model 5 Beta	Model 6 B^9
Block 1—Individual Variables						
Past experience	.042	− .179	− .176	− .181	− .112	− .180
Organizational tenure	− .064	.011	.007	.032	− .010	.039
Physician dummy variable	.059	.048	.004	.019	.006	.002
Block 2—Organizational Variables						
Performance achieved		− .204*	− .152	− .155	− .093	− .273
Dissemination intensity		.397**	.320**	.258**	.127	.072*
Teaching hospital dummy variable		− .050	− .020	− .037	− .062	− .246
Total revenue		− .058	− .098	− .097	− .036	− 2.65
Improvement culture		.248**	.285**	.235*	.156*	.157*
Block 3—Data Characteristic						
Report complexity			.329**	.250**	.077	.051
Block 4—Data Characteristic						
Data relevance				.244**	.006	− .023
Block 5—Data Characteristic						
Data quality					.677**	.901**
Block 6—Interactions						
Experience × ImpCulture						.044
Dissem × ImpCulture						.028
Report complexity × ImpCulture						− .053
Relevance × ImpCulture						− 0.61
Data quality × ImpCulture						− 0.43
Total R^2	.010	.287**	.385**	.427**	.691**	.706**
Change in R^2	.010	.278**	.098**	.042*	.264**	.014

*$p < .05$, **$p < .01$.

entered in the regression model next, explain 28.7 percent of the variance in perceived usefulness ($R^2 = .278$, $p < .01$). Finally, when data characteristic variables are added in Models 3, 4, and 5, they explain an additional 40 percent of the variance in perceived usefulness of performance data (Model 3 change in R^2 + Model 4 change in R^2 + Model 5 change in $R^2 = .404$). Because the R^2 change in Model 6 was not significant ($R^2 = .014$, n.s.) and because there is no strong theoretical argument for leaving the interaction terms in the analysis, they are removed in order "to permit more powerful tests of the other

effects" (Aiken and West 1991, p. 105). The total R^2 for the reduced model (5) is .691.

Although the regression coefficients reported in Table 4 appear to suggest that, in relative terms, data quality is a more important predictor of perceived usefulness than variables significant in Models 1 through 4, as Cooper and Richardson (1986) make clear, comparisons between predictor variables are often unfair because one variable may be procedurally or distributionally advantaged. For instance, the data quality factor has more items than the other variables, may be more reliable than other variables, and may well not be unidimensional, as noted in the description of the factor analysis results. Ultimately, these results indicate a significant improvement in the total R^2 with each model in the hierarchical analysis ending with the reduced model, Model 5.

DISCUSSION

The goal of this study was to try to understand and identify the factors that influence frontline and midlevel managers' perceptions of usefulness of comparative reports of hospital performance. To begin, the data gathered in this study revealed that 35 percent of line and midlevel managers in organizations who participated in *HR99* reported they were unfamiliar with the *HR99* results. Although somewhat disconcerting, this finding is consistent with data from Romano et al. (1999) showing that only 69 percent of nursing directors "received or discussed" the California Hospital Outcomes Project report. Other studies, reporting on the *use* of comparative reports of hospital performance for simple process changes, have shown that only about 30 percent of organizations report using the data (Longo et al. 1997; Rainwater, Romano, and Antonius 1998). As we try to gain insight into the impact of comparative reports of hospital performance, it seems clear that while comparative performance reports may land at the door of 100 percent of organizations in a certain jurisdiction, they find their way to the frontlines in only a subgroup of those organizations, and the data are put to use in an even smaller subgroup. Those who fund and design performance reports should therefore be reminded, as O'Connor et al. (1996) and Eddy (1998) have pointed out, that simply generating performance data will not, in and of itself, lead to improvement.

In this study it was hypothesized that several variables related to characteristics of the performance data, intensity of dissemination, and past

experience with performance data would be significantly related to perceived usefulness of performance data. A moderating effect for perceived improvement culture was also hypothesized. Results indicate that data quality, data relevance, report complexity, and intensity of dissemination are all significantly related to perceived usefulness of performance data. These study findings are now briefly discussed in the context of existing literature. Opportunities for future research are also suggested.

Results of this study showing that perceived data quality of performance reports is significantly related to perceptions of usefulness of those reports for line and midlevel hospital managers are highly consistent with literature suggesting that performance data need to be reliable and accurate if they are to be used constructively (Nelson et al. 1995; Schneider et al. 1999; Jennings and Staggers 1999). These findings are also consistent with literature showing that hospital executives and medical leaders have identified poor quality data as one of the biggest limitations of statewide performance reports in Pennsylvania (Schneider and Epstein 1996), California, and New York (Romano et al. 1999). Although the feedback literature suggests that assessments of data quality/accuracy are among the first cognitive reactions to feedback, ego defensive reactions to performance data may cause those who are the subjects of performance reports to question the quality or usefulness of the data (Taylor, Fisher, and Ilgen 1984). As indicated in Table 2, performance achieved is negatively related to perceived usefulness of performance data. Because assessments of performance data quality can be influenced by phenomenon other than one's strictly objective views about quality of the feedback, further research is needed to discover how to mitigate negative assessments of data quality (and subsequent negative assessments of perceived usefulness) that may result from factors such as poor performance. Not unrelated is the need for research on the effects of publicly reporting performance data. Health care performance reports designed with improvement aims in mind are often reported publicly despite evidence from the feedback literature which has shown that people prefer more private modes of obtaining performance feedback, particularly if they feel they may be performing poorly (Northcraft and Ashford 1990). The need to carefully study the effects of publicly reporting performance data has also been suggested by others (e.g., Marshall et al. 2000).

The relationship between perceived relevance of performance data and perceptions of usefulness of those data found in this study is consistent with claims in the improvement literature that, in order to be used constructively, report cards need to be relevant (Nelson et al. 1995) and with simple elements

of feedback theory which point out that people who fail to see the relevance of feedback will disregard the information (Taylor, Fisher, and Ilgen 1984). Other empirical studies conducted with management samples at different organizational levels have also demonstrated that there is a relationship between performance data relevance/research finding relevance and action that is taken in response to those data (Turpin et al. 1996; Weiss and Weiss 1981). Not unrelated are results indicating there is a significant relationship between performance report complexity and perceptions of usefulness—a finding consistent with Hibbard's work (1998) showing that people have difficulty processing large amounts of performance information. Studies conducted with senior hospital executives also showed that complexity was an important factor in assessments of California AMI mortality reports (Rainwater, Romano, and Antonius 1998) and Ontario's *HR99* (Baker and Soberman 2001).

Although findings indicating that data relevance and report complexity are important factors in line and senior managers' assessments of hospital performance data underscore the need to pay attention to the information needs of the intended users of performance reports, actually generating performance reports that are appropriately complex and relevant for line and senior management groups *alike* is highly problematic. Senior leadership values high-level summary reports documenting performance in multiple areas (Baker and Soberman 2001). Line managers want detailed, unit-specific data on a selected number of measures that pertain specifically to their area of work (Donaldson and Nolan 1997; Rogers and Smith 1999; Murray et al. 1997; Huberman 1994). Accordingly, what constitutes relevant and appropriately complex data will be different for line and senior management groups (never mind providers). In designing performance measurement initiatives it therefore becomes critical to define one stakeholder audience and to define that audience more narrowly than the "provider organization," a group recently identified as the most promising audience for performance data (Marshall et al. 2000). Despite strong arguments suggesting that multipurpose performance reports can create barriers to successful use of performance data (Solberg, Mosser, and McDonald 1997; Nelson et al. 1995; Rogers and Smith 1999; O'Leary 1995), such reports seem to be the rule rather than the exception. Empirical research is needed to critically examine the utility of performance reports designed with multiple purposes and multiple audiences in mind.

The hypothesis that past experience with performance data would be positively related to perceptions of usefulness of *HR99* was not supported.

Given the highly politicized nature of performance reporting at this time, it is possible that failure to support this hypothesis has to do with the fact that performance measurement has not yet become institutionalized (Meyer and Rowan 1977; DiMaggio and Powell 1983). Had the broader health policy and management environments in Ontario been genuinely "hot" on performance measurement, perhaps the effect of past experience would, as hypothesized, have been positive. The conceptual model designed and tested in this research was focused at the individual and organizational levels and therefore was not able to reasonably incorporate larger contextual issues such as the broader political environment for performance measurement. Future research might look at the macro and system level variables that influence perceptions and use of performance data.

Study findings demonstrating that intensity of dissemination is significantly related to perceptions of usefulness of comparative reports of hospital performance are consistent with aspects of Huberman's dissemination effort model for the utilization of research findings (1994). Huberman's model states that multiple modes and channels of communication and redundancy of key messages are two important features of dissemination competence. Others have also found that dissemination of performance data was perceived to be more effective if it involved the use of multiple strategies (Turpin et al. 1996). Future studies might investigate which dissemination methods are most effective under different conditions. As an example, such investigations could draw on the decision-making literature showing that "rich" communication strategies such as face to face contact are effective when there is personal or political opposition to the information being disseminated while "lean" communication mechanisms such as written documents are appropriate when value conflicts are not anticipated (Thomas and Trevino 1993; Nutt 1998).

Finally, this study looked at the impact of organizational improvement culture on perceptions of usefulness of hospital performance data. The findings confirm aspects of Huberman's (1994) dissemination effort model that suggest both the commitment of key administrators and the presence of an organizational mandate have an impact on the utilization of research findings. Qualitative, open-ended comments from the study questionnaires demonstrate that, for line managers, senior leadership plays an important "agenda-setting" role that, to a large extent, determines the response to performance data within organizations. Having a senior person "champion" performance data: (1) directs managers' attention to the data, (2) tells them the data are a priority, and (3) provides (or in some cases falls short of providing) the

required support to use the data (Baker and Soberman 2001). The significant moderator effect shown in Figure 2, demonstrating that with a strong improvement culture performance data are perceived as being useful even though they may not be seen as having much relevance, also speaks directly to the importance of improvement culture in shaping the perceptions, and perhaps behaviors, of organizational members. Although one could argue that this potentially profound influence might be a boon for bringing about improvement in hospitals, it also highlights a more contentious, controlling aspect of a strong improvement culture. This controlling aspect of TQM (total quality management) philosophy has been the subject of some criticism and is believed to be the source of serious problems for the successful implementation of TQM in industry (Hackman and Wageman 1995) and in health care organizations (Soberman Ginsburg 2001). The potential impact of improvement culture and senior leadership's role in promoting the use of performance data is a topic that is in need of further study. For instance, consistent with one avenue of research suggested earlier, it would be promising to look at whether improvement culture moderates the relationship between performance achieved and perceptions of data quality/perceived usefulness. Because of power problems associated with finding significant moderator effects with dichotomous variables (Aguinis and Pierce 1998) (performance achieved was a binary variable in this study), it wasn't possible to investigate these relationships in this study.

This study provides a useful model for thinking about and studying certain aspects related to the impact of performance measurement. Although the measures used were largely new and can benefit from further refinement and validation, this initial exploratory effort provides a useful model and set of measures that health services researchers can use to begin to quantitatively study the impact of performance measurement. Other limitations of this study include the fact that this study relied on self-report questionnaire data, which are subject to social desirability biases as well as common methods bias. Future studies might try to include unobtrusive measures alongside self-report measures. A final limitation surrounds the possibility that there was insufficient power to adequately test for some of the hypothesized moderating effects because such a high percentage of cases were lost to respondents who were unfamiliar with *HR99*.[7]

In terms of the generalizability of the study findings, because familiarity with *HR99* is likely to have had an effect on response rates, findings related to line manager levels of familiarity with *HR99* are not generalizable to acute care hospital clinical line managers.[8] Conversely, the *relatedness* of the study

variables is unlikely to influence response rates and findings related to model testing (e.g., relationships between the independent and dependent variables) should therefore be generalizable.

Given the absence of evidence demonstrating that performance reports are being widely used for improvement purposes, it is critical that we try to understand line managers' perceptions of and responses to these reports. Consumer choice models of performance improvement, which are quite common in the performance literature and suggest that performance measurement will lead to improvement via traditional market demand forces, have not been subjected to any rigorous empirical tests. A professional ethos model of performance improvement has not been subjected to proper empirical examination either. Admittedly, the performance and improvement literatures would be considered "new" relative to most work in the behavioral and organization sciences; it is therefore critical that researchers look to knowledge that exists in other fields as we try to understand how to truly use performance data for bringing about improvements in health care delivery. This study has attempted to move in that direction.

NOTES

1. Perceived usefulness was chosen as the dependent variable because the processes that occur between performance measurement, improvement activity, and seeing improvements on those baseline measurements are sufficiently complex that there is a great deal of opportunity for confounding to occur (Eddy 1998). Measures of perceived usefulness of performance reports can and have been studied as dependent variables (Turpin et al. 1996; Romano, Rainwater, and Antonius 1999; Longo et al. 1997). Moreover, in the domain of technology assessment, Davis (1989) found that the construct of perceived usefulness is significantly correlated with current use and a person's self-prediction of future use.
2. A full copy of *Hospital Report '99* is available at http//www.oha.com, then go to Hospital Reports.
3. A copy of the questionnaire is available from the author.
4. It is not appropriate to calculate a coefficient alpha for this variable because some of these items may be mutually exclusive (e.g. if I attended a presentation outside the organization, I may be less likely to attend a presentation inside the organization).
5. For example, use of the eigenvalue > 1 criteria (Kaiser 1960) or Scree test (Cattell 1966) as apposed to the method of parallel analysis (Cota, Longman, Holden, Fekken, and Xinaris, 1993) used to arrive at the two-factor solution.
6. The improvement culture of a hospital represents a contextual element *within which* dissemination takes place and *within which* managers assess data quality, relevance, and so on. From a theoretical standpoint, it is therefore appropriate to enter

improvement culture as the first step in the model followed by the focal study variables (Cohen and Cohen 1983). In addition, in order to reduce the correlation between the independent variables and the product term (and avoid multi-collinearity problems), the independent variables were centered and interaction terms were created from the centered variables (Aiken and West 1991).

7. Perceived improvement culture was not found to moderate the relationship between perceived usefulness and any of the proposed antecedents other than data relevance. However, given the measures used in this study, it is possible that genuine interaction effects may have been suppressed. For instance, to find a moderate effect, Aiken and West (1991) show that a sample size of 103 is necessary to achieve power = .80 for testing an interaction at alpha = .05. However, this sample size calculation of 103 assumes the absence of measurement error and it has been shown that when reliability drops to .80, variance accounted for by the interaction, power, and effect size are all reduced by approximately 50 percent (Aiken and West 1991; Evans 1985). Although the coefficient alpha for many of the study variables is quite strong (e.g., .87 for the perceived data quality scale), measurement error in the first order variables dramatically reduces the reliability of the product terms (Busemeyer and Jones 1983; Aiken and West 1991) and it is therefore possible that some of the true variance explained by the interaction terms in this research has been suppressed. Attenuation of the interaction terms is further possible given that there is likely correlated error among the study variables (Evans 1985).

8. However, even if all nonrespondents were familiar with *HR99* (an unlikely scenario), that still leaves 20.3 percent of line and midlevel managers unfamiliar with the performance report.

9. Although standardized Betas are reported for Models 1 through 5, unstandardized *b* is reported for Model 6 because the inclusion of interactions required centering all variables used to form interaction terms.

ACKNOWLEDGMENTS

REFERENCES

Aguinis, H. A., and C. A. Pierce. 1998. "Statistical Power Computations for Detecting Dichotomous Moderator Variables with Moderated Multiple Regression." *Educational and Psychological Measurement* 58 (4): 668–76.

Aiken, L. S., and S. G. West. 1991. *Multiple Regression: Testing and Interpreting Interactions.* Newbury Park, CA: Sage.

Baker, G. R., and L. R. Soberman. 2001. "Organizational Response to Performance Data: A Qualitative Study of Perceptions and Use of HR 99." Working paper, Department of Health Policy, Management, and Evaluation, University of Toronto.

Baker, G. R., G. M. Anderson, A. D. Brown, I. McKillop, C. Murray, M. Murray, and G. H. Pink, eds. 2000. *Hospital Report '99: A Balanced Scorecard for Ontario Acute Care Hospitals.* Toronto.

Baker, G. R., M. Murray, and K. Tasa. 1998. "The *Quality in Action* Instrument: An Instrument to Measure Healthcare Quality Culture." Working paper, Department of Health Administration, University of Toronto.

Bentley, J. M., and D. B. Nash. 1998. "How Pennsylvania Hospitals have Responded to Publicly Released Reports on Coronary Artery Bypass Graft Surgery." *Joint Commission Journal on Quality Improvement* 24 (1): 40–9.

Berwick, D. M., and D. L. Wald. 1990. "Hospital Leasers' Opinions of the HCFA Mortality Data." *Journal of the American Medical Association* 263 (2): 247–9.

Blais, R., D. Larouche, P. Boyle, R. Pineault, and S. Dube, 2000. "Do Responses of Managers and Physicians to Information on Quality of Care Vary by Hospital Performance?" *Abstract Book/AHSR* 16: 323.

Busemeyer, J. R., and L. E. Jones. 1983. "Analysis of Multiplicative Combination Rules When the Causal Variables Are Measured with Error." *Psychological Bulletin* 93: 549–62.

Caron, A., and D. Neuhauser. 1999. "The Effect of Public Accountability on Hospital Performance: Trends in Rates for Cesarean Sections and Vaginal Births after Cesarean Section in Cleveland, Ohio." *Quality Management in Health Care* 7 (2): 1–10.

Cattell, R. B. 1966. "The Scree Test for the Number of Factors." *Multivariate Behavioral Research* 1: 245–76.

Cohen, J., and P. Cohen. 1983. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* 2d ed. Hillsdale, NJ: Lawrence Erlbaum.

Cooper, W. H., and A. J. Richardson. 1986. "Unfair Comparisons." *Journal of Applied Psychology* 71: 179–84.

Cota, A. A., R. S. Longman, R. R. Holden, G. C. Fekken, and S. Xinaris. 1993. "Interpolating 95th Percentile Eigenvalues from Random Data: An Empirical Example." *Educational and Psychological Measurement* 53: 585–96.

Davis, E. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly* 13 (3): 319–40.

Dansky, K. H. 1999. "Electronic Medical Records: Are Physicians Ready?." *Journal of Healthcare Management* 44 (6): 440–55.

DiMaggio, P. J., and W. W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organization Fields." *American Sociological Review* 48 (April): 147–60.

Donaldson, M. S., and K. Nolan. 1997. "Measuring the Quality of Health Care: State of the Art." *Joint Commission Journal on Quality Improvement* 23 (5): 283–92.

Eddy, D. M. 1998. "Performance Measurement: Problems and Solutions." *Health Affairs* 17 (4): 7–25.

Evans, M. G. 1985. "A Monte Carlo Study of the Effects of Correlated Method Variance in Moderated Multiple Regression Analysis." *Organizational Behavior and Human Decision Processes* 36: 305–23.

Hackman, J. R., and R. Wageman. 1995. "Total Quality Management: Empirical, Conceptual and Practical Issues." *Administrative Science Quarterly* 40: 309–42.

Hannan, E. L., C. C. Stone, T. L. Biddle, and B. A. DeBuono. 1997. "Public Release of Cardiac Surgery Outcomes Data in New York." *American Heart Journal* 134 (1): 55–61.

Hibbard, J. H. 1998. "Use of Outcome Data by Purchasers and Consumers: New Strategies and New Dilemmas." *International Journal for Quality in Health Care* 10 (6): 503–8.

Huberman, M. 1994. "Research Utilization: The State of the Art." *Knowledge and Policy* 7 (4): 13–33.

Jennings, B. M., and N. Staggers. 1999. "A Provocative Look at Performance Measurement." *Nursing Administration Quarterly* 24 (1): 17–30.

Kaiser, H. F. 1960. "The Application of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement* 20: 141–51.

Longo, D. R., G. Land, W. Schramm J. Fraas, B. Hoskins, and V. Howell. 1997. "Consumer Reports in Health Care: Do They Make a Difference in Patient Care?" *Journal of the American Medical Association* 278 (19): 1579–84.

Marshall, M. N., P. G. Shekelle, S. Leatherman, and R. H. Brook. 2000. "The Public Release of Performance Data: A Review of the Evidence." *Journal of the American Medical Association* 283 (14): 1866–74.

Maxwell, C. I. 1998. "Public Disclosure of Performance Information in Pennsylvania." *Joint Commission Journal on Quality Improvement* 24: 491–502.

Meyer, J. W., and B. Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83: 340–63.

Miller, T., and S. Leatherman. 1999. "The National Quality Forum: A 'Me-Too' or a Breakthrough in Quality Measurement and Reporting?" *Health Affairs* 18 (6): 233–7.

Murray, M., P. Norton, L. Soberman, D. Roberge, M. Pineault, and R. Pineault. 1998. "How Are Patient Evaluation Survey Data Disseminated in Healthcare Organizations." Working paper, Department of Health Administration, University of Toronto.

Nash, D. B. 1998. "Report on Report Cards." *Health Policy Newsletter* 11 (2): 1–3.

Nelson, E. C., P. B. Batalden, S. K. Plume, N. T. Mihevc, and W. G. Swartz. 1995. "Report Cards or Instrument Panels: Who Needs What?" *Joint Commission Journal on Quality Improvement* 21 (4): 150–60.

Northcraft, G. B., and S. Ashford. 1990. "The Preservation of Self in Everyday Life: The Effects of Performance Expectations and Feedback Context on Feedback Inquiry." *Organizational behavior and Human Decision Processes* 47: 42–64.

Nutt, P. C. 1998. "How Decision Makers Evaluate Alternatives and the Influence of Complexity." *Management Science* 44 (8): 1148–66.

O'Connor, G. T., S. K. Plume, E. Olmstead, J. Morton, and C. Maloney, 1996. "A Regional Intervention to Improve the Hospital Mortality Associated with Coronary Artery Bypass Graft Surgery." *Journal of the American Medical Association* 275 (11): 841–6.

O'Leary, D. 1995. "Measurement and Accountability: Taking Careful Aim." *Joint Commission Journal on Quality Improvement* 21: 354–7.

Palmer, H. R. 1997. "Process-Based Measures of Quality: The Need for Detailed Clinical Data in Large Health Care Databases." *Annals of Internal Medicine* 127 (8, part 2): 733–8.

Rainwater, J. A., P. S. Romano, and D. M. Antonius. 1998. "The California Hospital Outcomes Project: How Useful Is California's Report Card for Quality Improvement?" *Joint Commission Journal on Quality Improvement* 24 (1): 31–9.

Rogers, G., and D. P. Smith. 1999. "Reporting Comparative Results from Hospital Patient Surveys." *International Journal for Quality in Health Care* 11 (3): 251–9.

Romano, P. S., J. A. Rainwater, and D. Antonius. 1999. "Grading the Graders: How Hospitals in California and New York Perceive and Interpret Their Report Cards." *Medical Care* 37 (3): 295–305.

Schneider, B. J. 1983. "Interactional Psychology and Organizational Behavior." In *Research in Organizational Behavior*, vol. 5, edited by L. L. Cummings and B. M. Staw, pp. 1–31. Greenwich, CT: JAI Press.

Schneider, E. C., and A. M. Epstein. 1996. "Influence of Cardiac-Surgery Performance Reports on Referral Practices and Access to Care." *New England Journal of Medicine* 335 (4): 251–6.

Schneider, E. C., V. Reihl, S. Courte-Weinecke, D. M. Eddy, and C. Sennett. 1999. "Enhancing Performance Measurement: NCQA's Road Map for a Health Information Framework." *Journal of the American Medical Association* 282 (12): 1184–90.

Schwartz, R. M., D. E. Gagnon, J. H. Muri, Q. R. Zhao, and R. Kellogg. 1999. "Administrative Data for Quality Improvement." *Pediatrics* 1031 (supplement): 291–301.

Soberman Ginsburg, L. 2001. "Total Quality Management in Health Care: A Goal Setting Approach." In *Advances in Health Care Management*, vol. 2, edited by J. D. Blair, M. D. Fottler, and G. F. Savage, pp. 265–90. Greenwich, CT: JAI Press.

Solberg, L. I., G. Mosser, and S. McDonald. 1997. "The Three Faces of Performance Measurement." *Joint Commission Journal on Quality Improvement* 23 (3): 135–47.

Taylor, M. S., C. D. Fisher, and D. R. Ilgen. 1984. "Individual's Reactions to Performance Feedback in Organizations: A Control Theory Perspective." In *Research in Personnel and Human Resources Management*, vol. 2, edited by K. Rowland, and G. Ferris, pp. 81–124. Greenwich, CT: JAI Press.

Thomas, J. B., and L. K. Trevin. 1993. "Information-Processing in Strategic Alliance Building—a Multiple-Case Approach." *Journal of Management Studies* 30 (5): 779–814.

Thomas, J. W. 1998. "Report Cards—Useful to Whom and for What." *Joint Commission Journal on Quality Improvement* 24 (1): 50–1.

Turpin, R. S., L. A. Darcy, R. Koss, C. McMahill, K. Meyne, D. Morton, J. Rodriguez, S. Schmalz, P. Schyre, and P. Smith, 1996. "A Model to Assess the Usefulness of Performance Indicators." *International Journal for Quality in Health Care* 8 (4): 321–9.

Tversky, A., and D. Kahneman. 1974. "Judgement under Uncertainty." *Science* 185: 1124–31.

Weiss, J., and C. Weiss. 1981. "Social Scientists and Decision-makers Look at the Usefulness of Mental Health Research." *American Psychologist* 36: 837–47.

Wicks, E. K., and J. A. Meyer. 1999. "Making Report Cards Work." *Health Affairs* 18 (2): 152–5.