A Task Environment Approach to Organizational Assessment

Alana Northrop, California State University, Fullerton
James L. Perry, University of California, Irvine

In recent years, interest has grown in the comprehensive and systematic measurement of organizational behavior. One reflection of this interest is the increased use of organizational assessment which is the process of measuring the effectiveness of an organization from a social systems perspective. An organizational assessment is grounded in conceptual models of organizations and is typically holistic, i.e., focused on "both the task-performance capabilities of the organization... and the human impact of the system on its individual members."1

When assessing the task-performance dimension of effectiveness, client evaluations and objective indicators are normally used to judge the extent to which an organization has met its goals.2 Input from officials in an organization's task environment is a third source of data that is traditionally ignored in organizational assessments, but one which is widely recognized as important.3

This study explores the use of effectiveness measures gathered from officials whose jobs require them to interact with federal organizations. By focusing on such inside respondents, a different view may be secured of federal agencies than that provided by private citizens.4 Respondents include members of Congress, state and local welfare officials, private contractors, and naval ship commanding officers. Being mainly government employees, they should have a relevant comparative perspective from which to evaluate other agencies' and workers' effectiveness, and they should be knowledgeable about service quality because of their stable interactions with the agencies. Of course, evaluations by both private citizens and "insiders" suffer from the limitations of perception and knowledge. But so far, evaluation studies have focused on private citizens with varying degrees of success.4 By using task environment respondents, we hope to provide another part of the evaluation picture, one previously overlooked, but one which can add to the external validity of organizational assessments.5

The concept of task environment not only directs use of a different set of respondents but also requires sensitivity to domains of activity as they vary across organizations.6 Prior organizational assessments have not explored how domain characteristics such as the location or type of service provided by the governmental unit may affect its perceived effectiveness. However, support exists for the view that the type of service provided by an agency can inherently affect its performance rating independently of the quality of the staff or of the organization's procedures.7 This is because relationships between goals of human service organizations and the means to achieve them are often not well understood. In contrast, relationships between goals of physical support organizations and the means to achieve them tend to be better understood.8 For example, methods for preventing soil erosion (a physical service) are based on years of scientific study while the process for making mortgage investments (a human service) where applicants' credit worthiness falls below market standards is much less well understood. This difference in the connection between means and ends would seem to have an important effect on the ability of an agency to be effective.

In addition, the hierarchical location of a governmental unit may also have an effect on the ability of an agency to be effective. In the federal government, headquarters offices are typically engaged in policy or rule-making.
making activity within politicized, complex, and ambiguous performance environments. However, the primary job of field offices is rule application. Therefore, the performance environment of field offices is ordinarily less politicized and more predictable. For instance, the routinization of field service provision for naval fleet support or social security benefits contrasts with the much less routine regulatory activities of GSA's transportation procurement offices. Hence, it seems reasonable that field offices have a greater potential for effective performance than do headquarters offices.11

It is also expected that location and type of service provided has smaller effects on evaluation of employees than on evaluation of agencies because one cannot completely hold employees responsible for any inherent problems in their work. Thus, employees are expected to get higher ratings than their agencies. As Michelson argues: "Government functionaries, though they work hard and long, accomplish little because they have inherently unproductive jobs."11 In explaining this "working bureaucrat, nonworking bureaucracy" thesis he writes: "This isn't sabotage. It's just plain bad management. It's not bad management at the functional level, assigning and monitoring doable tasks, although that may occur. It's bad management at the conceptual level, devising relevant tasks."13 This theme was reiterated in 1983 in a Wall Street Journal article by Clayton Christensen, a principal with the Boston Consulting Group and former White House Fellow.14 Mr. Christensen also argues that the Carter and Reagan administrations have been especially distrustful of bureaucrats, thereby restraining their productivity.

In sum, this study introduces the task environment as an alternative approach to organizational assessment. The task environment concept directs use not only of a different set of respondents than typically used in assessments, but also consideration of (1) whether the type of service provided, physical or human, affects performance ratings; (2) whether headquarters or field office status affects performance ratings; and (3) whether the focus of evaluation, employees or agencies, affects performance ratings. In the following pages, these questions are addressed using data gathered under both Carter's and Reagan's administrations.

Methods and Data

The data for this analysis were collected as part of a larger study of the effects of civil service reform in the federal government.15 Contractual and resource limitations confined the number of organizations participating in the study to seven. The following discussion describes the basis for selection of these organizations, the procedures for sampling respondents, and the construction of the evaluation indexes.

Site Selection

Four federal departments were chosen, and a fifth department was added in case one of the departments terminated its voluntary participation for unforeseen reasons. Through negotiations with the five departments, seven sites were chosen. The sample consisted of the following organizational units: the Transportation and Public Utilities Service (TPUS) of the General Services Administration, Washington, D.C. (headquarters/physical services); the Naval Ship Weapon Systems Engineering Station (NSWSES) in Port Hueneme, California (field/physical services); the NASA-Ames Research Center in Moffett Field, California (field/physical services); the Southern California network offices of the Social Security Administration (SSA) (field/human services); the national office of the Farmers Home Administration (FmHA) (headquarters/human services); and the national and California state offices of the Soil Conservation Service (SCS) (headquarters, field/physical services). All agencies were voluntary participants in the study.

Because sites were chosen expressly to test hypotheses about services and location, the primary site selection objective was internal validity. The extent to which sites were representative of all federal organizations was less of a concern because generalization to all federal organizations was not an immediate objective. The sites, however, appear to represent a range of federal organization characteristics, including military and civilian missions, of varying ages and sizes. While it is doubtful that a small sample could be representative of all federal agencies, this sample meets the explicit requirements of the research design and yet reflects the diversity found throughout the federal government.

Sample and Sampling Procedures

Members of each organizational unit's task environment—i.e., other federal officials, state and local officials, and oversight bodies—were surveyed. Those surveyed were either selected from archival information or were nominated by managers and staff from the focal agencies. The criterion for selection was that the individual interact with the focal agency in the performance of its tasks or mission (hence the "task environment" designation). We sought out individuals from each program unit to identify potential respondents to assure as representative a sample as possible.

Biases are possible when managers nominate respondents to evaluate their agencies. A manager could knowingly stack the deck with favorable informants or withhold identification of potentially critical information. Nominations could be biased in other ways, for example, by a manager's inability to remember potential
respondents. These biases would not necessarily affect the relative evaluation across agencies, but nevertheless efforts were made to minimize them. Archives were used for random sample selection whenever feasible, and with informants, emphasis was given to the non-threatening purposes for which the names of potential respondents were being solicited.

Potential respondents in positions with few incumbents were sampled disproportionately. Those in positions that had many incumbents (e.g., naval ship captains and contractors) were randomly sampled. Representative external roles that were included in each organizational unit's task environment are listed below.

TPUS: public utilities, state and local utilities agencies, General Accounting Office, transportation-related divisions and financial branches of federal departments and programs.

NSWSES: naval shipyards, weapons stations and other naval facilities, Department of Defense, Naval Sea Systems Command, contractors, ships' commanding officers.

Ames: contractors, NASA facilities, federal agency liaisons, military installations, high schools, colleges and universities, aircraft associations.

SSA: members of Congress, state and local welfare agencies.

FmHA: housing and investment associations, professional groups, federal departments and agencies, other Department of Agriculture agencies, public interest groups, health associations, water and waste associations.

SCS: professional groups and societies, environmental associations, colleges and universities, public interest groups, equal opportunity associations, federal, state and local agencies.

The sample of task environment members was developed in April and May of 1980 for the June 1980 survey administration, and it was revised before the June 1981 administration. Given the potential difficulty of resurveying the same individuals for the second wave, the probable attrition using a panel design, and the changes in the task environment, a modified panel sampling was used. The same organizational positions were sampled in each wave, rather than specific individuals. This assured comparability between the 1980 and 1981 samples, after adjusting for any structural changes in the task environments.

The survey was conducted by mail and accompanied by a letter of explanation. A follow-up letter and new survey were mailed to non-respondents about three weeks after the initial mailing. Table 1 summarizes the size of the task environment samples and the number of completed questionnaires.

### Construction of Evaluation Indexes

Ten indicators of employee and agency performance were factor analyzed. Two factors were obtained: (1) a rating of employee integrity and work, and (2) a rating of agency effectiveness. A score for each respondent on the two factors was obtained by averaging the category responses on the individual questions. Each index has a standardized item alpha of at least .80, which suggests their high scalability. The questions used to build each index are listed in Table 2.

### Findings

**Do Location and Type of Services Affect the Evaluation of Federal Agencies?**

We hypothesized that field offices and units which provide physical support services would be judged more

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Sample Size</th>
<th>Returned No.</th>
<th>% Ret.</th>
<th>Sample Size</th>
<th>Returned No.</th>
<th>% Ret.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farmers Home Administration</td>
<td>71</td>
<td>33</td>
<td>46</td>
<td>133</td>
<td>65</td>
<td>49</td>
</tr>
<tr>
<td>Soil Conservation Service Headquarters</td>
<td>8a</td>
<td>5</td>
<td>63</td>
<td>230</td>
<td>110</td>
<td>48</td>
</tr>
<tr>
<td>Soil Conservation Service California Field Office</td>
<td>38</td>
<td>28</td>
<td>74</td>
<td>63</td>
<td>45</td>
<td>71</td>
</tr>
<tr>
<td>NASA-Ames Research Center</td>
<td>236</td>
<td>184</td>
<td>80</td>
<td>182</td>
<td>130</td>
<td>71</td>
</tr>
<tr>
<td>Naval Ship Weapon Systems Engineering Station</td>
<td>130</td>
<td>94</td>
<td>72</td>
<td>223</td>
<td>119</td>
<td>53</td>
</tr>
<tr>
<td>Social Security Administration, Southern California Network</td>
<td>12</td>
<td>10</td>
<td>83</td>
<td>11</td>
<td>9</td>
<td>82</td>
</tr>
<tr>
<td>Transportation and Public Utilities Service</td>
<td>178</td>
<td>80</td>
<td>45</td>
<td>129</td>
<td>79</td>
<td>61</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>673</strong></td>
<td><strong>434</strong></td>
<td><strong>64%</strong></td>
<td><strong>971</strong></td>
<td><strong>557</strong></td>
<td><strong>57%</strong></td>
</tr>
</tbody>
</table>

aThe first wave of questionnaires was distributed only to a few external contacts for this agency because of the agency's apprehension about the consequences of such a survey. Since this agency's participation was voluntary, we chose not to press for a larger sample and potentially risk discontinuation of the long-term relationship. By the time the second year's survey was administered the agency's concerns were significantly reduced because the first administration had no negative effects and good working relationships had been developed with the agency. The sample size for the second administration, therefore, was increased considerably.
Again, on a scale of very negative (1) to very positive (7), agencies that provide physical services received a 3.9 mean rating in 1980 and a 4.1 mean rating in 1981 (Table 3). But agencies that provide human services maintain high standards of conduct. Employees in this organization have the skills necessary to do their jobs. Employees in this organization work hard. Employees in this organization have enough work to keep them busy.

Rating of Agency Effectiveness*  Alpha = .80

Overall, this organization is effective in accomplishing its objectives. In this organization, it is often unclear who has the formal authority to make a decision. It takes too long to get decisions made in this organization. The management in this organization is flexible enough to make changes when necessary.

*Response categories on a seven-point scale ranged from strongly disagree (1) to strongly agree (7).

The management in this organization is flexible enough to make changes when necessary. Hence, two different statistical tests. Yule's Q and t-test upon normal conventions given that Yule's Q varies from 0 (no relationship) to ± 1 (a perfect relationship). Thus, two different statistical tests, Yule's Q and t-test for difference in means, show that field offices receive significantly higher ratings than headquarters offices under two different administrations. And, looking at the ratings of the seven units individually confirms that the field/headquarters findings are not biased by the disproportionate sample size of any one unit; the individual headquarters ratings are lower than those of the field offices.

Considering the type of services provided by an agency, we find that they have a slightly weaker effect on an agency's effectiveness ratings than does location (Table 4). Again, on a scale of very negative (1) to very positive (7), agencies that provide human services received a 3.9 mean rating in 1980 and a 4.1 mean rating in 1981 (Table 3). But agencies that provide physical services received the higher mean rating of 4.7 in both 1980 and 1981 (Table 3). These differences in mean ratings by service type are statistically significant, and the magnitude of Yule's Q indicates a "moderate" to "strong" relationship between types of service provided and effectiveness ratings (Table 4). In addition, the different ratings for agencies that provide human services versus physical services are not due to the disproportionate sample size of any one unit.

In sum, field offices and units which provide physical services have a higher likelihood of being judged more effective than headquarters and units which provide human services. We believe that location affects performance rating because field offices have more routine and less politicized jobs, making their goals easier to accomplish. We also believe that type of services provided affects performance ratings because governmental agencies that provide physical services tend to have clearer means for achieving their goals than do agencies that provide human services, thereby making their goals easier to attain.

Do Location and Type of Services Affect the Evaluation of Federal Employees?

Respondents were also asked to assess the integrity and work effectiveness of federal employees (see Table 2 for questions). Again we wanted to look at the impact of location and type of services but this time on the evaluations of employees. We believed that location and type of services affected the evaluation of federal agencies, a spillover or tarnished halo effect would carry over to the federal employees. In addition, we expected that location and type of services provided would have a weaker effect on the evaluation of employees than on the evaluation of agencies. Specifically, we felt that employees can be viewed as distinct from the agency, and there could, therefore, be a rational tendency not to blame employees for the inherent administrative problems of their work.

On the scale of very negative (1) to very positive (7), our respondents gave headquarters employees a 4.8 mean rating in 1980 and a 4.9 mean rating in 1981 (Table 3). In contrast, employees in field offices received the significantly higher mean rating of 5.4 in both 1980 and 1981 (Table 3). These different ratings for employees in headquarters versus field offices are not due to any disproportionate sample size of unit respondents. Complimentarily, our additional statistical tests indicate that functional location has a statistically significant and "moderate" effect on the evaluation of federal employees (Table 4). We should also note, though, that location's moderate effect on the rating of employees is smaller than its effect on the rating of agencies (Table 4).

Using the same 1 to 7 scale, employees who are involved in human services received a 4.7 mean rating in 1980 and a 4.9 mean rating in 1981 (Table 3). As anticipated, employees who are involved in physical support services received the higher average ratings of 5.3 in 1980 and 5.2 in 1981 (Table 3). The effect of type of services provided on respondents' evaluations of federal employees is statistically significant and "moderate" in strength (Table 4). These findings again are not due to the disproportionate sample size of any one unit's respondents. In addition, type of services provided has a
A TASK ENVIRONMENT APPROACH TO ORGANIZATIONAL ASSESSMENT

TABLE 3
Means on Evaluation Indexes by Location and Service*

<table>
<thead>
<tr>
<th>Location and Service</th>
<th>Rating of Employeesb</th>
<th>Rating of Agenciesb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Headquarters/human</td>
<td>4.7</td>
<td>4.9</td>
</tr>
<tr>
<td>(N=115)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field/human</td>
<td>4.7</td>
<td>4.7</td>
</tr>
<tr>
<td>(N=43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Headquarters/physical</td>
<td>4.7</td>
<td>4.9</td>
</tr>
<tr>
<td>(N=9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field/physical</td>
<td>5.4</td>
<td>5.4</td>
</tr>
<tr>
<td>(N=185)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Respondents</td>
<td>5.2</td>
<td>5.2</td>
</tr>
<tr>
<td>(N=423)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Medians for location and service parallel the mean findings, so only one statistic is used for parsimony of data presentation.

bIndex ranges from (1) strongly disagree (very negative rating) to (7) strongly agree (very positive rating).

*p < .05, one-tailed t-test for difference in means within years.

smaller effect on the evaluation of employees than it does on the evaluation of the agencies in which they work (Table 4).

To summarize, employees in field offices and who are involved in physical service delivery receive higher evaluations in terms of integrity and quality of work than do employees in headquarters and those who are involved in human service delivery. Consequently, the location and type of services provided by an agency not only affect its image but also that of its employees. But being a headquarters office or providing human services has a greater effect on the evaluation of an agency than it does on the evaluation of an agency's employees.

Are Employees Rated Higher Than Their Agencies?

The third issue we wanted to address of importance to organizational assessment was whether evaluation of the employees of an agency differs from evaluation of their agency. We had hypothesized that it was quite conceivable that one would get different ratings for an agency than for its employees because quality of staff can be distinguished from the quality of the organization's procedures.

A large overlap was found in ratings of the employees and the agencies (1980 Q = .83*, 1981 Q = .76*). However, employees were consistently rated higher than their respective agencies (Table 3). This is true for each of the seven units individually as well as grouped by location or type of services (Table 3). Using our 1 to 7 scale, employees received a mean rating of 5.2 in both years while the agencies received the statistically significant lower mean rating of 4.6. The median scores as well as the mean scores were lower for agencies than for employees in both years.

Summary and Conclusions

The concept of task environment not only directed us to use a different set of respondents but also to be sensitive to domains of activity as they vary across organizations. As a result, we hypothesized that field offices and physical support organizations had an important edge in performance achievement over headquarters and human service organizations. Our data strongly support both of these hypotheses. Field offices and physical services organizations received much higher evaluations than did headquarters and human service
TABLE 4
Yule’s Qs Between Evaluation Indexes and Location and Service Measures

<table>
<thead>
<tr>
<th>Evaluation Index</th>
<th>Rating of Employees</th>
<th>Rating of Agencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headquarters/field</td>
<td>-.47*</td>
<td>-.47*</td>
</tr>
<tr>
<td>Human services/physical services</td>
<td>-.39*</td>
<td>-.28*</td>
</tr>
</tbody>
</table>

*aAll indexes ranged from 1 to 7 and were dichotomized by combining the positive ratings of 5, 6, and 7 and the negative and undecided ratings of 1, 2, 3, and 4. Variables were dichotomized because there were very few high or low scores on the 7-point scales, leading to the problem of insufficient units in table cells to be analyzed. Yule’s Q is an appropriate measure of association because the independent variables are measured at the nominal level and all variables are dichotomous. The data adhered to the 30:70 rule of marginal splits.

*p < .05.

organizations. Moreover, the findings remained the same even though one survey took place under the Carter administration and another under the Reagan administration. The stability of these findings suggests that systemic factors such as location and type of services provided may be important explanations of differences in performance ratings.

Furthermore, we found that the effects of location and type of services carry over, although not as strongly, to evaluations of federal employees. Employees in field offices and those who were involved in physical services received higher ratings than did employees in headquarters and those who were involved in human services.

Finally, we found that employees were consistently rated higher than their agencies. We think it is important that employees not only received different ratings from their agencies but also that they received higher ratings. And we believe that the higher employee ratings reflect problems—real and perceived—with the management of federal agencies. Such management problems appeared to be present under both the Reagan and Carter administrations.

In a speculative vein, these results have implications for the reorganization of the federal bureaucracy. Both Jimmy Carter and Ronald Reagan rode to victory as Washington outsiders. Carter viewed the problem of the federal bureaucracy as one of personnel and consequently focused on improving the quality of the workforce. These findings suggest that the system may be a more serious problem than the motivations or abilities of its personnel. Instead of merit pay and other employee oriented incentives, Congress and the present administration should consider decentralizing the federal bureaucracy. By expanding and creating new field offices, the image of the federal bureaucracy might improve, at least among government insiders and perhaps among a wider group of the citizenry. And to the extent that perceived effectiveness relates to effectiveness, the creation of new field offices could lead to payoffs in effectiveness. But please note we are suggesting decentralization through the creation of federal field offices and not decentralization through transferring federal programs to the states.

Our findings also suggest that cutting human services, a Reagan strategy, may not deal with the problem of human services agencies. Human services organizations appear to be perceived as less effective and thus are open to criticism due to the inherent nature of their tasks rather than to their present size, structure, or personnel. However, to the extent that the Reagan administration strategy leads to an aggregate shift toward physical services, the image of the federal government might be improved among government insiders and perhaps, in turn, among a wider group of the citizenry.

Future research will need to assess how results from this task environment approach may be used either in conjunction with or as a substitute for other sources of evaluation information, such as private citizen evaluations and objective indicators. One possible area for exploration is to see whether private citizens or objective indicators discriminate on the basis of location and type of services as did our task environment respondents. It is likely that this sample of federal agencies is not indicative of all federal agencies. Other studies will have to confirm whether our findings on location and type of services performed are unique to our agencies or are factors that are generalizable to all federal agencies and, for that matter, other similarly structured government bureaucracies.

Notes

2. Ibid., 6.
3. For a summary of many client-based evaluations, see Charles Goodsell, The Case for Bureaucracy (Chatham, N.J.: Chatham House Publishers, Inc., 1982), Chapter 2. Illustrative of these client surveys are Charles Goodsell, "Client Evaluation of Three"
A TASK ENVIRONMENT APPROACH TO ORGANIZATIONAL ASSESSMENT


4. The task environment concept was first introduced by William Dill to identify the parts of the environment which are relevant to goal setting and attainment. It has been used extensively in the organization theory literature. See William R. Dill, "Environment as an Influence on Managerial Autonomy," *Administrative Science Quarterly* 2 (March 1958), 409-443.


13. Ibid., 176.


16. A retest of the June 1980 survey using a sample of the original respondents, was conducted to determine the reliability of the initial questionnaire items. Twenty-one individuals were contacted and consented to complete the survey a second time, two weeks after the first survey. Comparison of an individual's responses at this short interval permits an assessment of the probable variance in responses due to measurement error alone. The appropriate statistic for assessing retest reliability is the Pearson correlation coefficient. A mean coefficient of .70 was achieved, indicating a high reliability for the survey items. For further discussion of this procedure, see Jum Nunnally, *Psychometric Theory* (New York: McGraw-Hill, 1977).
