REPORTING FOR EFFECTIVE DECISION MAKING

CSE Technical Report 298

Joan L. Herman
Lynn Winters
Shari Golan

UCLA Center for Research on Evaluation,
Standards, and Student Testing

March, 1989

The research reported herein was conducted with partial support from the U.S. Department of Education, Office of Educational Research and Improvement, pursuant to Grant No. G0086-003. However, the opinions expressed do not necessarily reflect the position or policy of this agency and no official endorsement by this agency should be inferred.

This report was presented at the 1989 Annual Meeting of the American Educational Research Asociation, San Francisco.

Please address inquiries to: CSE Dissemination Office, UCLA Graduate School of Education, 405 Hilgard Avenue, Los Angeles, California, 90024-1521

Drowning in Data

Educators are drowning in a sea of data. The amount of information collected by schools and districts for planning rivals the drops of water replenishing the oceans. While much of the data purportedly is useful for a variety of decision making purposes, the sheer volume of information available for school planning may well have overpowered its utility. For example, there were 66 different kinds of data presented in the district reports submitted in the Division H Awards Competition in 1988. While no one district reported all 66 pieces of information, on the average each report contained 16 types of descriptive information and 4 types of test data. And last year's reports did not even purport to present survey results which are becoming increasingly important in school and district level decision making. At best these 20 plus pieces of data represent an important resource for decision making and action; at worst, they form a tidal wave of disorganized messages. How can the potential power of this vast amount of information be harnessed? We as school district evaluators, R&E directors, and researchers have to channel it and moderate its flow so that those less sophisticated in data management and interpretation, school boards, district administrators, principals, and teachers, find it useful for making decisions about their schools.

Building a Life Raft

Other sectors have been quick to utilize the power of computer technology and its advanced presentation graphics to create management information systems to support decision making. We have been working on a project at the UCLA Center for Research on Evaluation, Standards, and Student Testing (CRESST) to explore the possibilities for school decision making. The project, the Multilevel Evaluation Systems Project funded by OERI, is investigating the feasibility of developing comprehensive information systems that might serve the planning and policy needs of school-based educators, district administrators, and school boards; one intent is to develop a set of design specifications for such systems. The current phase of the project is being conducted in three stages. Stage 1 was a multidisciplinary literature review to glean guiding principles for the design of school-based management information systems and presentation graphics. Stage 2 featured a review of existing district reporting practices. In Stage 3, which is in progress, we are conducting interviews with the potential user groups, school board members, district administrators, and principals to get a better picture of how they typically process and use information on school quality.

The First Plank: A Review of the Literature

Our literature review covered a number of topics: teacher planning, evaluation utilization, information representation and decision making, and computers and user-friendliness. We also incorporated research on management information systems (MIS), cognitive psychology, and marketing. The direction of the review was shaped by Lucas' (1975) observation that a major reason information systems often fail is that designers concentrate on technical aspects of system design and implementation, overlooking the potential user's organizational behavior and needs. Thus, we concentrated on findings related to the decision making behavior of school practitioners and how information could be organized and presented to meet their needs. This section highlights some of the major findings from the literature related to information representation and the use of data in decision making.

Evaluation Utilization

We are interested in how school planners use evaluation results to support discrete decisions, decisions related to school policy and improvement. While we are optimistic about the potential impact of data on decision making, we are aware that most school decisions and actions are based upon "ordinary knowledge," derived from a teacher's, principal's, or administrator's practical experience. This ordinary knowledge is widely shared among educators, is sensitive to the context of the schools, and is comprehensive (Lindbolm & Cohen, 1979). By contrast, school evaluations, as they are currently conducted, produce data that are context independent and selective rather than comprehensive. Thus, data contained in an information system will never supplant ordinary knowledge; rather, systems need to be designed to complement this existing base.

Research in evaluation utilization identifies some of the factors that will influence whether formal data will serve this complementary purpose. Among the time-worn variables that appear to be determinants of utility are: (a) quality and credibility of the data provided, (b) relevance to user needs, (c) quality and ease of communication, (d) timeliness, (e) user commitment to the evaluation process, (f) a political climate conducive to use, and (g) personal characteristics of the user (Cousins & Leithwood, 1986). Clearly, utility depends on more than the data and may in fact be at least equally dependent on the nature of the evaluation process and the social-political interactions it generates.

Technical quality is a characteristic we all hold dear—and well we should—but from a utility perspective. It is not the technical quality itself that is paramount, but the perceived quality of the data that is most important, as well as the data's sensitivity to users' needs in questions/issues addressed, its timeliness and accessibility of reported results, and its consistency with users' perceptions and beliefs. For example, it is of interest to note that studies related to the credibility of evaluations suggested that credibility hung on whether the findings were consistent or inconsistent with user expectations. Incongruent findings tended to be ignored (Cousins & Leithwood, 1986). On the other hand, data which challenged the status quo or were critical of current practice were very often utilized (Weiss & Bucuvalas, 1980), possibly because user expectations did not support the status quo. With user expectations providing such an important influence, effective information systems and data presentation techniques will need to incorporate a method for allowing users to make their "expectations" or hypotheses explicit and encourage them to review the data against these preconceptions.

Consistent with sensitivity to users' needs, it bears emphasizing that a useful system or a useful report is easy to use. An ideal system would generate reports using nontechnical language and the data in such reports would require little or no transformation for use by technically unsophisticated decision makers (Weiss & Bucuvalas, 1980). For a computer-based information system, the ease of use dimension probably implies that the system should be menu driven, have simple file transfer capabilities among school sites and between sites and a central computer, and have simple strategies for data input. (Cooley & Bickel, 1986). The system must be capable of anticipating and answering queries that users deem relevant, which means that an automatic process for local "customization" is essential . As for the context within which a truly useful evaluation data management system would operate, it almost goes without saying that the decision making and program planning functions should be separated from the accountability functions.

Our review of the evaluation utilization literature emphasizes the importance of users' perspectives and needs as they actually exist over hypothetical rational visions of what those needs are or should be. This conclusion prompted us to consider how we might introduce "user context" into our presentation of data and

maximize user engagement. For example, how can we help users make their expectations for certain results or outcomes explicit so that the data they are reviewing will be used? Some possible data presentation strategies that might enhance utilization are:

1. Prompt users to ask a question of themselves related to possible findings before viewing the data.

2. Present data tables and graphs with headings in the form of questions, such as, "What do you see happening to reading scores over time?," or, "Is there a relationship between the students' attitudes toward math and their scores?"

Information Representation and Decision Making

Our review of the evaluation utilization literature gave us a new perspective for approaching the literature on information representation. At first blush, it would appear that guidelines for presenting data would be objective, specific, and simply a recapitulation of "rules" for making charts, graphs, and tables. Because we found that the user's expectations, beliefs, and experience determine which data are attended to and how they are used, we decided to add to our review of data presentation formats a review of how people process numerical information. A sampling of our findings are summarized below.

Human Cognition and Data Display

The way numerical data are presented affects a user's comprehension and use of those data (MacDonald-Ross, 1977). Washburne (1927) found that both the visual and logical arrangement of data have an important effect upon information comprehension. Thus an understanding of the cognitive preferences and limitations of users as they relate to display formats is an important prerequisite to generating an effective data display.

While no catalog of user's cognitive preferences for processing information can be completely accurate or complete, some principles are emerging from the steadily growing body of information that we are accumulating about "how people think." Bettman and Kakkar (1977) and Ghani (1981) found that subjects' strategies for using information were based on the information that was most easily processed. An example of this is the findings of Slovic and Lichtenstein (1968) that people resist making even simple transformations of information. People tended to discount or ignore information requiring inference from an explicit display. Instead, they used only the information displayed and only in the form in which it was displayed. In keeping with the "keep it simple" maxim, it was found that users resisted making "cognitive shifts" when reviewing information. They tended to resist changes in information representation styles (Ghani, 1981). Thus, one cognitive preference of information users is that the data be presented in a simple fashion, requiring no transformation, and with consistent presentation formats.

A second preference concerns the balance between numeric and adjectival information. Scammon (1978) reported that subjects receiving adjectival information had more accurate aided recall and comprehended the overall meaning of data better than when information was represented numerically. His explanation was that adjectival descriptions are often inherently evaluative (e.g, "above average," "good," "poor") which made evaluation tasks easier. On the other hand, when presented with numerical information, subjects asked more questions and wanted more information. Further, numeric data is perceived to be more precise and accurate than adjectival information and allows for easier comparisons (Cherry,

1966). These seemingly conflicting findings raise questions: What is the tradeoff between credibility and comprehensibility? Can we have both?

Human cognitive limitations also suggest some guidelines for data presentation. We all know about the seven-plus or minus-two information retention hypothesis (and often assume that the nontechnical audience has a retention level at the minus-two end of the continuum). We also learned at one time (yet don't always keep in mind) how people make errors when interpreting statistics. The most common of these errors, as cataloged by Remus and Kottemann (1986), are: (a) users assume trends where in fact only random variation is occurring; (b) too much weight is given to findings based on small or unrepresentative samples (perhaps a result of the above discovery that numerical information appears more precise than adjectival); (c) users make incorrect cause-effect inferences from correlations or even independent events, and most importantly, (d) when faced with several sources of uncertainty (aren't we always careful to report findings with the subjunctive "may" or "might"), users simplify decision contexts by ignoring or discounting some of that uncertainty.

By taking into account the user's cognitive preferences and limitations when processing data, we might posit the following guidelines for making information more useful:

1. Stick to one format for tables or graphs.

2. Accompany data with explanations that will counter common statistical misinterpretations.

3. Select a presentation format that will allow the user to organize and conceptualize it in terms of a small number of chunks or categories.

Selecting and Generating Effective Display Formats

It will come as no surprise that we found no evidence that one graphic format is universally superior; however, each format has some advantages over others in specific situations. When reviewing our summary in this section, the reader should keep in mind that the literature related to effective data formats is fairly old and does not include some of the newer types of graphical displays (including dynamic displays) made possible by computers.

Tables provide an important vehicle for summarizing data because they are compact. However, they require more processing on the part of the user (MacDonald-Ross, 1977) which then engages the tradeoff between compactness and ease of understanding.

Vernon (1950) and Washburne (1927) found that particular types of graphs had specific strengths. Bar graphs are best for static comparisons, line graphs for dynamic, and pictographs for providing striking general impressions (which is why they find their way into *USA Today*). Schutz (1961) found that a single graph with multiple lines was better at representing multiple trends than multiple graphs with single lines. Croxton and Stein (1958) found bar charts easier to comprehend than stacked bars. Clearly these simple precepts about format selection are well known and were established a long time ago. It is our feeling that more research with the users of these graphic formats needs to be done to update format selection guidelines. We have begun to collect information about user format preferences and their utility. Our initial efforts will be described in the last section of this paper.

Guidelines for designing effective data displays abound. We have summarized these guidelines in Table 1. The plethora of rules presented in this

table could probably be best summarized by the following data display commandments:

1. Keep the data simple. Be sure what you are reporting has a straightforward interpretation that requires no transformations. For example, item-response-generated scores, standard deviation scores, and graphs summarizing "qualitative" information (such as percent of goals mastered when mastery means getting 75% correct on five out of seven objectives) should be reported with care.

2. Limit the representation formats. Stick to a few varieties of tables or graphs. Consistent formats allow easier comparisons across displays. Pause before selecting three dimensional or stacked bar displays.

3. Descriptive adjectives make numbers more meaningful. Numbers may lead users to believe that there are more distinctive differences between individuals or groups than really exist; adjectives accompanying the numerical categories (such as percentile ranks 1-25 being designated as "below average") may help counter this tendency.

4. Explain statistics clearly. Warn users when mini-interpretations are likely to occur and point out the limitations of the data. Distinguish between statistical and practical significance when presenting the results of research and evaluation studies. Remember that test scores can be especially difficult to understand. People often confuse percent correct with percentile and have difficulty understanding scale and standard scores.

The Second Plank: A Review of Current Reports

Although the literature review provides precepts, of real interest to us is what is occurring in practice. How do school districts present information to the public? What is the best of current practice? How does this coincide with what the literature suggests ought to be done to make information more useful? What promising practices not alluded to in the literature are occurring?

In an attempt to focus on the best of current practice, the review analyzed reports submitted to the American Educational Research Association Division H Award Competition in the category, "District Profiles." These reports were generally written for boards of education, the general public, and in some cases where detailed school profiles were provided, for school site administrators. Our review provides only a tentative picture of current practice. Limited to single reports which districts chose to submit as exemplars, it is clearly insensitive to the fact some districts produce several reports, each dealing with separate aspects of school quality. Further, because Division H membership is drawn heavily from large city school districts, our sample underrepresents the efforts of medium or small districts.

What Kinds of Data Were Reported?

It will come as no surprise that all reports featured test results, and all but one presented scores from nationally normed tests. Only one district of those analyzed used a locally developed test as its only measure of district performance. Most districts also reported a variety of test results, ranging from two to nine different types of tests. The types of tests most commonly reported in addition to nationally normed tests were state assessments, district competencies, advanced placement, and the Scholastic Aptitude Test.

Student background demographics also were featured in these reports. All reports contained information about enrollments by grade and ethnicity and all, save one, published average daily attendance rates. Over half reviewed disclosed socioeconomic status reported in various ways, from students receiving free and reduced priced meals to parent occupation to levels of parent education.

Attention to school completion, expressed in terms of drop-out rates or percentages of students receiving diplomas, suspension rates, enrollment in specific courses, and post-secondary plans were included in many reports. Percentage of students enrolled in special programs such as bilingual, gifted or special education were commonly reported. Interestingly, given the attention of both School Improvement and Chapter VII Programs to student attitude information, we found no survey data in the district profiles. Most likely this is because these reports are generated for a wide, rather vague audience. Reports generated to meet school-level planning needs would be more apt to have survey information summarized.

Information about teachers and staff was provided. Class loads, pupil/staff ratios, and description of staff ethnicity, experience and training were commonly included in the reports.

How Was This Information Presented?

Data were displayed via tables of descriptive statistics and brief phrases. We found fewer graphs than anticipated. The most common statistic reported was a "test score average." In some cases the average was the median percentile rank, in others it was the mean, and in still others it was not clear which "average" was being used. Reports were also silent as to whether the ranks described represented individual pupil ranks or building ranks, although from their magnitude it appeared as though individual ranks were used. Comparisons of these averages to some desired standard such as "performing above the norm" were infrequent.

There was a great amount of space devoted to longitudinal data, but these data were generally lacking visual support. Some reports did use frequency polygons to depict trends, but many did not.

How Well Did Reports Adhere to Data Presentation Guidelines?

Perhaps the most frequently violated guideline was the "keep it simple" precept. Individual pages were chock full of information in a variety of formats, mostly tabular. In some cases, one page would contain a pie chart, two or three tables, and a bar chart. Some reports featured tables, box and whisker plots, and frequency polygons along with brief descriptive information on a page. Of course, there was a reason for data overload; the intended audience for these reports—mainly board members, oftentimes realtors—are busy people who want the most information possible in the smallest amount of space. This raises an issue: Does "keep it simple" mean presenting data argument one at a time as argued in the literature, or has this maxim evolved to mean "keep it simply to one page"?

A second violation of the "keep it simple" rule was the use of complex transformed scores without accompanying explanation. Often "composite" scores were reported without an indication of how they were derived, a situation that could be masking methodological improprieties. Item response theory scores were presented without interpretation; since these are often seen as standard deviation scores by the uninitiated, this omission is misleading.

There was an overall lack of explanatory material accompanying the displays. Often explanations of how scores were derived and how data was collected and interpreted were lengthy and informative, but set apart from the display in either

an introduction or appendix. Even when these were read by users, the limits of short-term memory (even for the seven-plus-two-chunk users) would suggest that the explanations had only a slight chance of being used. Typically, much technical jargon was embedded in the explanatory material and the prose was less than lively. Although it's easy to take a cheap shot at the writing style in technical reports, the impact of writing style on utility needs to be considered.

What Did We Learn from Current Reports?

Given the political context in which schools must operate and the undefined audiences for which most school and district reports are generated, following ideal practice in data display will be difficult. However, our small review does suggest that technically correct reports are a necessary but not sufficient condition to usefulness. We may wish to present fewer data, in simpler form but with more context, both in the form of non-technical explanation and questions to engage the reader with the data. There was little attempt to explore relationships among different pieces of data in the reports; yet, given the kinds of data presented, it was clear that some information was selected because of implied relationships among data such as class size, ethnicity, language background, and staff experience. Given that users tend to make errors in interpreting statistics, much less teasing out cause-effect relationships, we might serve ourselves better to explicitly investigate these relationships in our reports and to lay the erroneous assumptions to rest once and for all.

A Guidance System for Keeping Afloat: What Users Want

We believe that data can tell important stories. We don't always know, however, how people who use data construct these stories. Based on our review of evaluation utilization literature, the focus of our study shifted from an exclusive reliance on "scientific" guidelines for data reporting to a broader perspective which includes the interaction between the user's expectations and ordinary knowledge and presentation formats as a variable in information use. We felt that neither the literature on data representation nor evaluation utilization provided sufficient information about the user's perceptions. What was missing were guidelines for reporting that would help us present data to users in such a way that it would be credible, understandable, and easily incorporated into or compatible with their ordinary knowledge. Much as an understanding of reading rests upon a knowledge of a person's metacognitive processes, an understanding of data utilization is founded upon knowledge of how people make sense of empirical information in formal reports.

Thus, we designed the third stage of this project to study how perceptions of quality are formed, what information currently available in district reports is used, and what strategies school planners use to make sense of the data we prepare for them in formal reports. This study of how users process empirical information began last fall with the identification of a voluntary sample of six school districts selected to reflect different socioeconomic levels, ethnic compositions and geographical settings (urban, suburban and rural). Because we want to extend our findings to school districts that don't have extensive research and evaluation expertise, the majority of the sample consists of districts in the small (under 5,000 a.d.a.) to medium (5,000-20,000 a.d.a.) range. Thus far the participating districts include: one large urban district with an a.d.a. of 66,000; a mid-sized rural district with an a.d.a. of 15,000; and four mid-sized districts with a.d.a. ranging from 9,000 to 20,000. Within each district we plan to interview three school board members and three principals, two high school and one elementary. The data from each of these groups will be separately analyzed, with separate reports on implications for school-based planning and for information systems for policy use.

Administrators from the selected districts and CSE staff are working together to conduct clinical interviews using a common interview protocol. Collaboration with district personnel in these interviews has two important benefits: (a) in-house staff can act as key informants who have a detailed knowledge of district context, and (b) the information gathered will be of direct use in the participating districts.

We have developed a draft interview protocol, which appears in Table 3, designed to find out what information school planners use for decision making, how exactly the information is used, how planners "read" and interpret actual reports, and what necessary information is missing from current reports. To date, CSE and the district administrators collaborating in the Multilevel project have conducted three focus groups at the Annual Meeting of the California School Boards Association (CSBA) and six clinical interviews with administrators or board members in four school districts.

The interviews are being conducted according to guidelines for clinical interviewing established by Mary Lee Smith. They essentially consist of three parts: (a) the actual interview which is taped, (b) field notes summarizing both the interview and context necessary to interpret the tape, and (c) a protocol analysis of the tape itself. The interview is a 30-minute session with an informant, either a school board members or principal, structured to elicit answers to the questions outlined in the interview protocol. Immediately after the interview, "field notes" are jotted down. These notes include attitude of the interviewee, rapport, a summary of the "answers" to the five interview questions, and any other information that would help "explain" the respondent's concerns and remarks.

Once interviews are completed in the six sample districts, we will transcribe and attempt to identify consistent trends relating to how school planners process information currently available in formal reports and what information is most useful for decision making.

Preliminary Findings

Preliminary findings raise some important questions for data reporting. From our school board focus groups, we found that there was wide disparity in background and experience in working with tables, graphs, and school district technical reports as well as in sophistication in interpreting data. When presented with what we considered both typical and novel data presentation formats, responses ranged from disinterest in reading tables and graphs (e.g., "Just tell me how many students are on grade level.") to the rare few who waxed enthusiastic about the opportunities box and whisker plots provided for examining different segments of the student score distribution. The disparity in responses to different presentation formats raises this issue: How do we identify "best" or "most effective" reporting strategies when faced with such differences in user preference?

We have found one consistency among users, despite their level of sophistication: All board members and even principals interviewed were trying to answer a general, implicit question as they read each report: "How well are students doing compared to some standard I hold for good student performance?" While the specific standards varied somewhat from district to district and from person to person, three general standards have emerged:

1. Are students performing at or above grade level?

2. Are students performing "adequately" when compared with similar schools or districts?

3. Are students maintaining or improving their performance over time?

4. Given the national emphasis on minimum competency testing and the emphasis in California on absolute standards as embodied in the Model Curriculum, our informants' almost exclusive reliance on norm-referenced interpretations was surprising. Only rarely did respondents ask, "What do these scores mean in terms of actual student competence?" when reviewing data.

A second concern of the more sophisticated board members and virtually all of the school administrators was the need for prescriptive information. They frequently reviewed district reports with the question: "What do the data tell me about how to improve my school?" The concern was not only for what was wrong (or right) but what was causing the inadequate performance and what should be done to improve it. Some board members in our focus groups suggested that district reports include a list of "remedial" steps that would be taken to respond to school weaknesses.

Our clinical interview protocol was designed to probe more fully some of the concerns that arose in the CSBA focus groups. Preliminary analysis of the first six interviews suggests that much of the information used to make decisions about schools is in fact based on "ordinary knowledge" and experience. Respondents claimed to use a variety of sources for decision making with personal experience and direct observation playing a larger role than empirical data. Despite the existence of survey data in all of these districts, neither governing board members nor principals mentioned it as a source for decision making. A surprising finding was that even though our respondents have been frequently exposed to data in the form of technical reports from their own districts or their state, many or even most appeared uncomfortable in interpreting data. Many said that they "relied on the professionals" to summarize and interpret the information for them. They stated that they wanted brief explanations for data discrepancies or unanticipated findings and that these would preferably be presented in a brief executive summary. They suggested that the data tables and graphs that currently make up the bulk of district level reports be appended to these brief explanatory summaries for reference or credibility purposes.

Data Reporting for Effective Decision Making

Our early results suggest that, contrary to our expectations, school planners either don't know how or don't care to use the hypothetical-deductive model of problem solving when using data to answer questions about their schools. Instead they come to data-laden reports with a series of implicit questions and expectations and assimilate those data according to how well they fit extant knowledge. If these findings are replicated over the rest of our sample, the implications for present reporting methods are tremendous. Some possible recommendations from such a finding would be:

1. We will need to provide brief interpretive information about our data in nontechnical language.

2. We will need to key our interpretations to answer the kinds of questions school planners pose when reading our reports.

3. We may need to work with our audiences to educate them as to how data can better inform their own experience and observations.

Before concluding, we would like to re-emphasize the tentative and preliminary nature of our findings about how decision makers process empirical data. At the same time, we would like to extend an invitation for you to join us in this

third phase of the project by conducting interviews with your own governing boards and principals. If, together, we can extend and validate consistent patterns of data use and preference for certain reporting methods, we may perhaps be able to keep our schools afloat and on course.

References

Betteman, J., & Kakker, P. (1977). Effects of presentation format on consumer information acquisition strategies. *Journal of Consumer Research, 3,* 233-240.

Cherry, C. (1966). *On human communication.* Boston, MA: MIT Press.

Cooley, W., & Bickel, W. (1986). *Decision-oriented educational research.* Boston, MA: Kluwer-Nijhoff.

Cousins, J., & Leithwood, K. (1986). Current empirical research on evaluation utilization. *Review of Educational Research, 56*(3), 331-364.

Croxton, F., & Stein, H. (1958). Graphic comparisons by bars, squares, circles, and cubes. *Journal of the American Statistical Association, 27,* 54-60.

Ghani, J. (1981). *The effects of information representation and modification.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Lindbolm, C. & Cohen, D. (1979). *Usable knowledge.* New Haven, CT: Yale University Press.

Lucas, H. (1975). *Why information systems fail.* New York: Columbia University Press.

MacDonald-Ross, M. (1977). How numbers are shown: A review of research on the presentation of quantitative data in texts. *Audio Visual Communication Review, 25* (4), 359-409.

Remus, W., & Kottemann, J. (1986). Toward intelligent decision support systems: An artificially intelligent statistician. *Management Information Systems Quarterly, December,* 403-418.

Scammon, D. (1978). Information load and consumers. *Journal of Consumer Research.*

Schutz, H. (1961). An evaluation of formats for graphic trend displays: Experiment II. *Human Factors,* July.

Slovic, P., & Lichtenstein, S. (1968). The importance of variance preferences in gambling decisions. *Journal of Experimental Psychology, 78,* 646-654.

Vernon, M. (1950). The visual presentation of factual data. *The British Journal of Educational Psychology,* February.

Washburne, J. (1927). An experimental study of various graphic, tabular and textual methods of presenting quantitative material, Part I. *Journal of Educational Psychology, 18,* 361-376.

Weiss, C., & Bucuvalas, M. (1980). *Social science research and decision making.* New York: Columbia University Press.

Table 1

Guidelines for Creating Effective Graphs and Tables

General Guidelines

- Do not write headings in all capitals. Use both upper and lower case.
- Sentences worded negatively are stronger than those worded negatively are stronger than those worded positively. These are best used to present information that is contrary to a reader's expectations.
- Because negatively worded sentences are more difficult to understand and are so emphatic,use them sparingly.
- Questions at the beginning of textual explanations help readers focus on and remember important information.
- Color coded bars and lines on graphs help readers compare information more easily than lines and bars in the same color.

Bar Charts

- Information should be labeled directly on the chart rather than through a "key" or "legend".
- Horizontal bar charts leave more room for labels but may be more difficult to read.

Line Graphs

- Avoid clutter, both of lines and tic marks.
- Choose the range of tic marks to include the entire range of the data; avoid scale breaks.
- If scale breaks are necessary, do not connect numerical values on either side of a break.
- Choose a scale so that data fill up as much of the region as possible.
- Zero does not always have to be included in a scale.
- Use a logarithmic scale when presenting multiplicative factors.

Pictographs

- Symbols should be self-explanatory and easily differentiated from one another.
- Quantity is better represented by increasing the number rather than the size of symbols.

Tables

- Round numbers to no more than two significant digits.
- Provide row and column averages.
- Use columns to display the most important comparisons.
- Order rows and columns by size of numbers rather than alphabetical order.
- Set rows and columns compactly so that the eye can make easy comparisons rather than spacing them out across the page.

Table 2

What Do 'Typical' District Testing Reports Contain?

Based on a review of entries submitted to Division H of the American Educational Research Association for its annual award for "Best Testing Report," we found the following:

- 89% of the district presented results from standardized, nationally-normed tests. Only one district used a locally developed test as its only, measure of district performance.

- 89% of the districts presented reports from a variety of tests. These most often were minimum competency tests, Advanced Placement examinations and state assessment programs. Scholastic Aptitude Tests results were included in a majority of these reports as well.

- The number of different tests reported ranged from one to nine.

- Included with district test results was a variety of information, up to twenty-five different kinds. Among the kinds of information reported in district testing results were:

  1. Teacher/staff information

     — 100% provided number of staff and types of positions
     — 56% provided a breakdown of staff ethnicity and gender and pupil/staff ration
     — 44% gave teacher experience
     — 22% included teacher salaries, staff turnover, and median days of teacher absence

  2. Student information

     — 100% provided enrollment by grade and ethnicity
     — 89% gave teacher experience
     — 56% reported socio-economic status
     — 44% displayed class size and student mobility
     — 33% reported number and percent of LEP, Chapter 1 students
     — 22% provided numbers of students receiving reduced price meals and students enrolled in special education

  3. Other measures of "school effectiveness"

     — 67% gave dropout rates by grade and ethnicity
     — 56% provided graduation rates
     — 33% reported post high school plans of seniors; suspension, retention, and failure rates

Table 3

Interview Protocol

1. "Briefly, how do you know how good a job your school(s) are doing for students?" (List sources of information mentioned, e.g., parent phone calls, newspaper articles/editorials, personal observation of schools, test scores, etc.).

   a. For each source mentioned, probe for how influential that source is in their judgment of their school(s) and why it carries that weight. ("You mentioned a number of sources of information, I'd like to know a little more about how much importance you place in that source and why. How about _____ Would you say it's of overpowering importance, important, or only somewhat important? Why?")

   b. If test scores are not initially mentioned, probe: "Do you use any test data?" (If no, probe for why not; if yes, probe for which specific ones are used)

2. I'd like to know a little more about how you use this information to judge your school(s).

   a. For each source mentioned as important above, ask:
   "If your schools were doing a good job, what would you expect to see in [the information source]?"
   "What in [the information source] signals to you that there is a problem or that some change is needed in your school/ district?"

   b. "Suppose there's some discrepancy between these various sources of information. For example, suppose you thought your math program was pretty good, but your math test scores are relatively low, what would you do/think?"

3. (Show district/school testing report) "There's a lot of information in reports like this and not everyone who reads these reports goes about it in the same way. When you get a report like this, how do you attack it to make some sense out of it? What's the process you go through to find out what you want to know?"

   (Probe if necessary with questions such as: "What's the first thing that gets your attention? And then what? Is there anything else that particularly draws your attention? What questions are you asking yourself as you review such reports

4. When you look at information such as this, are you interested in knowing how different groups of students within the district perform (e.g., how LEP students perform, how Hispanic, Black, Asian, Caucasians perform, how girls vs. boys perform?)

   a. If yes, which subgroups are of most interest to you?

   b. If yes, how is this information useful to you?

5. Let's think a little more about the information that's presented in school/district testing and evaluation reports. We're interested in knowing how to do this better.

   a. First, about how the information is presented. Reports like this typically include narrative text, data tables, graphs and the like. What's the easiest way for you to get information?

   b. Is there some part of this report that is particularly effective? What makes it effective?

   c. (Probe for format comments, e.g., are there any displays that you find particularly informative?)

6. Would you say that test scores capture most of what's important in schooling? [probe for whether it represents most of what's important in students' academic achievement.