Educational Evaluation and Policy Analysis
Summer 1993, Vol. 15, No. 2, pp. 195–207

Data Analysis Strategies for Mixed-Method Evaluation Designs

Valerie J. Caracelli
U.S. General Accounting Office

Jennifer C. Greene
Cornell University

Four integrative data analysis strategies for mixed-method evaluation designs are derived from and illustrated by empirical practice: data transformation, typology development, extreme case analysis, and data consolidation/merging. The appropriateness of these strategies for different kinds of mixed-method intents is then discussed. Where appropriate, such integrative strategies are encouraged as ways to realize the full potential of mixed-methodological approaches.

A formal acknowledgment of the increasing practice of using multiple methods in program evaluation appeared in the 1984 *Evaluation Studies Review Annual:*

> The challenge is to mix the best parts of multiple methods to accomplish our evaluation tasks. Thus far there are more calls for the use of multiple methods than actual examples of how this can be accomplished successfully. Nonetheless, this important shift in thinking is a necessary precondition for the development of new models. Consequently, we anticipate that some very creative multiple method models will begin to appear in the [next] few years. (Connor, Altman, & Jackson, 1984, p. 17)

Since this time, a burgeoning literature has developed around issues pertinent to the use of multiple methods in evaluation and applied research, including triangulation (Mathison, 1988), multiplism (Cook, 1985; Mark & Shotland, 1987; Shadish, Cook, & Houts, 1986; Shotland & Mark, 1987), mixing methods and paradigms (Guba, 1990; Kidder & Fine, 1987; Rossman & Wilson, 1985; Smith & Heshusius, 1986), and mixed-method typologies (Greene & McClintock, 1985; Mark & Shotland, 1987; Maxwell, Bashook, & Sandlow, 1986). Each of these works builds on and extends the classic theoretical literature that underlies interest in multiple research strategies (Campbell & Fiske, 1959; Denzin, 1978; Reichardt & Cook, 1979; Webb, Campbell, Schwartz, & Sechrest, 1966). Only recently, however, has the challenge of developing new models for mixed-method evaluation designs—which fall under the umbrella of multiple methods—been addressed.

Mixed-Method Evaluation Designs in Theory and Practice

Greene, Caracelli, and Graham (1989) reviewed much of the theoretical literature just cited, as well as a purposive sample of 57 mixed-method evaluation studies, in order to begin developing a conceptual framework for mixed-method evaluation designs. In that work, mixed-method designs are defined as including at least one quantitative method (designed to collect numbers) and one qualitative method (designed to collect words), where neither type of method is inherently linked to a particular inquiry paradigm or philosophy.[1] Greene et al. concentrated this conceptual work on clearly differentiating al-

ternative *purposes* for combining qualitative and quantitative methods in program evaluation and on identifying *elements of design choice* related to mixed methodology.[2]

Greene et al. (1989) identified five purposes for mixed-method evaluations, grounded both in the theoretical literature and in evaluation practice as represented by the 57 empirical studies reviewed: triangulation, complementarity, development, initiation, and expansion. In the classic sense, *triangulation* seeks convergence, corroboration, and correspondence of results across the different method types (Campbell & Fiske, 1959; Cook, 1985; Denzin, 1978; Shotland & Mark, 1987; Webb et al., 1966). A *complementarity* purpose is indicated when qualitative and quantitative methods are used to measure overlapping, but distinct facets of the phenomenon under investigation. Results from one method type are intended to enhance, illustrate, or clarify results from the other (Greene & McClintock, 1985; Mark & Shotland, 1987; Rossman & Wilson, 1985). In *development* designs the different method types are used sequentially. The intent, based on the work of Sieber (1973) and Madey (1982), is to use the results of one method to help develop or inform the other method. Development is broadly construed to include sampling and implementation, as well as measurement decisions. Rossman and Wilson (1985) demonstrate that the iterative use of both method types can intentionally seek the discovery of paradox and contradiction. Such *initiation* designs are meant to be provocative through the recasting of questions or results from one method type with questions or results from the contrasting method type. Finally, combining methods for purposes of *expansion* occurs when inquirers extend the breadth and range of inquiry by casting the method types for different inquiry components. In evaluation, quantitative methods frequently play the leading role in assessing program outcomes, while qualitative methods are chosen for the supporting role of examining program processes.

For each of the five purposes a recommended design was also elaborated in terms of seven design elements identified as relevant to mixed methodology. These elements encompass characteristics of methods, the phenomena under investigation, paradigmatic framework, relative status of the different methods, and criteria for implementation.

Greene et al. (1989) further grouped the mixed-method data analysis and interpretation/reporting approaches used in the 57 evaluations reviewed into four categories: (a) no integration, analyses and interpretation of qualitative and quantitative data conducted separately; (b) analyses separate but some integration during interpretation; (c) integration during both analyses and interpretation; and (d) analysis procedures not reported. These findings were crosstabulated by mixed-method purpose.

The results showed that the authors of the majority of empirical studies reviewed either did not report how they conducted their data analyses ($n = 9$) or kept both analysis and interpretation of the two data types separate ($n = 25$). This was especially true for studies that combined methods for the purpose of expansion. When data types were integrated, it was most often at the level of interpretation ($n = 18$) and much more rarely during the analysis process itself ($n = 5$). The paucity of instances of meaningful integration of qualitative and quantitative data at the analysis stage was perplexing given the intentional mixed-method design of these studies.

We believe that a comprehensive conceptual framework for mixed-method evaluations must consider planning for data analysis as a task concomitant with planning the design of a program evaluation. Hence, the present discussion focuses on elaborating the mixed-method analytic strategies used in the handful of evaluations reviewed that did, effectively and at times creatively, integrate quantitative and qualitative data during data analysis, interpretation, and reporting (Hall, Hord, & Griffin, 1980; Louis, 1981; Schermerhorn, Williams, & Dickison, 1982; Talmage & Rasher, 1981). Four major strategies were gleaned from this review: data transformation, typology development, extreme case analysis, and data consolidation/merging. Although these strategies are not new, per se, it is useful to view these analytical techniques in the context of a mixed-method framework. We believe that a closer look at

these studies, supplemented by other examples, can contribute to a rudimentary repertoire of promising analytic strategies for mixed-method evaluations. These strategies are defined in Table 1 and further elaborated in the ensuing discussion.

Integrative Strategies for Mixed-Method Data Analysis

Data Transformation

One means by which qualitative and quantitative data can be integrated during analysis is to transform one data type into the other to allow for statistical or thematic analysis of both data types together.

Larner, Nagy, and Halpern (1987) used this integrative data analysis strategy in their implementation evaluation study in which different methods were used to assess different aspects of program implementation (an expansion purpose). These investigators studied the Rural Alabama Pregnancy and Infant Heath Program (part of the Ford Foundation's Child Survival/Fair Start initiative) in order to determine which factors most importantly affected a client's level of participation in this home visiting program. A quantitative measure was used by the home visitor to assess the level of a mother's participation with respect to nine central aspects of a home visit; for example, how often does the mother ask questions about the lesson? How often does she share personal problems? In addition, home visitors classified the type of relationship they shared with the client as social worker, teacher, or friend. Interviews with clients were also conducted to assess client demographics and two indicators of client social support.[3]

Qualitative data from the interview were transformed into numeric ratings so that all

TABLE 1
Analytical Strategies for the Integration of Qualitative and Quantitative Data

1. Data Transformation—The conversion or transformation of one data type into the other so that both can be analyzed together:
   - Qualititative data are numerically coded and included with quantitative data in statistical analyses.
   - Quantitative data are transformed into narrative and included with qualitative data in thematic or pattern analysis.

2. Typology Development—The analysis of one data type yields a typology (or set of substantive categories) that is then used as a framework applied in analyzing the contrasting data type.
Examples:
   - A set of conceptual dimensions resulting from a factor analysis of quantitative data is incorporated into the categorical analysis of qualitative data (i.e., category development and coding).
   - A respondent or site-level typology resulting from analysis of qualitative data forms a "group" explanatory variable for statistical analyses of quantitative data (e.g., ANOVA, regression analysis) or, as another possibility, is combined with other quantitative explanatory variables for the statistical analysis of qualitative (categorical) data (e.g., logit analysis).

3. Extreme Case Analysis—"Extreme cases" identified from the analysis of one data type and pursued via (additional data collection and) analysis of data of the other type, with the intent of testing and refining the initial explanation for the extreme cases.
Examples:
   - Extreme cases in the form of high residuals from a regression analysis of quantitative data are pursued via (collection and) analysis of qualitative data, the results of which are used to refine the original explanatory model.
   - Extreme cases identified from constant comparative analysis of qualitative data are further examined via analysis of quantitative data, the results of which are used to refine the original interpretation.

4. Data Consolidation/Merging—The joint review of both data types to create new or consolidated variables or data sets, which can be expressed in either quantitative or qualitative form. These consolidated variables or data sets are then typically used in further analyses:
   - Qualitative and quantitative data are jointly reviewed and consolidated into numerical codes or narrative for purposes of further analysis.

variables could be included in the study's main analyses investigating which specific client characteristics were related to their level of participation in the program. A stepwise multiple regression was performed, yielding an overall R^2 of .67. Among the significant predictors, one was the home visitors' perceptions of their role as a friend to the client. From these results, relationship-building became one of the essential components of the home visitor training program.

Thus, in this example, the conjoint analysis of qualitative and quantitative data provided an enriched understanding of factors affecting a client's program participation and served to redirect the home visitor training program to incorporate these significant factors.

Typology Development

In the typology development mixed-method analysis strategy, the analysis of one data type considers the homogeneity within and heterogeneity between subgroupings of data on some dimension of interest, yielding a set of substantive categories or typology. This typology is then incorporated into the analysis of the contrasting data type.

Hall, Hord, and Griffin (1980), using a mixed-method design primarily for development purposes, illustrate the use of this strategy. The authors present results from a 3-year longitudinal study of the implementation of a science curriculum innovation for grades 3–6 in the Jefferson County School District, a large suburban system in Colorado. Both quantitative and qualitative methods were used to determine the factors that influenced teacher change in relation to the new curriculum. The intention was to integrate both types of data at the level of analysis in order to provide "more powerful insights about the change process than either could have produced alone" (p. 3).

The Concerns-Based Adoption Model (Hall, Wallace, & Dossett, 1973) served as a framework for the study. This model assumes that change is carried out by individuals and emphasizes two dimensions that are central to understanding the adoption and implementation of a curriculum innovation: Stages of Concern About the Innovation (SoC) and Levels of Use of the Innovation (LoU). In the curriculum evaluation, existing quantitative measures from this theoretical adoption model were used to classify teachers on these two dimensions. In addition, 41 volumes of ethnographic reports were compiled by a full-time ethnographer located on-site to assess implementation during the first 2 years of the curriculum innovation. District and research staff also contributed observational data and document reviews. Minicase studies were developed from this qualitative data base for nine schools that were representative of the varied phases of implementation across the school district.

The following data analysis strategy was used to achieve integration. A typology of the nine schools was created by placing them in one of three groups based on their SoC profiles. These three groupings represented management-concerned schools, impact-concerned schools, and schools concerned with both management and impact of the innovation. Analysis of the qualitative case study data then concentrated on searching for commonalities within these types of schools, as well as differences among them. Attention was focused on discerning factors (e.g., activities of district and school staff) that affected change among teachers. It was found that the principal's level of support for the innovation and his or her activities as a change facilitator were the main factors influencing the schools' SoC classification.

Thus, the integrated analysis yielded important factors explaining variation in teacher concern for and use of the new science curriculum. Although the primary purpose for using a mixed-method strategy in this evaluation was clearly in line with our definition of development, the actual strategies employed to combine qualitative and quantitative data added hypothesis-generating and initiation components to this evaluation.

A reverse sequence in this integrative analysis strategy can also be used. Patton (1980) discusses the development of emergent typologies from qualitative data. Implementing this procedure, Caracelli (1988) reviewed interview transcripts of adult reentry women in Fordham University's EXCEL program to create a typology of women representing differences in career goal focus. Women with

focused career goals and women with unfocused career goals were then contrasted on data derived from quantitative measures, such as GPA, self-esteem, personality variables, and college satisfaction. The comprehensive portraits of these reentry women that resulted from combining qualitative and quantitative data sources through typology development had implications for program planning. The integrated analysis provided evidence that advisory and counseling needs differed for women depending upon their career goal focus.

In an important mixed-method paper, Rossman and Wilson (1985) illustrate a variation of the typology development analytic strategy, appropriate when the mixed-method purpose is triangulation (which they label *corroboration*). As required for this purpose, analyses of the different data types are conducted independently and then compared for convergence at the level of conclusions and interpretations. Rossman and Wilson's examples are drawn from a large-scale, 3-year evaluation of regional educational service agencies (RESAs). A qualitative review of documents indicated that the RESAs could be categorized as primarily oriented toward either assistance or enforcement activities. Quantitative surveys were then used to probe employees of each agency on the extent to which their work activities emphasized assistance and enforcement. Bivariate plots of mean agency scores from the survey data revealed two clusters that matched the qualitative categorization of RESAs into primarily assistance or enforcement roles for promoting educational reform. The quantitative results were therefore used in a triangulation framework to corroborate the qualitative typology.

These examples suggest that the typology development analysis strategy may be appropriately used for a variety of mixed-method purposes and contexts. Although not illustrated by these examples, one important feature of this strategy is its potential for iteration. A typology could be created from one data type and applied to an analysis of the other data type, the results of which could, in turn, be used to refine and elaborate the typology. This enriched typology could then be reapplied in further analyses of either data type, and so forth, further explicating the initial analyses. Iteration is also a potential feature of the next analysis strategy.

Extreme Case Analysis

A third empirically derived strategy for mixed-method data analysis involves the identification and further analysis of extreme cases. Such cases are identified through analysis of one data type and then further investigated through (additional data collection and) analysis of the other data type. An enhanced understanding of these cases contributes to clarification and refinement of inquiry interpretations.

Rossman and Wilson (1985) also illustrate this analytic strategy from their RESA evaluation, again for the mixed-method purpose of triangulation or corroboration. Data from a survey of local school administrators were used to identify RESAs at both extremes of a continuum of "perceived usefulness." These RESAs were then investigated more intensively through qualitative case studies, the results of which were used to corroborate the survey findings.

In a variation of extreme case analysis, Fry, Chantavanich, and Chantavanich (1981) conducted three mixed-method cross-cultural studies in Thailand. These studies led the authors to espouse the technique of ethnographic residual analysis which, as an integrated analysis strategy, is closely aligned with our identified design purpose of initiation. In the context of cross-cultural educational research, the authors propose the technique in order to obtain "new ideas, insights, hypotheses, and understandings" (p. 153) and "a deeper and better understanding of the complex interrelationships among educational inputs, educational processes, and educational outcomes" (p. 155).

With this approach a school's expected effectiveness is specified by quantitative indicators, such as students' cognitive and noncognitive abilities, values, and attitudes. Then a school's expected effectiveness is compared with its actual effectiveness through multiple regression techniques that incorporate explanatory factors such as the socioeconomic background of students, school financial re-

sources per person, and the teacher-student ratio. Anomalies in terms of schools that have either unusually high or unusually low quality, relative to their educational inputs, are then examined ethnographically to try to ascertain qualitative factors accounting for statistical deviance and unexplained error variance.

Ethnographic investigators, who are blind to the quantitative findings, are sent to these "extreme case" schools to study their educational process. The ethnographic analysis is specifically expected to generate insights that foster the development of new concepts or categories. For example, differences in teaching methods, principal characteristics, or community support for education may be factors important in assessments of school effectiveness. Finally, these ethnographic variables are incorporated back into the regression model in an effort to increase the explanatory power of the model, and thus the depth of conceptual understanding.

As a mixed-method integrated analysis strategy, ethnographic residual analysis has the potential for recasting or elaborating the theory that directs the initial analysis. It can be viewed as a mixed-method counterpart to the use of negative case analysis by participant observers (Kidder, 1981). Both negative case analysis and ethnographic residual analysis systematically search for cases that may provide disconfirming evidence for the hypothesis under investigation, leading to refinements of the hypothesis.

Data Consolidation/Merging

Our final mixed-method analysis strategy, data consolidation or merging, involves the more sophisticated, joint use of both data types to create new or consolidated variables or data sets. These consolidated data types can be expressed in either quantitative or qualitative form, and would be appropriately used in further analysis. As illustrated by the following studies, this data analysis strategy may be especially suitable for mixed-method designs with initiation intents (i.e., the use of mixed methods to uncover fresh insights or new perspectives). The more extensive examples offered for this strategy are intended to underscore its unique reliance on multiple, varied sets of data.

Talmage and Rasher's (1981) approach for merging qualitative and quantitative data at the level of analysis is explicitly linked to our mixed-method purpose of initiation. Their work demonstrated how the integration of both types of data could generate new variables, extending the scope of their data base to address elusive evaluation problems in a school setting. This formative evaluation assessed in 11 urban schools the Model Builders Project (MBP), a 3-year arts-in-the-schools program, and the factors that influenced implementation and program effects.

The authors summarize their dialectic approach to integrating both types of data in terms of a "spiral effect." In the first year of the evaluation, quantitative data were gathered in the form of self-report questionnaires, structured interviews, and structured classroom observations. The quasi-experimental design permitted comparisons among MBP participants, nonparticipants at MBP sites, and two comparison schools on students' perceptions of their classroom learning environment, degree of program implementation, art-related activities, and course evaluations. No program effects were discerned. Nevertheless, the evaluators "felt" there was a program impact that was escaping traditional instrumentation, and so they shifted the methodological thrust of the evaluation.

During the second and third years, mini-case studies were completed to examine the implementation and impact of the program within and across sites. Using semistructured, open-ended observation and interview instruments, trained data collectors gathered data from a wide range of sources, including administrators, teachers, artists/instructors, school support personnel, students, and parents. As the case studies were prepared, it was evident that much of the qualitative data had quantifiable aspects that not only were relevant to the case study but also could serve to augment the first-year quantitative data base. Thus, patterns within the qualitative data were transformed into quantitative form through categorization and ratings.

However, the quantitative data base was not augmented solely with the *addition* of transformed qualitative data. Rather, some

new variables were created through a merging of both qualitative and quantitative data. The authors note that in addition to the two data forms providing [supplementary and] supportive information, a spiral effect was occurring; each type of information, when combined, displayed a dynamic interconnectedness. The integration was leading to a synthesis that produced a new variable (p. 9).

The example given by the authors focused on the creation of the variable "principal support." Qualitative data from principal interviews and quantitative data from teacher and artist/instructor questionnaires were both assessed in order to determine a quantifiable rating that would capture the level of "principal support" (1 = minimal to 4 = extensive) for the MBP project. This *merged-data* variable was found to be significantly correlated with the extent of implementation ($r = .74$, $p < .01$), thereby furnishing critical information that was not apparent from the independent analyses of either quantitative or qualitative data alone.

Schermerhorn, Williams, and Dickison (1982) provide a further example of this data consolidation and merging analysis strategy in their multi-purpose,[4] mixed-method evaluation of Project COMPAS. Project COMPAS (Consortium for Operating and Managing Programs for the Advancement of Skills) represented a cooperative effort among seven community colleges to develop cognitive skills programs for entry-level freshman. Again, initiation elements are evident in the authors' reflections on their analysis process:

> The reconciliation of the two data sets is thus more a task *of weaving together* the influences resulting from each set than of confirming one inference with supportive evidence from a second perspective [as in triangulation]. In some instances a common theme is discerned, though for most concerns *only questions arise* as the two data sets are merged. (p. 95, italics added)

In other words, a recasting of questions concerning program impact for future evaluations of Project COMPAS was one important outcome of this study.

In this initial evaluation, the process of weaving together the qualitative and quantitative data sources resulted in the discovery of an important factor that had not been considered in the original evaluation design—the degree to which students were immersed in the program. To capture this phenomenon, an "immersion" variable was constructed using both quantitative and qualitative data sources. An important outcome of subsequent analyses was the finding that immersion contributed to or moderated program outcomes and the attendant recognition that measurement of the level of a student's immersion in the program would need to be refined and included in future assessments of project impacts.

Louis (1981, 1982) describes an interactive analytic model, again with clear links to initiation intents. Louis's model is explicitly focused on integrating the data obtained from different instruments, respondents, and observers. The model evolved during a multisite longitudinal evaluation of the Research and Development Utilization Program (RDU). This $8 million demonstration project was funded by the National Institute of Education between 1976 and 1979 to promote the adoption and implementation of new curriculum and staff development materials in 300 local schools.

A variety of data collection methods were used throughout the project, including miniethnographies based on interviews, observations, and document analysis; case study writers' surveys; standardized site-visit field reports; "event-triggered" reports monitoring a school's progress through the project; and formal principal and teacher surveys. Site-level data were thus rich and diverse; however, no more than 20% of the sites had a complete data set, which seriously constrained cross-site analysis possibilities. To overcome this constraint, these evaluators created a transformed and consolidated site-level data set via the development and application of a "consolidated coding form" (CCF). The form constituted 240 dichotomous or Likert scale items, which were scored by senior staff members who had visited at least four of the sites and were involved in an intensive 2-day session in which common interpretations for consolidated coding were reached. Included on the CCF were variables that could not be readily obtained through

traditional survey methods, for example, quality of the decision-making process and patterns of influence of different actors over decisions at various stages in the change process. Moreover, the consolidated data base reflected the holistic knowledge the site-visit team brought to the cases, as well as the reliability of standardized data, integrated both within and across sites.

The level of integration of qualitative and quantitative data achieved in the RDU evaluation is captured in the following summary statement:

> Can a database composed of numbers that is entirely dependent on the iterative, holistic judgments of experienced site field teams be described as only quantitative? While the analysis procedures used to manipulate the data are statistical, the data itself, and any interpretation of results, is totally conditioned by its origins. On the other hand, as we approach any given analysis using case materials rather than quantified data, it has become genuinely impossible not to embed that activity in our knowledge of the descriptive statistics and correlational relationships that were available to us well before data collection had ended. (Louis, 1981, p. 21)

Louis cautions that this comprehensive, interactive approach to analytic integration requires constant attention by staff members who are skilled in both quantitative and qualitative data analysis techniques. Low rates of turnover among project staff, who are relatively free of paradigmatic preferences, would also be essential to achieving the high level of integration that was obtained in this evaluation.

In these three examples of mixed-method inquiry, the data consolidation/merging analysis strategy was used effectively. Through data consolidation/merging, the authors of all three of these studies were able to create new variables and conduct a more comprehensive analysis, which served to provoke insights and new perspectives on planned evaluation foci. The exigencies of multisite data coordination and analysis may have been the imperative behind the iterative nature of the analyses and maximal use of both data types. In Louis (1981, 1982) the development of a consolidated coding form provided a means by which a data base could be created from a plethora of sources that compensated for missing data, reflected the holistic knowledge of the site team, and ensured standardized data across sites. In both Talmage and Rasher (1981) and Schermerhorn, Williams, and Dickison (1982), new variables were constructed out of the joint use of both data types and were subsequently quantified for further analysis. It is certainly also possible to generate new themes or patterns from a merged analysis of quantitative and qualitative data and then use them in further qualitative analyses.

Discussion

This paper constitutes a continuing response to Connor et al.'s (1984) challenge "to mix the best parts of multiple methods to accomplish our evaluation tasks." Under the umbrella of multiple methods, mixed-method evaluations now frequently dot the landscape of evaluation research. Our work is focused on providing a conceptual framework that can effectively guide mixed-method evaluation practice. Previously, we identified distinct mixed-method *purposes* and relevant *design elements*. The present discussion contributes an initial repertoire of four *data analysis strategies* appropriate for mixed-method studies. The critical defining characteristic of all four strategies is their integration of the different data sets during the analysis process itself. Some form of integration, we believe, constitutes the essence of a mixed-method approach.

As noted earlier, the data analysis strategies presented here are not, in and of themselves, new. It is common research practice, for example, to code numerically qualitative data for purposes of statistical analysis and to single out extreme cases or residuals for more intensive scrutiny. What is new, we believe, is the collection of these data analysis strategies within a mixed-method framework. This framework highlights the integrative potential of these strategies, and underscores their potential power not only to incorporate qualitative data into quantitative analyses, but also vice versa, and, even beyond, to spiral iteratively around the different data sets, adding depth of understanding with each cycle.

Yet, as noted, in our prior review of empirical studies mixed-method evaluation practice rarely incorporated an integrative analysis strategy. From this disjuncture, two questions arise: (a) When is an integrative analysis strategy appropriate, and (b) why is integrative analysis so rare in practice?

First, to the practical question of contextual appropriateness, we offer provisional guidelines that are linked to mixed-method purpose, for this remains the cornerstone of our conceptual framework. We suggest that, in general, integrative analytic strategies are appropriate when methods are mixed for purposes of initiation, expansion, or development, but less useful when triangulation is the mixed-method intent.

The studies reviewed in this article clearly illustrate the value of integrative analyses for initiating fresh insights and new perspectives that enhance conceptual understanding. The examples included initiation uses of three of the four analytic strategies discussed—typology development, extreme case analysis, and data consolidation/merging—suggesting a particularly strong match between initiation mixed-method designs and integrative data analysis strategies.

The Larner, Nagy, and Halpern (1987) evaluation of the rural Alabama home visitor health program provides an example of using data transformation as an analytic strategy in an expansion mixed-method design. In this study, qualitative data on program implementation were numerically coded and incorporated, along with quantitative implementation data, into a regression analysis predicting client program participation. Data transformation is perhaps the most obviously useful analysis strategy in expansion designs, where different methods are employed to increase the breadth and scope of the inquiry. Data transformation would enable analyses of the relationships between typically qualitative information on program processes and typically quantitative information on program outcomes. More effective use of this design, however, could be enabled by creative applications of other integrative data analysis strategies. For example, outcome data analyses could signal residuals or outliers or generate a typology useful for more intensive implementation analyses. Different data on different program components could even be consolidated or merged to capture dynamic patterns of program experiences. Expansion designs dominated our prior empirical review of mixed-method evaluation practice, composing nearly half the sample. Yet the potential power and benefits of such designs may well remain unfulfilled without more conscientious attention to integrative analyses.

A salient characteristic of development mixed-method designs is their *sequential* character, where the results of the first method are used to inform the development of the second (including instrumentation, sampling, and administration decisions). Two of the integrative analysis strategies described here are also sequential—typology development and extreme case analysis—and thus potentially strong analytic approaches for development designs. The Hall et al. (1980) evaluation of an innovative science curriculum illustrates the analytic value of typologies for mixed-method development studies. Specifically, a typology of schools created from the quantitative data usefully framed and focused the subsequent qualitative analysis in this study. And extreme case analysis can be viewed as a special case of the classic development design in which the results of the first method are used to select the samples for the second (Sieber, 1973). Because the different data types are processed simultaneously in the other two integrative analysis strategies, these are unlikely to support a mixed-method development design.

In evaluations where complementarity is the primary purpose for mixing methods, the decisions guiding separate versus integrative processing of the different data types are not as clear-cut. In part, this is due to the particularly wide contextual variability possible in the design of such studies. In complementarity designs, different methods are used to measure overlapping, but also distinct facets of a given phenomenon. The greater the overlap in the conceptualizations of the phenomenon guiding each method, the closer this design is to a triangulation design, for which we believe integrative analysis strategies are not generally useful. The less the

overlap, the closer this design is to an expansion design, for which we believe integrative analysis strategies can offer strong support. The present review did not include an example of integrative analysis in a complementarity study. Logic nonetheless clearly suggests that there should be many cases where the joint analysis of data from methods implemented to develop an elaborated, enriched understanding of a phenomenon would serve well to do just that.

Finally, in contrast to our promotion of integrative analysis strategies for evaluations that mix methods for initiation, expansion, development, and complementarity purposes, the very concept of data integration is less meaningful when methods are mixed for purposes of triangulation. The underlying logic of triangulation requires independence of methods through data analysis and interpretation. Arguments for convergent validity of findings from different methods are stronger when such independence can be claimed. Hence, to integrate different data sets intentionally during data analysis is to undermine the potential power of a triangulation design.

To the second question of why integrative analyses in mixed-method evaluation practice are still a rarity, we offer two sets of speculations, one pertaining to the evolving contexts of program evaluation and the other to methodological stances within the field.[5] Contextually, funding has been reduced for large-scale, multisite evaluations that are conducive to thoughtful mixed-method designs and, as illustrated by the Louis (1981) study, invoke the need for integrative analyses. Smaller data sets from single sites may be more readily managed and understood without the felt need for coordination.

Methodologically, there are three recognized stances within the community of program evaluators that mitigate against meaningful integration in mixed-method practice. First, mixed-method designs are often inaccurately equated with the *in vogue* concept of triangulation. Integral to this concept is strong independence of the different methods used. So evaluators employing mixed-method designs who adopt the rubric of triangulation, even when theirs is not a triangulation design, may eschew or fail to even consider the potential of integrative strategies. Second, this is an era of dizzying pluralism in social inquiry approaches and justifications (see Guba, 1990). For many, this pluralism connotes a basic acceptance of diverse ways of knowing and diverse things worth knowing about, from propositional causal claims to experiential meaning and to critical sources of distortion in communications. In accepting diversity, however, many social inquirers have effectively retrenched, rejecting either the possibility or the desirability of integrative rapprochement among different kinds of knowledge claims. This kind of climate—where some are gathering with their own behind barricades, propelling philosophical and political salvos (Sechrest, 1992) over to the other side—is surely not very hospitable to the concept of integration. Yet, this challenge of paradigmatic integration remains important for future mixed-method development. Finally, in several currently popular evaluation approaches, integration is a sensible concept but not necessarily through data bases or analytic methods. These alternatives include interpretivist, qualitative approaches (from Stake, 1975, through Guba & Lincoln, 1989), in which integration and synthesis are intellectual tasks demanded of the individual evaluator; more openly ideological approaches like Schwandt's (1989), which calls for moral evaluation, and Sirotnik's and others' (Sirotnik, 1990), which call for an evaluation practice oriented around social justice in which integration becomes meaningful only in the service of some ideological aim.

In summary, we have identified four integrative data analysis strategies and provisionally argued for their value in many mixed-method evaluation designs. We believe that the intentional use of such strategies can significantly augment the power of these designs to advance conceptual understanding and insight. Clearly, further work is also needed. Concerted attention must be directed to the role of inquiry paradigms in integrative data analysis strategies and mixed-method inquiry more generally. How can contrasting epistemological assumptions and worldviews be integrated or reconciled within a mixed-method framework? From the present work, the data consolidation and merging approach

emerged as a promising strategy for data integration yet perhaps also the strategy most vulnerable to abuse from conceptualization and measurement perspectives. For example, how, if at all, should different data types be weighted when consolidating or merging them (Cordray, 1986; Jick, 1983)? More examples of successful practice employing integrative data analysis strategies in mixed-method contexts are also needed; this work, in particular, relies on an iterative interplay of theory and practice.

In this pluralistic era in applied social inquiry, mixed-method approaches are likely to continue to increase in desirability and frequency. The power and added value of such designs can be realized only if mixed-method decisions are systematic and explicit. The integrative data analysis strategies presented here are offered as contributions toward that end.

Notes

The statements and opinions expressed in this article do not represent official policy of the General Accounting Office. The authors would like to extend grateful thanks to Leslie J. C. Riggin, Robert A. Johnson, and two anonymous reviewers for their constructive contributions to this manuscript.

[1]The key distinction here is between methods that yield numerical data and those that yield narrative or other forms of data (see note 2). Although such different methods are often linked to different inquiry philosophies, these linkages do not inhere in the methods (Bednarz, 1983; Reichardt & Cook, 1979). In the Greene et al. mixed-method framework, this issue is addressed through the design element of inquiry paradigm, where the paradigm guiding each method type is delineated.

[2]As envisioned, the full conceptual framework also includes broader issues related to evaluation context, purpose (e.g., formative, summative, critical), audience, and intended uses. Work on the conceptual framework to date has concentrated on mixing qualitative and quantitative inquiry methods. While this emphasis matches current practice, it is not intended to exclude the emerging importance of methods reflecting critical perspectives or perspectives drawn from the humanities.

[3]The classification of measures as quantitative or qualitative by a third party is sometimes a matter of judgment. For the Larner et al. study, a personal communication with the principal author revealed that the quantitative rating scale was derived from qualitative focus groups conducted with the home visitor staff. With this procedure, ratings would reflect the experiential nature of client-staff relationships as perceived by the home visitors. The interviews, which assessed demographics and social support, could actually be considered more quantitative than the rating scale.

[4]The authors' stated primary purpose for using a mixed-method design was complementarity. We inferred, however, that a development purpose evolved from the study design and that both initiation and expansion purposes emerged during the analysis phase. In Greene et al. (1989) over a fifth (n = 13) of the evaluations were rated for both primary and secondary mixed-method purposes.

[5]We are indebted to an anonymous reviewer for many of these ideas.

References

Bednarz, D. (1983). *Quantity and quality in evaluation research: A divergent view*. Revised version of paper presented at the Joint Meeting of the Evaluation Network and the Evaluation Research Society, Chicago.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105.

Caracelli, V. J. (1988). *Adult women in college. Stability and change in identity and personality*. Unpublished doctoral dissertation, Cornell University, Ithaca, NY.

Connor, R. F., Altman, D. G., & Jackson, C. (1984). 1984: A brave new world for evaluation. In R. F. Connor, D. G. Altman, & C. Jackson (Eds.), *Evaluation Studies Review Annual: Vol. 9* (pp. 13–22). Beverly Hills, CA: Sage.

Cook, T. D. (1985). Postpositivist critical multiplism. In R. L. Shotland & M. M. Mark (Eds.), *Social science and social policy* (pp. 21–62). Beverly Hills, CA: Sage.

Cordray, D. S. (1986). Quasi-experimental analyses: A mixture of methods and judgment. In W. M. K. Trochim (Ed.), *Advances in quasi-experimental design and analysis. New Directions for Program Evaluation 31* (pp. 9–27). San Francisco: Jossey-Bass.

Denzin, N. K. (1978). *The research act: An introduction to sociological methods* (chap. 10). New York: McGraw-Hill.

Fry, G., Chantavanich, S., & Chantavanich, A. (1981). Merging quantitative and qualitative research techniques: Toward a new research paradigm. *Anthropology and Educational Quarterly, 12*, 145–158.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*, 255–274.

Greene, J. C., & McClintock, C. (1985). Triangulation in evaluation: Design and analysis issues. *Evaluation Review, 9*, 523–545.

Guba, E. G. (1990). *The paradigm dialog*. Newbury Park, CA: Sage.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

Hall, G., Hord, S. M., & Griffin, T. H. (1980, April). *Implementation at the school building level: The development and analysis of nine mini-case studies*. Paper presented at the Annual Meeting of the American Educational Research Association, Boston. (ERIC Document Reproduction Service No. ED 207 170)

Hall, G. E., Wallace, R. C., & Dossett, W. A. (1973). *A developmental conceptualization of the adoption process within educational institutions*. Austin, TX: The University of Texas, Austin, Research and Development Center for Teacher Education.

Jick, T. D. (1983). Mixing qualitative and quantitative methods: Triangulation in action. In J. Van Maanen (Ed.), *Qualitative methodology* (pp. 135–148). Beverly Hills, CA: Sage.

Kidder, L. H. (1981). *Selltiz, Wrightsman and Cook's research methods in social relations*. (4th ed.). New York: Holt, Rinehart & Winston.

Kidder, L. H., & Fine, M. (1987). Qualitative and quantitative methods: When stories converge. In M. M. Mark & R. L. Shotland (Eds.), *Multiple methods in program evaluation. New Directions for Program Evaluation 35* (pp. 57–75). San Francisco: Jossey-Bass.

Larner, M., Nagy, C., & Halpern, R. (1987, October). *Inside the black box: Understanding home visiting programs*. Paper presented at the annual meeting of the American Public Health Association, New Orleans.

Louis, K. S. (1981, April). *Policy researcher as sleuth: New approaches to integrating qualitative and quantitative methods*. Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles. (ERIC Document Reproduction Service No. ED 207 256)

Louis, K. S. (1982). Sociologist as sleuth: Integrating methods in the RDU study. *American Behavioral Scientist, 26*(1), 101–120.

Madey, D. L. (1982). Some benefits of integrating qualitative and quantitative methods in program evaluation, with illustrations. *Educational Evaluation and Policy Analysis, 4*, 223–236.

Mark, M. M., & Shotland, R. L. (1987). Alternative models for the use of multiple methods. In M. M. Mark & R. L. Shotland (Eds.), *Multiple methods in program evaluation. New Directions for Program Evaluation 35* (pp. 95–100). San Francisco: Jossey-Bass.

Mathison, S. (1988). Why triangulate? *Educational Researcher, 17*(2), 13–17.

Maxwell, J. A., Bashook, P. G., & Sandlow, C. J. (1986). Combining ethnographic and experimental methods in educational evaluation. In D. M. Fetterman & M. A. Pitman (Eds.), *Educational evaluation: Ethnography in theory, practice, and politics* (pp. 121–143). Beverly Hills, CA: Sage.

Patton, M. Q. (1980). *Qualitative evaluation models*. Beverly Hills, CA: Sage.

Reichardt, C. S., & Cook, T. D. (1979). Beyond qualitative *versus* quantitative methods. In T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research* (pp. 7–32). Beverly Hills, CA: Sage.

Rossman, G. B., & Wilson, B. L. (1985). Numbers and words: Combining quantitative and qualitative methods in a single large-scale evaluation study. *Evaluation Review, 9*, 627–643.

Schermerhorn, L. L., Williams, L., & Dickison, A. (Eds.). (1982). *Project COMPAS [Consortium for operating and managing programs for the advancement of skills]: A design for change*. East Peoria, IL: Illinois Central College. (ERIC Document Reproduction Service No. ED 219 100)

Schwandt, T. A. (1989). Recapturing moral discourse in evaluation. *Educational Researcher, 18*(8), 11–16.

Sechrest, L. (1992). Roots: Back to our first generations. AEA 1991 presidential address. *Evaluation Practice, 13*(1), 1–7.

Shadish, W. R., Jr., Cook, T. D., & Houts, A. C. (1986). Quasi-experimentation in a critical multiplist mode. In W. M. K. Trochim (Ed.), *Advances in quasi-experimental design and analysis. New Directions for Program Evaluation 31*, (pp. 29–46). San Francisco: Jossey-Bass.

Shotland, R. L., & Mark, M. M. (1987). Improving inferences from multiple methods. In M. M. Mark & R. L. Shotland (Eds.), *Multiple methods in program evaluation. New Directions for Program Evaluation 35* (pp. 77–94). San Francisco: Jossey-Bass.

Sieber, S. D. (1973). The integration of field work and survey methods. *American Journal of Sociology, 78*, 1335–1359.

Sirotnik, K. A. (Ed.). (1990). *Evaluation and social justice. New Directions for Program Evaluation 45*. San Francisco: Jossey-Bass.

Smith, J. K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate. *Educational Researcher, 15*(1), 4–12.

Stake, R. E. (1975). *Evaluating the arts in education: A responsive approach*. Columbus, OH: Merrill.

Talmage, H., & Rasher, S. P. (1981, April). *Quantifying qualitative data: The best of both worlds*. Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles. (ERIC Document Reproduction Service No. 204 396)

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures* (chap. 1). Chicago: Rand McNally.

Authors

VALERIE J. CARACELLI, Senior Social Science Analyst, U.S. General Accounting Office, Program Evaluation and Methodology Division, 441 G St., NW, Washington, DC 20548. *Specializations:* multiple methods evaluation research, applied social science research methods.

JENNIFER C. GREENE, Associate Professor of Human Service Studies, Department of Human Service Studies, N136B MVR Hall, Cornell University, Ithaca, NY 14853. *Specializations:* evaluation theory and practice, participatory inquiry.

Received July 27, 1992
Revision received February 18, 1993
Accepted February 22, 1993