Research in Nursing & Health, 2000, 23, 246–255

Focus on Research Methods

Combining Qualitative and Quantitative Sampling, Data Collection, and Analysis Techniques in Mixed-Method Studies

Margarete Sandelowski*

University of North Carolina at Chapel Hill, School of Nursing,
7460 Carrington Hall, Chapel Hill, NC 27599
Received 12 July 1999; accepted 13 January 2000

Abstract: Researchers have increasingly turned to mixed-method techniques to expand the scope and improve the analytic power of their studies. Yet there is still relatively little direction on and much confusion about how to combine qualitative and quantitative techniques. These techniques are neither paradigm- nor method-linked; researchers' orientations to inquiry and their methodological commitments will influence how they use them. Examples of sampling combinations include criterion sampling from instrument scores, random purposeful sampling, and stratified purposeful sampling. Examples of data collection combinations include the use of instruments for fuller qualitative description, for validation, as guides for purposeful sampling, and as elicitation devices in interviews. Examples of data analysis combinations include interpretively linking qualitative and quantitative data sets and the transformation processes of qualitizing and quantitizing. © 2000 John Wiley & Sons, Inc. Res Nurs & Health 23:246–255, 2000

Keywords: mixed-method studies; combination studies; qualitative and quantitative research; sampling; data collection; data analysis; triangulation

The idea of mixing qualitative and quantitative methods has stimulated much interest and debate (e.g., Greene & Caracelli, 1997a; Sandelowski, 1995; Swanson, 1992; Tashakkori & Teddlie, 1998). Researchers increasingly have used mixed-method techniques to expand the scope of, and deepen their insights from, their studies. As advocates of mixed-method research have argued, the complexity of human phenomena mandates more complex research designs to capture them. Despite this interest, there is still relatively little direction on and much confusion about how to accomplish mixed-method studies. In this paper, I discuss the where, what, why, and especially, the how of studies combining qualitative and quantitative techniques.

THE WHERE, WHY, AND WHAT OF COMBINATIONS

Combination or mixed-method studies are concretely operationalized at the technique level, or the shop floor, of research: that is, at the level of sampling, data collection, and data analysis. Mixed-method studies are not mixtures of paradigms of inquiry per se, but rather paradigms are

*Professor.

reflected in what techniques researchers choose to combine, and how and why they desire to combine them.

At the Paradigm Level

As many scholars (e.g., Guba & Lincoln, 1994; Heron & Reason, 1997) have variously named and described them, paradigms of inquiry are worldviews that signal distinctive ontological (view of reality), epistemological (view of knowing and the relationship between knower and to-be-known), methodological (view of mode of inquiry), and axiological (view of what is valuable) positions. Indeed, paradigms of inquiry are best understood as viewing positions: ways, and places from which, to see. Because different paradigms, such as (neo- or post-) positivism, constructivism, critical theory, and participatory inquiry, entail contradictory viewing positions, combinations at the paradigm level are not true combinations, mergers, or reconciliations of worldviews, but rather the explicit framing of inquiry in two or more worldviews, each of which remains distinct from the other. That is, for example, it is not possible to combine, merge, or reconcile a view of reality as singular and objective (positivist) with views of it as multiple and individually or culturally constructed (constructivist), or as historically contingent (critical theory). A critical theorist will frame the issues around hormone replacement for midlife women differently from a neopositivist.[1] That is, they will each *see* different things and, therefore, ask different questions that will, in turn, require the use of different methods and techniques to answer them. Such a paradigm combination may be used to elicit two or more perspectives on hormone therapy, or for the purpose of "initiation" (Greene, Caracelli, & Graham, 1989, p. 259), to surface paradoxes and contradictions that surround hormone therapy. Accordingly, two or more paradigms of inquiry can be used to frame the same target phenomenon (in this case, hormone replacement). Yet, arguably, the positivist and the critical theorist may not really be studying the same phenomenon, because to see a phenomenon in a certain way is to change that phenomenon. As Pearce (1971, p. 2) observed, "a change of world view can change the world viewed." Moreover, it is debatable whether one researcher can hold two different viewing positions, albeit at different times. That is, it is uncertain whether a researcher can conduct a project framed first, for example, in positivism and then later, in critical theory. Such worldviews, like all strongly held belief systems, are not easily exchanged.

As data collection and analysis techniques are not linked to paradigms (Berman, Ford-Gilboe, & Campbell, 1998; Sandelowski 1995), both the researcher in a positivist viewing position and the researcher in a critical theory viewing position may use interviews and even the very same standardized measures to answer their questions, but they will employ these techniques and, more importantly, analytically treat their results differently. Indeed, arguably the key difference between "qualitative" and "quantitative" (when these words are used to signal different inquiry paradigms or research traditions) researchers is in their attitude toward and treatment of data. Accordingly, although techniques can be mixed, the resulting mix will reveal the researcher's viewing position (or, in cases of mixed-up research, the researcher's futile effort to mix the research equivalent of oil and water).

At the Method Level

Combinations at the method level can be used to expand the scope of a study as researchers seek to capture method-linked dimensions of a target phenomenon (Greene et al., 1989, p. 259). As Wolfer (1993) proposed, different aspects of reality lend themselves to different methods of inquiry. For example, the physiology of hormone replacement therapy for women lends itself to study via physiologic measures, while the debate around the benefits and liabilities of hormone replacement therapy lends itself to discourse, semiotic, and other cultural analyses. In the study I conducted with my colleagues of the transition to parenthood of infertile couples (Sandelowski, Holditch-Davis, & Harris, 1992), we used grounded theory methodology with naturalistic, ethological observation to capture different features of this phenomenon, including couples' understandings of their struggle to become parents and their interaction with their babies. Grounded theory and ethological observation share a common naturalist imperative (that is, not to manipulate, or impose a priori conceptualizations on, the target phenomenon) that make them compatible for use in one study. But grounded theory is particularly well-suited to theorizing human understandings, while ethological observation is especially well-suited to theorizing human behavior.

Although the various dimensions of phenomena may be method-linked in that different dimensions may be best captured by different methods,

[1] I use terms like "positivist" and "positivism" to quickly communicate complex and admittedly controversial ideas for the purposes of this paper, not to stereotype researchers or viewing positions.

methods, like paradigms, are not specifically linked to techniques. Grounded theory may be generated using an array of qualitative and quantitative data collection techniques and sources. Moreover, methods are not uniformly linked to paradigms (Greene & Caracelli, 1997b). Grounded theory may be conducted in a neopositivist or constructivist paradigm (Annells, 1996). A neopositivist researcher conducting grounded theory believes in an external and objectively verifiable reality. In contrast, a constructivist conducting grounded theory believes in multiple, experientially based, and socially constructed realities. For the neopositivist, concepts emerge or are discovered, as if they were there to be found. The act of discovery is separate from that which is discovered. For the constructivist, concepts are made, fashioned, or invented from data. What constructivists find is what they made. For the constructivist, all human discovery is creation.

At the Technique Level

The technique level of research is the site where combinations actually occur and is what is most often referred to in discussions of mixed-method research. Such combinations entail the use of sampling, data collection, and data analysis techniques commonly (although not necessarily) conceived as qualitative or quantitative. Because techniques are tied neither to paradigms nor to methods, combinations at the technique level permit innovative uses of a range of techniques for a variety of purposes. Three purposes include (a) triangulation, to achieve or ensure corroboration of data, or convergent validation; (b) complementarity, to clarify, explain, or otherwise more fully elaborate the results of analyses; and (c) development, to guide the use of additional sampling, and data collection and analysis techniques (Greene et al., 1989, p. 259). The rest of this paper is devoted to illustrating how such purposes can be achieved.

THE HOW OF COMBINATIONS

Mixed-method studies entail concrete operations at the technique level of research by which "qualitative" and "quantitative" techniques are used together and either remain distinct design components, or are explicitly integrated (Caracelli & Greene, 1997). As shown in Figure 1, either qualitative or quantitative approaches to sampling, data collection, and data analysis may prevail or have equal priority in a study, and both qualitative and quantitative approaches may be used sequentially, concurrently and iteratively, or in a sandwich pattern.

Combining Sampling Strategies

One of the most important features distinguishing what is commonly referred to as qualitative from quantitative inquiry is the kind of sampling used. While qualitative research typically involves purposeful sampling to enhance understanding of the information-rich case (Patton, 1990), quantitative research ideally involves probability sampling to permit statistical inferences to be made. Although purposeful sampling is oriented toward the development of idiographic knowledge — from generalizations from and about individual cases — probability sampling is oriented toward the development of nomothetic knowledge, from generalizations from samples to populations. Notwithstanding these key differences, purposeful and probability sampling techniques can be combined usefully.

Criterion sampling. For example, in design template 2 shown in Figure 1, in which the use of quantitative techniques precede the use of qualitative techniques, research participants' scores on the instruments used to collect data in the quantitative portion of the study can be used to initiate a criterion sampling strategy. Criterion sampling is a kind of purposeful sampling of cases on preconceived criteria, such as scores on an instrument. Cases may be chosen because they typify the average score; this kind of sampling may also be referred to as typical case sampling. Cases may be chosen because they exemplify extreme scores; this kind of sampling may also be called extreme or deviant case sampling. (The term deviant here refers to any departure from a specified norm.) Such cases are highly unusual. Or, cases may be chosen because they show a variable intensely, but not extremely; this kind of sampling may also be referred to as intensity sampling (Patton, 1990, pp. 182–183).

Researchers using scores on instruments as the criterion for purposeful sampling may wish to collect more data (e.g., via interviews or observations) from the chosen participants for the purpose of triangulation: that is, to discern whether a typical, extreme, or intense case of something on a standardized test is also a typical, extreme, or intense case using other data collection techniques. Or, researchers may use the criterion of scores (typical, extreme, or intense scores) for the purpose of complementarity: that is, to find out more about what makes a case typical, extreme, or in-

Templates	Qualitative/Quantitative Relationship: Priority & Temporality	Use of Qualitative Adjunct:	Use of Quantitative Adjunct:
Template #1 Template #1a	QUAL>quan Or QUAL+quan		-measured description -validation -formal generalization
Template #2 Template #2a	QUAN>qual Or QUAN+qual	-explanation -validation	
Template #3	quan>QUAL		-guide purposeful sampling -focus information-seeking -suggest analytic paths
Template #4	Qual>QUAN	-generate items, variables -generate hypotheses	
Template #5	Quan wave 1 wave 2 wave 3 Qual on-going field work	-explanation -validation -generate items, variables -generate hypotheses	-measured description -validation -formal generalization -guide purposeful sampling -focus information-seeking -suggest analytic paths
Template #6	Qual>Quan>Qual		-instrumental bridge
Template #7	Quan>Qual>Quan	-fieldwork bridge	

[1]Constructed from information in Miles & Huberman (1994), Morgan (1998), Morse (1991), and Tashakkori & Teddlie (1998).

> indicates sequential relationship

CAPITALS indicate priority

+ indicates concurrent relationship

Arrows suggest a rolling wave

FIGURE 1. Hybrid, combination, or mixed-method design templates.

tense. In the process of sampling for complementarity, researchers will inevitably also obtain information on convergent validity and, thereby, also achieve the purpose of triangulation. That is, in the process of obtaining fuller information on why persons scored as they did, they will also obtain information on whether persons look the same on interview or observation as they did on the quantitative measure of a target phenomenon. Researchers will sample participants in scoring categories until the point of informational redundancy; that is, until they have collected information from enough cases in each scoring category to allow them to draw conclusions about the validity of the result (if they are seeking convergent validity), or to elaborate on and clarify the result. Sampling on the basis of scores is an especially useful strategy in clinical trials of interventions to validate or clarify their effects on different participants.

Random purposeful sampling. Another example of the combined use of probability and purposeful sampling is random purposeful sampling, which may also be used in design template 2 or 3. This sampling strategy is employed when there is a very large pool of potentially information-rich cases and no obvious reason to choose one case over another. For example, in a clinical trial of an intervention to reduce pain with 500 people, 300 of them scored as having less pain on a standardized measure of pain, 150 scored as having as much pain as they had before the treatment, and 50 scored as having more pain. These numbers are too large for any purposeful sampling strategy oriented toward the intensive study of the particulars of each case. Accordingly, cases can initially be chosen from each of these three scoring groups (the criterion for sampling here is again the scores) by assigning all the cases in each group a number

from a random number table and then drawing them in turn. Each case drawn must meet the minimum criterion in all purposeful sampling: namely, that it is an information-rich case. And, only so many cases are drawn in each scoring group that will permit researchers to make the kinds of inferences they wish to make: for example, concerning convergent validity, or fuller description or explanation of cases.

Stratified purposeful sampling. Another kind of combination of sampling techniques is stratified purposeful sampling, where the researcher wants to ensure that certain cases varying on preselected parameters are included. This strategy can potentially be used in any of the design templates shown in Figure 1. Although this kind of sampling is — from a probability sampling standpoint — statistically nonrepresentative (Trost, 1986), it is, from a purposeful sampling standpoint, informationally representative. Each case represents a prespecified combination of variables, the distinctive confluence of which is the focus of study. The researcher wants to explain how these variables come together to make a case the case that it is.

For example, a researcher studying caregiving may wish to ensure that s/he has 1-2 cases that minimally to maximally vary on four parameters: type of caregiving dyad, role of each member of dyad, mortality of illness, and transmissibility of illness. Caregiving dyads include parents caring for their children, adult or minor children caring for their parents, husbands caring for their wives, wives caring for their husbands, heterosexual partners caring for each other, male and female homosexual partners caring for each other, brothers caring for sisters, sisters caring for brothers, grandparents caring for grandchildren, grandchildren caring for grandparents, friends caring for friends, strangers caring for strangers, paid caregivers caring for patients, and so on. In each of these dyads, one or the other member can be the caregiver or the cared-for; accordingly, wives care for or are cared for by husbands. Finally, illnesses can be classified as mortal or nonmortal, and transmissible or nontransmissible.

Accordingly, a researcher may want to have the most variation s/he can achieve on each of these four parameters, or limit variation on one or more of these parameters. In a maximally varied form of purposeful stratified sampling, the researcher might want to have 1–2 cases of every combination of variations s/he can find. The researcher may want to have 1–2 cases of a husband caring for his wife, who suffers from a mortal and nontransmissible illness; 1–2 cases of a husband caring for his wife, who suffers from a nonmortal and nontransmissible illness; 1–2 cases of a husband caring for his wife, who suffers from a mortal and transmissible illness; and, 1–2 cases of a husband caring for his wife, who suffers from a nonmortal and transmissible illness. The researcher may also want to have cases of wives caring for husbands, which show the variations just itemized on illness mortality and transmissibility. In addition, the researcher may want to include these variations for as many of the other kinds of caregiving dyads previously listed as s/he can. Or, the researcher may want to restrict sampling only to one kind of caregiving dyad (eg., parent–child caregivers) varying on the parameters of member role and type of illness.

In short, in stratified purposeful sampling, researchers want to fill each sampling cell with 1-2 cases that exemplify the kinds and degree of variation they surmise are relevant to understanding a target phenomenon, such as caregiving. This kind of sampling typically involves empirical, as opposed to theoretical, cases that represent combinations of demographic (e.g., age, sex, income, and education) and other low-inference variables. Low-inference variables are variables concerning which there is likely to be consensus about what they are. That is, there is likely to be no disagreement on whether a case represents a wife caring for her sick husband. In contrast, an empirical case of Mrs. Jones caring for her husband, Mr. Jones, who is dying from cancer, might also represent a theoretical case of "reluctant caregiving," as opposed to "engaged caregiving." Although researchers can use a stratified purposeful sampling strategy to fill sampling cells with such theoretical cases, they must first have made the case for such theoretical cases. Figure 2 illustrates a stratified purposeful sampling plan restricted to one kind of caregiving dyad (parent–child), but varying on member role and type of illness. A total sample size of 16 would be required to fill each sampling cell with one case. Figure 2 also illustrates why certain sampling cells may be difficult or impossible to fill; few parent–child dyads will be comprised of a minor child caring for her/his own minor ill child.

Combining Data Collection Techniques

Another set of concrete operations at the technique level of research entail the combined use of data collection techniques that are commonly (but not necessarily) associated with either qualitative or quantitative research, such as open-ended and unstructured interviewing and structured questionnaires, respectively. (Some would argue that all data collection techniques in human subjects research, including instruments, are "qualitative" in

Transmissible Illness								Non-Transmissible Illness							
Mortal				Not-mortal				Mortal				Not-mortal			
Ill parent		Ill child		Ill parent		Ill child		Ill parent		Ill child		Ill parent		Ill child	
AC[1]	CC[2]	AC	CC[3]	AC	CC	AC	CC	AC	CC	AC	CC	AC	CC	AC	CC

[1]AC=Adult caregiving child

[2]CC=Child caregiving child.

[3]There will be few parent-child dyads comprised of a minor age child caring for her/his own minor age ill child.

FIGURE 2. Illustration of stratified purposeful sampling plan.

that they involve verbal data, which are only later transformed into numbers.) Researchers' viewing positions will influence how they use these techniques. For example, for many researchers in a positivist viewing position, data collection techniques vary in the degree to which they yield objective data. Observations of behavior are generally thought to be more objective than self-reports of behavior. When a target phenomenon can be observed, observation is often the criterion measure against which self-report is judged. Accordingly, researchers in a positivist viewing position will often seek "more objective" measures to evaluate the validity of "more subjective" measures. Moreover, whenever there is a discrepancy between what participants do and say they do, what observers see participants doing is generally considered a more accurate reflection of reality than self-report. The self-report is typically called into question.

In contrast, from a constructivist viewing position, there is no hierarchy of data collection techniques whereby one technique is judged to yield more objective (or more accurate or truer) data than another. If the results from two data collection techniques do not converge, these results are treated as interpretive opportunities: either to show that no true discrepancy exists or to suggest the phenomenon that accounts for the apparent discrepancy. For example, Silverman (1993) demonstrated how a discrepancy between parents' accounts of their behavior and observations of their behavior in an examining room could be accounted for without resorting to judgments about these parents as unreliable informants. Indeed, he advised against the kind of "simple-minded 'triangulation'" whereby one kind of data is used simply to corroborate or refute the results of another without attention to the "embedded, situated nature of accounts" (p. 200). Helitzer-Allen and Kendall (1992) described how "discrepancies" between data derived from surveys and laboratory tests were resolved by the analysis of ethnographic interviews, which provided them the opportunity to deepen their understanding of the use of antimalarial chemoprophylaxis during pregnancy. This study also shows the inherently multi-technique nature of ethnographic studies.

The Varied Uses of Instruments

Instruments can be used in combination studies to fulfill a variety of objectives. They can be used to provide fuller description of cases in areas suggested by interview or observation data, as, for example, when interview data suggest that participants are depressed and the researcher decides to administer a depression inventory. Each participant's score may be compared to the normative score for that instrument or to the scores of other participants in the study. Instruments are used here to make case-bound or idiographic generalizations —that is, generalizations about each participant in the study—not to make nomothetic generalizations, or generalizations from the study sample to populations. This use of instruments exemplifies design template 1 shown in Figure 1, where the decision to use an instrument is made on analytic grounds developed from data collected in a qualitative study. This use also exemplifies the "development" purpose for mixed-method research, whereby the results of using one kind of data collection technique informs or guides the use of another kind. Researchers may also use instruments for description (as opposed to statistical inference) concurrently with interviews or observations for the purpose of complementarity, as exemplified in design template 1a, or as a development bridge between the qualitative exploration of a target phenomenon and the further explanation of this phenomenon, as shown in design template 6. The development bridge here—the instrument—directs researchers toward whom they will sample

and what data they will collect in the second qualitative portion of a study.

Another use of instruments that fulfills the purpose of development is to guide purposeful sampling. The results of instruments can direct researchers more precisely to the kinds of participants they may wish to recruit and the nature of information they will want to obtain from them. As described previously, instrument scores can be the basis for sampling in a follow-up qualitative study. Results from a survey can direct researchers toward the most fruitful variables and associations to examine further both qualitatively, as shown in design template 3 in Figure 1, or first qualitatively and then again, quantitatively, as shown in design template 7 in Figure 1. Ornstein and his colleagues (1993) initiated focus groups after learning from the results of a telephone survey that there were three groups of "nonresponders" to a reminder letter for cholesterol screening. The focus groups were organized to elicit more information concerning the lack of response. This information could then serve as the basis for revising the reminder protocol and then qualitatively or quantitatively evaluating its effectiveness in reducing the numbers of nonresponders. For example, the new reminder protocol could be tested in a clinical trial and the scores on the instruments used to appraise the protocol could then be used as the basis for criterion sampling to further explain those scores. This combined example illustrates how a quan > Qual study (survey>focus group) can lead to a Quan > qual (experiment > criterion sampling + interviews) study. Together, these studies can be represented as quan>Qual> Quan>qual.

Instruments can also be used as elicitation devices in interviews concerning both the target phenomenon and the instrument itself. For example, a researcher may use participants' responses on a depression inventory to trigger more thoughts and feelings about depression in individual interview sessions. Participants often require some assistance to articulate inchoate thoughts or to speak the unspeakable. The use of an instrument as an elicitation device can serve this purpose, just as the use of projective techniques (e.g., Ornstein et al., 1993) can help participants language, and focus on, a target experience.

Researchers may also ask participants to comment specifically on their view of each item on an instrument in order to appraise its content and construct validity. In the infertility study mentioned previously (Sandelowski et al., 1992), we used a symptom inventory to appraise the type and intensity of symptoms women and men were experiencing as they awaited birth or adoption. One of the symptoms listed was "headache." One waiting father spontaneously mentioned to me that he had experienced a headache from hitting his head on a pipe doing a home repair, but, on the symptom inventory, did not mark headache as one of his symptoms. As this man explained it, he surmised that we were not interested in symptoms that derived from home repairs or other events apparently unconnected to waiting for a child. What this man clarified for me was all of the judgments participants make when they encounter any item on an instrument of which researchers are completely unaware. Accordingly, researchers may want to explicitly use instruments as data elicitation devices to determine exactly what participants see and, therefore, respond to in each item. Because participants may also offer reasons for responding as they did, in addition to explanations for how they and persons like them are likely or ought to respond, researchers may obtain information not only on the content validity of an instrument, but also on its construct validity.

Instrument scores can also be used for qualitative profiling. Qualitative profiling, which is discussed further in the section on "qualitizing," entails theoretically grouping or typing participants according to their scores on two or more instruments. These typologies can be used to guide theoretical sampling in grounded theory studies designed to develop theory.

Combining Data Analysis Techniques

Qualitative and quantitative data sets can be linked, preserving the numbers and words in each data set. Or, these data can be transformed to create one data set, with qualitative data converted into quantitative data, or quantitative data converted into qualitative data (Caracelli & Greene, 1993).

Linking the results of qualitative and quantitative analysis techniques is accomplished by treating each data set with the techniques usually used with that data; that is, qualitative techniques are used to analyze qualitative data and quantitative techniques are used to analyze quantitative data. For example, constant comparison, qualitative content, and narrative analysis techniques are used to analyze interview data, whereas one or more statistical techniques are used to analyze data from instruments. The results of the qualitative analysis of qualitative data and of the quantitative analysis of quantitative data are then combined at the interpretive level of research, but each data set remains analytically separate from the other.

In contrast to the process of linking data are treatments of data that transform one kind of data

into another kind to create one data set. Tashakkori and Teddlie (1998, p. 126) referred to the conversion of qualitative data into quantitative data as "quantitizing," and to the conversion of quantitative data into qualitative data as "qualitizing."

Quantitizing. Quantitizing refers to a process by which qualitative data are treated with quantitative techniques to transform them into quantitative data. The researcher must first reduce verbal or visual data (e.g., from interviews, observations, artifacts, or documents) into items, constructs, or variables that are intended to mean only one thing and that can, therefore, be represented numerically. One of the most commonly used examples of this process (and of design template 4 in Figure 1) is the creation of items for an instrument from interview data. Although the researcher's intention is to "preserve qualitative meaning" in the development of instruments (Fleury, 1993), any one item can only have one meaning to function well psychometrically. A valid item in an instrument cannot be meaning*ful*. That is, it cannot be full of meaning; it cannot mean different things at the same or different times to the same or different people.

Another example of a quantitative transformation aimed at producing a one-meaning unit is the reduction of narrative data to a variable that can be correlated with other variables. In a study exemplifying mixed-method design template 4 in Figure 1, in which quantitative methods prevail but follow and depend on the results of qualitative methods, Borkan, Quirk, and Sullivan (1991) elicited narrative data from elders concerning how they viewed the hip fractures they had suffered. Using a form of narrative analysis, they determined that there were two major narrative emplotments of hip fracture: the mechanical and the organic. They conducted a series of reliability tests to ensure that each narrative was consistently classified as mechanical or organic. They then grouped elders according to whether they had told a largely mechanical or organic story about their hip fracture, and then conducted a correlation study to determine whether and in what direction the emplotments-turned-into-variables — that is, *mechanical narrative* and *organic narrative* — predicted functional outcomes. They found that the organic story predicted poorer outcomes than the mechanical story.

This study is an especially good illustration of the axiom that methods are not uniformly linked to paradigms. Borkan and his colleagues used a narrative methodology for positivist ends. They sought to create a predictor variable by reducing the interview data they had obtained to units with only one meaning and then appropriately used reliability measures to validate them. In contrast, a constructivist using narrative methodology would emphasize the meaningfulness of the data and would treat those data as inherently revisionist and, thus not amenable to reliability testing.

Quantitative treatments of qualitative data can also be used to extract more information from qualitative data, and to confirm researchers' impressions from these data. For example, from interview data with infertile couples, we suspected an association between physicians' encouragement to have an amniocentesis performed and couples' decision to undergo the procedure (Sandelowski, Harris, & Holditch-Davis, 1991). Accordingly, in order to validate our impression, we created a data display to discern the congruence between decision and advice by counting the numbers of couples who were: (a) encouraged to have the procedure and had it ($n = 7$); (b) encouraged to have the procedure but did not have it ($n = 3$); (c) not encouraged or discouraged to have the procedure and did not have it ($n = 7$); and, (d) not encouraged or discouraged to have the procedure but did have it ($n = 3$). The display visually confirmed our impression of an association between the decision to have the procedure and the physician's advice. We then analyzed these data using Fisher's exact probability test, which showed a nonsignificant statistical relationship ($N = 20$; $p = .078$), which might reach significance in another study with a larger sample size.

Qualitizing. Qualitizing refers to a process by which quantitative data are transformed into qualitative data. As in quantitizing, qualitizing can be used to extract more information from quantitative data, or to confirm interpretations of it.

An example of this process is the use of scores on instruments to profile participants — to create verbal portraits or typologies of them — around target phenomena. Tashakkori and Teddlie (1998, pp. 130–133) described five kinds of narrative or qualitative profiling: modal, average, comparative, normative, and holistic. A modal profile is a verbal description of a group of participants around the most frequently occurring attributes. For example, if most of the participants in a group are in their 50s, the group can be described as middle-aged. If most of the participants in a group score in a certain range on a depression inventory, they can be described as mildly or severely depressed. Similarly, an average profile is a verbal description of a group of participants around the mean of an attribute. Although not usually identified in research reports as qualitative profiling, this kind of qualitizing is commonly done to describe samples and to interpret research results.

A comparative profile is a verbal description based on the comparison of participants to each other on one or more sets of scores. In contrast, a normative profile is a verbal description based on the comparison of participants' scores to the normative scores for one or more instruments. Both of these kinds of qualitative profiling depend on the results of quantitative cluster analyses, which entail a set of statistical techniques to identify homogeneous groups of subjects, or distinctive subgroups in which subjects may be placed because they are more similar to each other than to subjects in other subgroups. For example, Rothert and her colleagues (1990) identified four groups of women on the basis of their responses to eight scenarios concerning hormone replacement therapy. That is, they created a typology of information use that showed how differently people can respond to the same sets of information, and that these differences could be clustered in ways that distinguished women from each other. After they quantitatively identified these groups, they interviewed three women from each of these groups for further clarification and to confirm their typology. Their work is an example of design template 2 in Figure 1.

Qualitizing entails further studying these theoretical groups to explicate the features of members in each group that make them like each other and of the groups themselves that make them different from each other, and to choose a name that will capture these features. This kind of investigation lends itself especially well to subsequent grounded theory study, as the clusters can provide the basis for theoretical sampling and for typology development. Researchers can then validate these groupings with further qualitative study or quantitative study, as exemplified in design templates 5 or 7.

Finally, a holistic profile is a verbal description based on impressions rather than specific attributes or scores. Holistic profiles may also be comprised of various combinations of modal, mean, comparative, and normative profiles.

CONCLUSION

Mixed-method research is a dynamic option for expanding the scope and improving the analytic power of studies. When done well, mixed-method studies dramatize the artfulness and versatility of research design. Mixed-method research operationally includes an almost limitless array of combinations of sampling, and data collection and analysis techniques, only a few of which could be described here. Indeed, closer examination of research reports will show researchers have long been using combinations of qualitative and quantitative techniques, but the qualitative techniques have not been featured, have been inappropriately or inadequately used or explained, or have been explicitly minimized. Sutton (1997) lauded the "virtues of closet qualitative research" whereby researchers "conceal or downplay" (p. 97) the use of qualitative techniques in order to have their research reports published!

The combined use of qualitative and quantitative techniques will inevitably be informed by the researcher's viewing position, which shapes what techniques will be combined, and how and why they are combined. Accordingly, if researchers want to combine "stories and numbers. . . without compromise" (Ford-Gilboe, Campbell, & Berman, 1995), use "numbers and words" in a "shamelessly eclectic" manner (Rossman & Wilson, 1994), and ease the "uneasy alliance" (Buchanan, 1992) between qualitative and quantitative methods, they must have a clear view of their viewing positions and what dynamic mixes they suggest or permit.

Researchers must also resist the "'mix and match syndrome'" (Leininger, 1994, p. 103) that can result in a "qualitative quagmire" (Barbour, 1998) and that mandates that all research be mixed-method research. As Chen (1997) suggested, mixed-method research is in danger of becoming the new ideology. Mixed-method research should never be used because of the misguided assumptions that more is better, that it is the fashionable thing to do, or, most importantly, that qualitative research is incomplete without quantitative research (Morse, 1996). Indeed, qualitative techniques have been used to "salvage" quantitative studies (Weinholtz, Kacer, & Rocklin, 1995). The "completeness" of any individual study, no matter what kind it is, must be judged without resorting to methodological fads or fetishes.

REFERENCES

Annells, M. (1996). Grounded theory method: Philosophical perspectives, paradigm of inquiry, and postmodernism. Qualitative Health Research, 6, 379–393.

Barbour, R.S. (1998). Mixing qualitative methods: Quality assurance or qualitative quagmire? Qualitative Health Research, 8, 352–361.

Berman, H., Ford-Gilboe, M., & Campbell, J.C. (1998). Combining stories and numbers: A methodologic approach for a critical nursing science. ANS: Advances in Nursing Science, 21(1), 1–15.

Borkan, J.M., Quirk, M., & Sullivan, M. (1991). Finding meaning after the fall: Injury narratives from elderly hip fracture patients. Social Science & Medicine, 33, 947–957.

Buchanan, D.R. (1992). An uneasy alliance: Combining qualitative and quantitative research methods. Health Education Quarterly, 19, 117–135.

Caracelli, V.J., & Greene, J.C. (1993). Data analysis strategies for mixed-method evaluation designs. Educational Evaluation and Policy Analysis, 15, 195–207.

Caracelli, V.J., & Greene, J.C. (1997). Crafting mixed-method evaluation designs. In J.C. Greene & V.J. Caracelli (Eds.), Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms (pp. 19–32). San Francisco: Jossey-Bass.

Chen, H. (1997). Applying mixed methods under the framework of theory-driven evaluations. In J.C. Greene & V.J. Caracelli (Eds.), Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms (pp. 61–72). San Francisco: Jossey-Bass.

Fleury, J. (1993). Preserving qualitative meaning in instrument development. Journal of Nursing Measurement, 1, 135–144.

Ford-Gilboe, M., Campbell, J., & Berman, H. (1995). Stories and numbers: Coexistence without compromise. ANS: Advances in Nursing Science, 18, 14–26.

Greene, J.C., & Caracelli, V.J., (Eds.). (1997a). Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. San Francisco: Jossey-Bass.

Greene, J.C., & Caracelli, V.J. (1997b). Defining and describing the paradigm issue in mixed-method evaluation. In J.C. Greene & V.J. Caracelli (Eds.), Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms (pp. 5–17). San Francisco: Jossey-Bass.

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.

Guba, E.G., & Lincoln, Y.S. (1994). Competing paradigms in qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds.), Handbook of qualitative research (pp. 105–117). Thousand Oaks, CA: Sage.

Helitzer-Allen, D.L., & Kendall, C. (1992). Explaining differences between qualitative and quantitative data: A study of chemoprophylaxis during pregnancy. Health Education Quarterly 19, 41–54.

Heron, J., & Reason, P. (1997). A participatory inquiry paradigm. Qualitative Inquiry, 3, 274–294.

Leininger, M. (1994). Evaluation criteria and critique in qualitative research studies. In J.M. Morse (Ed.), Critical issues in qualitative research methods (pp. 95–115). Thousand Oaks, CA: Sage.

Miles, M.B., & Huberman, A.M. (1994). Qualitative data analysis: An expanded sourcebook (2nd ed.). Thousand Oaks, CA: Sage.

Morgan, D.L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. Qualitative Health Research, 8, 362–376.

Morse, J.M. (1991). Approaches to qualitative-quantitative methodological triangulation. Nursing Research, 40, 120–123.

Morse, J.M. (1996). Is qualitative research complete? Qualitative Health Research, 6, 3–5.

Ornstein, S.M., Musham, C., Reid, A., Jenkins, R.G., Zemp, L.D., & Garr, D.R. (1993). Barriers to adherence to preventive services reminder letters: The patient's perspective. Journal of Family Practice, 36, 195–200.

Patton, M.Q. (1990). Qualitative evaluation and research methods (2nd ed.). Newbury Park, CA: Sage.

Pearce, J.C. (1971). The crack in the cosmic egg: Challenging constructs of mind and reality. New York: Washington Square Press.

Rossman, G.B., & Wilson, B.L. (1994). Numbers and words revisited: Being "shamelessly eclectic." Quality & Quantity, 28, 315–327.

Rothert, M., Rovner, D., Holmes, M., Schmitt, N., Talarczyk, G., Kroll, J., & Gogate, J. (1990). Women's use of information regarding hormone replacement therapy. Research in Nursing & Health, 13, 355–366.

Sandelowski, M. (1995). Triangles and crystals: On the geometry of qualitative research. Research in Nursing & Health, 18, 569–574.

Sandelowski, M., Harris, B.G., & Holditch-Davis, D. (1991). Amniocentesis in the context of infertility. Health Care for Women International, 12, 167–178.

Sandelowski, M., Holditch-Davis, D., & Harris, B.G. (1992). Using qualitative and quantitative methods: The transition to parenthood of infertile couples. In J.F. Gilgun, K. Daly, & G. Handel (Eds.), Qualitative methods in family research (pp. 301–322). Newbury Park, CA: Sage.

Silverman, D. (1993). Interpreting qualitative data: Methods for analyzing talk, text and interaction. London: Sage.

Sutton, R.I. (1997). The virtues of closet qualitative research. Organization Science, 8, 97–106.

Swanson, S.C. (1992). A cross-disciplinary application of Greene, Caracelli and Grahams's conceptual framework for mixed method evaluation. Unpublished doctoral dissertation, Georgia State University, Atlanta, GA.

Tashakkori, A., & Teddlie, C. (1998). Mixed methodology: Combining qualitative and quantitative approaches. Thousand Oaks, CA: Sage.

Trost, J.E. (1986). Statistically non-representative stratified sampling: A sampling technique for qualitative studies. Qualitative Sociology, 9, 54–57.

Weinholtz, D., Kacer, B., & Rocklin, T. (1995). Salvaging quantitative research with qualitative data. Qualitative Health Research, 5, 388–397.

Wolfer, J. (1993). Aspects of "reality" and ways of knowing in nursing: In search of an integrating paradigm. Image: Journal of Nursing Scholarship, 25, 141–146.