Why Triangulate?

SANDRA MATHISON

This article discusses triangulation as a strategy for increasing the validity of evaluation and research findings. Typically, through triangulating we expect various data sources and methods to lead to a singular proposition about the phenomenon being studied. That this is not the case is obvious to most researchers and evaluators. Given that this expectation is unrealistic, an alternative perspective of triangulation is presented. This alternative perspective takes into account that triangulation results in convergent, inconsistent, and contradictory evidence that must be rendered sensible by the researcher or evaluator.

Good research practice obligates the researcher[1] to triangulate, that is, to use multiple methods, data sources, and researchers to enhance the validity of research findings. Regardless of which philosophical, epistemological, or methodological perspectives an evaluator is working from, it is necessary to use multiple methods and sources of data in the execution of a study in order to withstand critique by colleagues. The experimentally inclined are enjoined to use qualitative research methods to help conceptualize their studies and ethnographers are often expected to conduct surveys to corroborate observational data. We all agree that triangulation is a good thing and that research and evaluation will be improved by such a practice. The issue explored in this paper is how triangulation actually improves our research practice, and what the reasons are for employing triangulation.

Triangulation is typically perceived to be a strategy for improving the validity of research or evaluation findings: "...triangulation is supposed to support a finding by showing that independent measures of it agree with it or, at least, don't contradict it" (Miles & Huberman, 1984, p. 235). It is essentially a strategy that will aid in the elimination of bias and allow the dismissal of plausible rival explanations such that a truthful proposition about some social phenomenon can be made (Campbell & Fiske, 1959; Denzin, 1978; Webb, Campbell, Schwartz, & Sechrest, 1966).

Historically, triangulation is a new concept in the social science repertoire dating back to a paper published by Campbell and Fiske in 1959. In this paper, Campbell and Fiske discuss establishing validity of measures through the application of a multitrait-multimethod matrix, a procedure which examines both convergent and discriminant validation of measures of traits. While the procedure is presented in a mathematically elegant fashion, its basic idea is that in the development of measures of psychological traits, several methods should be employed to measure several traits simultaneously. Through correlational analyses the independence of methods *and* traits can be established.

Triangulation has arisen as an important methodological issue in the evaluation literature as well. In particular, naturalistic and qualitative approaches to evaluation have demanded attention to controlling bias and establishing valid propositions because traditional scientific techniques are incompatible with these alternate epistemologies. Guba and Lincoln (1981) discuss strategies such as prolonged engagement in a site, peer debriefing, and establishing structural corroboration as means for improving the credibility of evaluation findings. Establishing structural corroboration, which is "...a process of gathering data or information and using it to establish links that eventually create a whole that is supported by the bits of evidence that constitute it" (Eisner, 1979, p. 215), can be, they suggest, greatly enhanced by using triangulation. Patton (1980) also emphasizes the centrality and problematic nature of triangulation. "There is no magic in triangulation. The evaluator using different methods to investigate the same program should not expect that the findings generated by those different methods will automatically come together to produce some nicely integrated whole" (Patton, p. 330). The point of triangulation, he suggests, is "...to study and understand when and why there are differences" (p. 331). Smith and Kleine (1986) suggest that the use of multimethods results in "different images of understanding" thus increasing the "potency" of evaluation findings. The prominence of triangulation in discussions of naturalistic and qualitative evaluation suggest both the importance and problematic nature of this methodological technique.

Although Campbell and Fiske (1959) introduced the idea of using multiple methods, Webb et al. (1966) coined the term "triangulation" in their treatise on nonreactive measures in the social sciences. Their discussion centers around the establishment of the validity of propositions which they claim could be aided by using a variety of methods, and particularly nonreactive measures. If Webb et al. (1966) labelled this new technique, it was Denzin (1978) who provided a detailed discussion of how to triangulate. In his explication of how to use triangulation as a research strategy, Denzin outlines four types of triangulation: (a) data triangulation including time, space, and per-

SANDRA MATHISON *is Evaluation Coordinator, University of Chicago School Mathematics Project, Rm. 9, Judd Hall, 5835 S. Kimbark Ave., Chicago, IL 50637. She specializes in program evaluation and qualitative research methods.*

son, (b) investigator triangulation, (c) theory triangulation, and (d) methodological triangulation (pp. 294–307). In actual fact, Denzin only seriously suggests three types of triangulation since the notion of theoretical triangulation is problematic at best, and likely impossible in reality. Even Denzin questions the plausibility of such a notion in the notes at the end of the chapter.

> My use of theoretical triangulation must in no way be construed as a defense of ecelecticism. Indeed, sociologists committed to a given perspective will probably not employ theoretical triangulation. The great value of this strategy, as I see it, however, is its assurance that no study will be conducted in the absence of some theoretical perspective. In this sense it is most appropriate for the theoretically uncommitted, as well as for analysis of areas characterized by high theoretical incoherence. (p. 307)

Data triangulation refers simply to using several data sources, the obvious example being the inclusion of more than one individual as a source of data. However, Denzin expands the notion of data triangulation to include time and space based on the assumption that understanding a social phenomenon requires its examination under a variety of conditions. So, for example, to study the effect of an inservice program on teachers, one should observe teachers at different times of the school day or year and in different settings such as the classroom and the teachers' lounge.

Investigator triangulation, which involves more than one investigator in the research process, is also considered good practice. This perhaps more than other types of triangulation is usually built into the research process because most studies simply require more than one individual to accomplish the necessary data collection. However, the decision about who these multiple researchers should be and what their roles should be in the research process is problematic (Denzin, 1978; Miles, 1982). How much hands-on data collection the principal investigator needs to do in order to analyze the data, and how much data analysis is relegated to field workers because much of the analysis occurs as data are collected, are both relevant and not easily answered questions.

Methodological triangulation is the most discussed type of triangulation and refers to the use of multiple methods in the examination of a social phenomenon. Psychologists have long used Denzin's notion of within-method triangulation in assessing psychological traits. Multiple scales comprise a psychological assessment such as an intelligence test in an effort to assess the different aspects of intelligence. The lie detector scale in some psychological inventories is another example. Denzin suggests that the within-methods triangulation approach has limited value, because essentially only one method is being used, and finds the between-methods triangulation strategy more satisfying. Other researchers seem to follow this lead and focus primarily on between methods triangulation. "The rationale for this strategy is that the flaws of one method are often the strengths of another: and by combining methods, observers can achieve the best of each while overcoming their unique deficiences" (Denzin, 1978, p. 302). It is with this type of triangulation that Denzin relies most heavily on the work of Webb et al. (1966) to suggest that the use of appropriate multiple methods will result in more valid research findings.

The discussions of triangulation as a research strategy are based on some fundamental assumptions that should be explicated in order to understand why a different conception of triangulation will be introduced later in this paper. First is the assumption that the bias inherent in any particular data source, investigator, and particularly method will be cancelled out when used in conjunction with other data sources, investigators, and methods. The second, and related, assumption is that when triangulation is used as a research strategy the result will be a convergence upon the truth about some social phenomenon.

The assumption that bias will be eliminated in a multimethod research design is a puzzling one that frequently goes unexamined. Jick (1983, p. 138) summarizes the assumption as follows: "...triangulation purports to exploit the assets and neutralize, rather than compound, the liabilities." Jick recognizes the problematic nature of this assumption particularly as it relates to his use of triangulation in a study of the effects of a merger on company employees. Denzin (1978) explicates the value of five different methods (experiment, survey, participant observation, unobtrusive methods, and historical methods) given his theoretical perspective and the factors associated with internal and external validity according to Campbell and Stanley. This discussion suggests that different methods produce different understandings of a social phenomenon but Denzin does not address how these differences are to be reconciled. And the discussion certainly does not justify the assumption that these differences constitute bias which will be cancelled out in an overall design.

It would seem that methods are considered to be subjective in much the same way that individuals are subjective, that is, individuals make claims which are the sole property of that person. Scriven's (1972) distinction between the quantitative and qualitative senses of objectivity illustrate some of the confusion here. In the quantitative sense of objectivity, an individual's view is considered to be subjective and the collective view of many individuals is an objective one. In the qualitative sense of objectivity, "[b]eing objective means that the observation is factual, while being subjective means that the observation is biased in some way" (House, 1980, p. 86). The quantitative notion of objectivity seems to underlie discussions of triangulation at the expense of the qualitative sense of objectivity. So the use of any single method, just like the view of any single individual, will necessarily be subjective and therefore biased. Triangulation is the methodological counterpart to intersubjective agreement and, just as with individuals, reliability is confused with validity. The evidence produced by different methods might be different because of bias in the measures but it is also possible that different methods tap different domains of knowing. Until the argument that bias cancels itself out in a multimethod design is more fully explicated, there seems little reason to pursue a triangulation strategy based on this assumption.

The assumption that the bias in methods, data sources, and investigators cancels out permits the assumption that what is left is the truth about what is investigated. This assumption suggests that when a triangulation strategy is used the result will be *convergence* on a single perspective

of some social phenomenon. And, if one examines the metaphors in discussions of triangulation it is apparent that convergence on a point or object is the desired goal. For example, in the commonly used navigational metaphor the process involves using multiple reference points to determine the location of yet another point, usually a place or object. Military strategists use such a model in locating enemy targets. So, given certain information one can accurately (and accuracy is very important if one wants to destroy the enemy before they destroy you) locate an object. Extending this metaphor to social phenomena, the researcher (navigator or military strategist) can use several kinds of information to determine the truth (the location) about some social phenomenon (an island or the enemy). This does not suggest that the metaphor itself makes researchers adhere to this assumption, but rather the metaphor used reflects the assumption and in its use reinforces the assumption.

Some researchers, and particularly those with a qualitative bent, have begun to question the metaphors used because actual research practice does not conform to the anticipated outcomes (Jick, 1983; Miles & Huberman, 1984). Miles and Huberman (pp. 234-235) suggest alternative metaphors such as detective work and the work of car mechanics and general practitioners. Using these forms of work as metaphors suggests a different process for triangulating involving intuition, an ever-present degree of uncertainty, and purposive human action. These are indeed more descriptive metaphors which may or may not become dominant in future methodologial discussions.

An Alternative Conception of Triangulation

The second part of this paper will outline an alternative reason for using a triangulation strategy which does not rely on the assumptions discussed above. The examples in this part of this paper are based on the evaluation of a large scale mathematics curriculum and teacher development project. This project provided ample opportunity to employ triangulation, and while the outcome was occassionally convergence upon a single conclusion, more frequently the outcome was inconsistent findings or, on some occasions, contradictory findings. I suspect the degree to which these alternative outcomes occurred in this particular evaluation, which actually included four different projects, reflects their occurrence generally.

In practice, triangulation as a strategy provides a rich and complex picture of some social phenomenon being studied, but rarely does it provide a clear path to a singular view of what is the case. I suggest that triangulation as a strategy provides evidence for the *researcher* to make sense of some social phenomenon, but that the triangulation *strategy* does not, in and of itself, do this.

Because of the predominance of the assumption that triangulation will result in a single valid proposition we look for the convergence of evidence and miss what I see as the greater value in triangulating. More accurately, there are three outcomes that might result from a triangulation strategy. The first is that which is commonly assumed to be the goal of triangulation and that is *convergence.* The notion of convergence needs little explanation: data from different sources, methods, investigators, and so on will provide evidence that will result in a single proposition about some social phenomenon.

A second and probably more frequently occurring outcome from a triangulation strategy is *inconsistency* among the data. When multiple sources, methods, and so on are employed we frequently are faced with a range of perspectives or data that do not confirm a single proposition about a social phenomenon. Rather, the evidence presents alternative propositions containing inconsistencies and ambiguities. With this outcome it is not clear what the valid claim or proposition about something is.

A third outcome is *contradiction.* It is possible not only for data to be inconsistent but to actually be contradictory. When we have employed several methods we are sometimes left with a data bank that results in opposing views of the social phenomenon being studied.

If one were to accept the assumptions that triangulation should result in a single claim because bias is naturally cancelled out, outcomes of the second and third type would not be useful in the research process. We all know, of course, this is not the case. We do, in fact, utilize not only convergent findings but also inconsistent and contradictory findings in our efforts to understand the social phenomena that we study. The value of triangulation is not as a technological solution to a data collection and analysis problem, it is as a technique which provides more and better evidence from which researchers can *construct meaningful propositions* about the social world. The value of triangulation lies in providing evidence such that the researcher can construct explanations of the social phenomena from which they arise.

This section of the paper, drawing on the evaluation of the mathematics project previously mentioned, gives an example of each of the three possible outcomes from a triangulation strategy. The outcomes are not, however, an end in themselves. The researcher is left with the task of making sense of the evidence *regardless* of what the outcome is. So whether the data converge, are inconsistent, or are contradictory the researcher must attempt to construct explanations for the data and about the data. Following each of the examples is a possible, and hopefully plausible, explanation for the data presented. The value of triangulation lies in providing evidence—whether convergent, inconsistent, or contradictory—such that the researcher can construct good explanations of the social phenomena from which they arise.

- ***Convergence:*** When data from different sources or collected from different methods agree, the outcome is convergence.

Example: A teacher development project attempted, among other goals, to increase the amount of time teachers spent teaching mathematics. Teachers' self-reports at the end of the year indicate, on average, they taught math for 48 minutes per day and the average length of observed math classes was 45 minutes.

Explanation: We were skeptical about the validity of either teacher self-reports (teachers often responded to such questions as time by giving district guidelines) or classroom observations (we felt teachers may have extended their math instruction because we were observing their class). The fairly close estimates of time spent on math may suggest that such skepticism was ill-founded. Both estimates indicate a small increase in time spent on math from the previously spent self-reported average of 38 minutes per day. The validity of the proposition is also enhanced by

knowing that the district time allotment for math instruction (30 minutes) was less than either of these estimates.

- ***Inconsistency:*** The data obtained through triangulation may be inconsistent, not confirming but not contradictory.

Example: Primary teachers were provided with a set of cards containing short math activities which could be done during transitional and typically non-instructional times of the day. Teachers reported using the activity cards extensively but in over 200 classroom observations only 14 such activities occurred.

Explanation: An attempt to explain the inconsistency of these results requires going beyond the data obtained from these methods. We trusted our observational data *and* the teachers' self-reports; these two sources of data and data collection methods also resulted in convergence on other issues. By delving into the background understanding the evaluator acquires during an extended and intimate interaction with project participants it is possible to construct a plausible explanation of these inconsistent findings. If one separates teachers' actions from their thoughts a plausible explanation might be as follows. As an abstract idea, teachers were very positively disposed toward using non-instructional time for math activities, a sentiment expressed in many ways throughout the school year. However, in practical terms, such an idea is not so easily enacted. For one thing, teachers would have no unplanned time (even a short activity shares many of the characteristics of a complete lesson) and time spent on other schooling goals would be usurped. So, teachers support the idea through their responses, but have difficulty implementing the idea as we observed in classrooms.

- ***Contradictory:*** At times, data are not simply inconsistent but are actually contradictory, leading the researcher to incommensurable propositions.

Example: An evaluation of a math curriculum was conducted in seventh, eighth, and ninth grade classes and because of the model of instruction inherent in the curriculum, pace and content coverage were important issues. As one moves from seventh to eighth to ninth grade, the pace at which the teachers reported covering the material slows from one lesson per day to one lesson every two days. When teachers indicated how many of the total 13 chapters were covered, the median number for ninth grade was 12, for eighth grade it was 11, and seventh grade classes got about half way through the 11th chapter.

Explanation: Typically, the seventh and eighth grade classes in this evaluation study were average to above average ability level groups whereas ninth grade classes were average to below average ability level groups. Knowing this provides a way of making sense of the pace variation across grades; teachers with lower ability students take longer to cover the material. Why then did teachers with the lower ability students cover more of the total curriculum? One might postulate that there are fundamental differences among seventh, eighth, and ninth grade teachers and combined with the variance in the type of students, it is possible to construct an explanation that accomodates these contradictory data. In these ninth grade classes teachers proceeded more slowly because many of the low ability students will perhaps take one more math course, likely a consumer math or applications course. The need to ensure that students fully comprehended the concepts was minimized by the future math aspirations for these students; little time was spent on students obtaining full mastery of the curriculum. For the seventh and eighth grade classes, on the other hand, there was a preoccupation with ensuring mastery of the content by the higher ability students since many will proceed to a college-prep math course sequence. It appeared that more time was taken for quizzing and enrichment activities in these classes, and therefore they did not get as far into the curriculum as ninth grade classes.

How is it that the preceding explanations are derived? The traditional view of triangulation would have us put forward propositions about the social world that all of our data sources, methods, and so on agree on. By doing this, one would necessarily be unduly restricted in making quite valid claims about social phenomena. One is not, nor should one be, shackled by the immediately available empirical evidence in rendering some social phenomenon understandable.

A close examination of the explanations offered here suggests that several levels of evidence are required for the researcher to construct plausible explanations. There are obviously the data on hand. There is, also, a holistic understanding of the project itself, its history, the intentions of the developers, the ongoing relationships within the project, and so on. This understanding about a project or program is frequently unarticulated, a part of that vast body of tacit knowledge that we all have. And lastly, the researcher/evaluator has a store of knowledge and understandings about the social world which allows such projects and evaluations of them to exist. In the examples given this is a store of information about how schools operate, what schooling goals are, how classrooms operate, what teachers do, and so on. It takes all of these levels to provide good explanations around the data collected through triangulation strategies.

It is a mistake to assume that only inconsistent and contradictory data need to be explained by the application of these various levels of knowledge. All of the outcomes of triangulation, convergent, inconsistent, and contradictory, need to be filtered through knowledge gleaned from the immediate data, the project/program context, and understandings of the larger social world. So, the example of convergent data, the amount of time teachers spend teaching mathematics, requires explanation in much the same way as the contradictory evidence about pacing and content coverage in the third example. It will ordinarily, but probably not always, be easier to construct meaningful explanations in cases where the evidence is convergent.

A crucial question arises from this reformulation of the value and purpose of triangulation: By what standards should the explanations constructed by the researchers be judged? Miles and Huberman (1984) suggest

> . . . triangulation is a state of mind. If you self-consciously set out to collect and double-check findings, using multiple sources and modes of evidence, the verification process will largely be built into the data-gathering process, and little more need be done than to report on one's procedures. (p. 235)

As I have suggested, triangulation extends beyond data collection but Miles and Huberman's comment is still instructive because it obligates the researcher to be explicit, as much as possible, about the research process. Not only

must the researcher report his or her data collection procedures but also the three levels of information from which explanations about social phenomena are constructed. If these levels of information are explicated, the logic and plausibility of explanations are public and open to discussion, a minimal criterion for social science research. Without fully developing the criteria which could be applied to such explanations, one would certainly be concerned about the quality of the data, plausibility, coherence, accommodation of counter-factual evidence, and perhaps predictive ability.

Why Triangulate?

This alternate conception of the value of triangulation explicates problems that previously existed but were unarticulated. Practicing researchers and evaluators know that the image of data converging upon a single proposition about a social phenomenon is a phantom image. More realistically, we end up with data that occasionally converge, but frequently are inconsistent and even contradictory. And we do not throw our hands up in despair because we cannot say anything about the phenomenon we have been studying. Rather, we attempt to make sense of what we find and that often requires embedding the empirical data at hand with a holistic understanding of the specific situation and general background knowledge about this class of social phenomena. This conception shifts the focus on triangulation away from a technological solution for ensuring validity and places the responsibility with the researcher for the construction of plausible explanations about the phenomena being studied.

The examples in this paper are taken from the evaluation of the University of Chicago School Mathematics Project which is supported by grants from the Amoco, Carnegie, and National Science Foundations. An earlier version of this paper was presented at the 1986 American Evaluation Association meeting, Kansas City. The comments of James Burry and an anonymous reviewer are appreciated.

---

[1]For readibility, *researcher* is used throughout this paper but the argument applies equally to research and evaluation, as it also applies equally to quantitative and qualitative researchers and evaluators.

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105.

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

Eisner, E.W. (1979). *The educational imagination.* New York: Macmillan.

Guba, E.G., & Lincoln, Y.S. (1981). *Effective evaluation.* San Francisco: Jossey-Bass.

House, E.R. (1980). *Evaluating with validity.* Beverly Hills: Sage.

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

Miles, M.B. (1982). A mini-cross-site analysis. *American Behavioral Scientist, 26*(1), 121-132.

Miles, M.B., & Huberman, A.M. (1984). *Qualitative data analysis.* Beverly Hills: Sage.

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

Scriven, M. (1972). Objectivity and subjectivity in educational research. In L.G. Thomas (Ed.), *Philosophical redirection of educational research.* National Society for the Study of Education.

Smith, M.L., & Kleine, P.L. (1986). Qualitative research and evaluation: Triangulation and multimethods reconsidered. In D.D. Williams (Ed.), *Naturalistic evaluation* (New Directions for Program Evaluation, No. 30). San Francisco: Jossey-Bass.

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

AERA Summer Institutes

Alternatives to Recitation and Teacher Dominated Interaction: The Instructional Conversation and Responsive Questioning

Directors: Ronald Gallimore, *University of California, Los Angeles;* Roland Tharp, *University of Hawaii, Honolulu;* Stephanie Dalton, *Kamehameha Elementary Education Program, Honolulu;* Robert Rueda, *University of Southern California*

San Diego, CA; **July 15–17;** $215 members, $250 nonmembers

Recitation and other forms of teacher-dominated interaction are the norm in U.S. schools. Recitation is useful to determine what students already know, but it does little to teach them what they need to learn. Good teaching occurs when teachers assist students to perform at levels they cannot achieve alone. The workshop will describe and demonstrate classroom teaching that "assists performance." The presentation is based on a 20-year program of research and development, the Kamehameha Elementary Education Project (KEEP).

A significant means of teacher assistance is responsive questioning in an "instructional conversation" through which teachers assist students to use higher order skills to understand text and perform tasks. The instructional conversation and responsive questioning require highly professional teaching skill. The workshop will present a description of a classroom environment which supports such teaching. Topics include ways trainers can assist teacher performance; case study of a teacher who, by having her own performance assisted, becomes competent in assisting the performance of students; and the issues in assisting the assistors of teachers—the next level in the chain of teaching including program development, assessment, evaluation, and administration.

Analyzing Multi-Level Longitudinal and Educational Data: Applications and Techniques

Directors: Anthony Bryk, *University of Chicago;* Stephen Raudenbush, *Michigan State University*

Keystone, CO; **July 28–30;** $215 members, $250 nonmembers

As a result of recent developments in the theory and application of hierarchical linear models (HLM), a satisfactory solution to two methodological problems, the measurement of change and the assessment of multi-level effects, now exists. This training session will introduce participants to the statistical theory of hierarchical linear models and demonstrate its application to a variety of research problems including the measurement of change, the study of school effects, and research synthesis. The primary purpose of this session is to enable participants to formulate educational research questions in terms of HLM and to pose and test hypotheses within this framework. The session has been designed to accommodate participants with a range of methodological backgrounds; however, knowledge of multiple regression is necessary.

HLM and HLM3, statistical software for the analyses of two- and three-level models, will also be illustrated with a portable micro-computer. Participants will have an opportunity to use these interactive programs during the workshop and will be provided with a draft of a monograph on applications of HLM, statistical theory, and data analysis advice.

Please contact AERA for application:
1230 17th St. NW, Washington, DC; 202-223-9485