Analytic Induction as a Qualitative Research Method of Analysis

Donald E. Ratcliff

(The University of Georgia, 1994)

Analytic induction is a method of data analysis described by Florian Znaniecki (1934) who named the method and systematized many of the associated ideas. However Znaniecki was careful to note that the essence of analytic induction has been used repeatedly throughout history (pp. 236-237), particularly by scientists in the physical sciences (he cites numerous examples from physics and biology). That essence involves " . . . inducing laws from a deep analysis of experimentally isolated instances" (p. 237).

Analytic induction can be contrasted with defining and using terms in advance of research (p. 240). Instead, definitions of terms are considered hypotheses that are to be tested (p. 241). Inductive, rather than deductive, reasoning is involved, allowing for modification of concepts and relationships between concepts occurs throughout the process of doing research, with the goal of most accurately representing the reality of the situation.

The goal of research is making universal statements that may need to be modified later if exceptions are discovered (pp. 232-233), but ultimately can reflect fairly exhaustive knowledge of what is researched (pp. 249, 274-275). Causation is a potential goal of such knowledge, although it is causation that can include numerous exceptions (p. 305). Those exceptions, however, add to the base of knowledge as the generalizability of the construct is determined, and a new, more comprehensive law of behavior can be generated that accounts for the exception (p. 306, 272-273). Such laws are marked by comprehensiveness and parsimony that bring meaning to the various parts of the whole of a phenomenon (pp. 257-258). Yet no analysis can be considered final, since reality is "inexhaustible" and constantly changing (p. 256). The net result of attempts at comprehensive laws of human behavior, suggest Znaniecki, is a creativity and inventiveness in the social sciences that has more commonly marked discoveries in the natural sciences (p. 295). [It is important to note that most advocates of analytic induction have repudiated Znaniecki's implication that certain and absolute laws of human behavior can be attained (Robinson, 1953).] The essential characteristics of a system must be distinguished from irrelevant details that may co-occur with that system (p. 252). Furthermore it is not sufficient to state those characteristics or constructs in isolation from one another, they must be interrelated to one another through comprehensive, logical theories and classifications (p. 253). Classifications are theories, they are not preliminary to theory (p. 254).

Znaniecki (1934, p. 261-262) lists several steps in the analytic induction process: 1) develop a hypothetical statement drawn from an individual instance, 2) compare that hypothesis with alternative possibilities taken from other instances. Thus the social system provides categories and classifications, rather than these being imposed upon the social system (p. 264). Progress in the social sciences is escalated further by comparing aspects of a social system with similar aspects in alternative social systems (p. 265). The emphasis in the process is upon the whole, even though elements are analyzed as are relationships between those elements (pp. 266-267). It is not necessary that the specific case being studied be "average" or representative of the phenomena in general. What is crucial, rather, is that a given case have the essential characteristics, that it function as a pattern or mold by which future cases are defined (p. 251; also see Robinson, 1951). A decade and a half later Cressey (cited by Robinson, 1951) summarized Znaniecki's rather complex and detailed description of analytic induction in the form of six steps. These are: 1) a phenomenon is defined in a tentative manner, 2) a hypothesis is developed about it, 3) a single instance is considered to determine if the hypothesis is confirmed, 4) if the hypothesis fails to be confirmed either the phenomenon is redefined or the hypothesis is revised so as to include the instance examined, 5) additional cases are examined and, if the new hypothesis is repeatedly confirmed, some degree of certainty about the hypothesis results, and 6) each negative case requires that the hypothesis be reformulated until there are no

exceptions. Robinson (1951) emphasizes that analytic induction underscores the value of a hypothesis is not confirmed because it orients observations towards the development of better hypotheses that fit reality better. Indeed the nature of the divergence provides information regarding how the hypothesis needs to be revised. Thus deviant cases are specifically sought so that the theory in development can be maximally generalized to apply to all relevant examples.

While Znaniecki criticized quantitative methods because they often overlooked inductive analysis, Robinson notes that the best quantitative researchers and theorists use negative results from individual studies to modify hypotheses. More distinctive to analytic induction, on the other hand, is redefining what is studied so that exceptions are excluded. This delimiting of the phenomenon also delimits how widely a hypothesis can be applied. Robinson describes the delimiting of Isaac Newton's theories by the development of relativity theory as an example of such a delimitation. Robinson states that analytic induction as described by Znaniecki only provides the necessary conditions for a phenomenon, but does not give the sufficient conditions for its presence. Consequently this method of data analysis does not allow for prediction of the phenomenon studied; analytic induction provides only a partial explanation for the phenomenon which may be supplemented by enumerating data to provide a fuller explication. Yet analytic induction avoids probabilistic statements of what is most likely; the desired end product is certainty that fits all known cases. Robinson concludes that enumerative methods begin where analytic induction ends, since it provides the necessary but not sufficient conditions for a phenomenon. Consequently analytic induction can be combined with quantitative approaches (to which Znaniecki reacted) for a fuller view of what occurs. In reacting to Robinson's critique, Lindesmith (1952) reacts to the combining of quantitative and analytic inductive methods of analysis, emphasizing that each relates to very different kinds of problems. Lindesmith underscores the distinctive aspect of analytic induction in that disconfirming cases are overtly sought by the researcher so that the developing theory can be more comprehensive. This is in direct contrast to probabilistic sampling which marks most quantitative approaches; it is not the typical case but the atypical exception that is sought. Likewise, the physical sciences are most likely to advance when evidence that disconfirms existing theory is sought; not when scientists seek typical cases (see Hawking, 1988, for a description of how Newton's theory had to be modified due to the results of a 1919 experiment). Lindesmith also notes that statistical methods do not prove causation, only that they may contribute to an analysis of causation.

Howard Becker (1958) describes the analysis used in his now famous studies of medical school students, part of which he links with analytic induction (pp. 653 & 658, footnotes 2 & 13). However, it should be noted that he blended several ways of doing analysis in the study, including enumeration of qualitative data. In Becker's description of analytic induction, data analysis begins while data is being gathered, unlike most quantitative approaches where analysis begins subsequent to data acquisition. Analyses are tentative and provisional throughout the study and only become comprehensive once the data is completely collected. It is crucial in the process of revising tentative hypotheses that specific indicators of a phenomenon be identified, indicators that are more observable than the more abstract phenomenon itself. It may be necessary to search for those indicators, or it may be that the indicators are observed initially and later come to be understood as pointing to a larger phenomenon. Initially, however, data is used only for speculating about what might be, possibilities that are likely to be discarded or significantly modified before the research is completed.

Becker suggests the possibility of enumerating the distribution of events in the categories developed. These data may not allow for precise quantification or meet the requirements for statistical tests, but rather give some indication as to how likely a conclusion is applicable. These are termed "quasi-statistics," representing the degree of confidence in the findings. Multiple forms of evidence are preferred, with greater confidence resulting when multiple varieties of evidence point to the same conclusion. Partial models are built in the process of research, again making use of the search for negative cases. Eventually potential connections between partial models are explored until an more comprehensive model of the social organization of behavior becomes possible. Checking and rechecking of models may result in the need to rebuild one's analysis in the attempt to account for every aspect of the analysis. Becker notes that often some of the best evidence is accumulated when

the observer records what occurs without reflecting carefully on it at the time. Since the observations are not being placed within a system of understandings, they is less likely to be biased by the researcher. This early experience may help in the formulating of early hypotheses which are tested by the search for negative cases, as alternative hypotheses are considered in light of the evidence. Becker concludes that researchers should include a natural history of the various stages in arriving at conclusions, so that readers of the report are able to judge the adequacy of the procedure and the evidence for conclusions.

Glaser and Strauss (1967, pp. 103-105), in their now famous introduction to constant comparative analysis, briefly describes analytic induction and contrasts it with their own method of analysis. They emphasize that analytic induction involves generating theory as well as testing theory in a provisional manner, whereas their own approach emphasizes the generating function (and thus is termed "grounded theory") without testing. The theory produced by analytic induction is universal, precise, and limited. They underscore that all data available must be used to test hypotheses, in contrast with constant comparison which requires that data only be used until categories become saturated. Glaser and Strauss also differentiate their approach by the concentration upon causation in analytic induction, while constant comparison adds properties such as processes, types, consequences, dimensions, and conditions. [This contrast seems to me a truncated understanding of Znaniecki. While Znaniecki does give strong emphasis to causation, he also gives many examples of the other properties Glaser and Strauss list, though they are taken from analytic induction as used in the natural sciences rather than from examples in the social sciences.] More recently Hugh Mehan (1979, pp. 16-22) describes his own data analysis approach, termed "constitutive ethnography," as analogous to Znaniecki's and Robinson's analytic induction. Aspects of analytic induction incorporated into Mehan's model include the constructing of a model that includes and explains all data relevant to the research question, including anomalies. The process of analysis is initiated with analysis of a small data set from which a tentative hypothetical framework is generated. Comparisons are made with additional forthcoming data resulting in changes in the framework until a group of "recursive rules" are developed that comprehensively describe the phenomenon. Mehan repeats a crucial distinction between analytic induction and other approaches: residual cases are explained by the inclusion of additional variables in other kinds of analyses, whereas analytic induction changes the analytic scheme to account for exceptions to the provisional framework. Judith Goetz and Margaret LeCompte (1981) outline analytic induction in six steps, summarizing much of the previous literature on this method of analysis [an additional step described in an adjacent paragraph of their discussion is included here at its respective position]: 1) scanning of data collected in field notes to identify categories and attributes, 2) additional scanning of the data for other examples of categories, 3) creating typologies for categories, 4) determining the relationships that exist between categories, 5) creating hypotheses from the relationships discovered, 6) seeking examples that contradict hypotheses, and 7) continually refining hypotheses until all examples are accounted for and explained.

Goetz and LeCompte note that while analytic induction is a generative approach, it can be combined with enumerative strategies [here they follow Becker's and Mehan's extensions of Znaniecki]. They also note that analytic induction is particularly useful in the initial phases of research and is not limited to a specific variety of analytic unit [thus correcting Glaser and Strauss' overemphasis upon causation]. Bogdan and Biklen (1982, pp. 65-68) emphasize that the initial hypotheses in analytic induction are derived from data, particularly from initial observation and interviews. In subsequent phases of a study researchers make use of purposeful sampling, where specific individuals are selected for inclusion because they are most likely to help broaden the theoretical explanation. The result is the opposite of "funneling"—instead of conclusions becoming restricted, a more encompassing yet refined theory is produced.

Katz (1983) emphasizes that several hypotheses are held at the beginning of a study and "a mass of hostile evidence" is confronted very early in the analysis (p. 132). Inevitably the researcher flounders in the selection of which hypothesis is to be selected, what datum to consider as a negative example, and whether the theory needs to be changed or the definition of what is being observed should be altered. While supportive of analytic induction, Katz describes some of the weaknesses of analytic

induction. One obvious difficulty is that the quest for universal explanation is elusive. He believes that the test of success is not a comprehensive explanation, but rather the degree to which changes in constructs and hypotheses explain the phenomenon with few exceptions. The desire for an exhaustive description should be considered a research strategy, not the criterion for success. Katz criticizes Znaniecki's unnecessary hostility toward quantitative approaches [in contrast with later users who added enumeration to analytic induction]. He also notes that the product of analytic induction is not causal explanation, as claimed by Znaniecki and some of his followers, but at best good definitions. Because everyday life is a continual process of ongoing symbolization, concepts used tend to have boundaries that are marked by vagaries. "If analytic induction follows the contours of experience, it will have ambiguous conceptual fringes" (p. 133). This ambiguity is accounted for through the probabilistic conclusions of quantitative researchers and by the difficulties of coding examples on the boarders of categories in analytic induction; both approaches reflect uncertainty and the possibility of error. A key value of the approach is that findings are more likely to be generalized since numerous examples must be explained through successively qualified versions of the hypotheses.

References

Becker, H. S. (1958). Problems of inference and proof in participant observation. American Sociological Review, 23, 652-660.

Bogdan, R. C., & Biklen, S. K. (1982). Qualitative research for education. Boston: Allyn & Bacon.

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory. New York: Aldine.

Goetz, J., and LeCompte, M. (1981). Ethnographic research and the problem of data reduction. Anthropology and Education Quarterly, 12, 51-70.

Katz, J. (1983). A theory of qualitative methodology. In R. M. Emerson (Ed.), Contemporary field research: A collection of readings (pp. 127-148). Prospect Heights, Ill.: Waveland.

Lindesmith, A. R. (1952). Two comments on W. S. Robinson's "The logical structure of analytic induction." American Sociological Review, 17, 492-493.

Lofland, J. (1971). Analyzing social settings. Belmont, Calif.: Wadsworth.

Lofland, J., & Lofland, L. (1984). Analyzing social settings, 2nd ed. Belmont, Calif: Wadsworth.

Mehan, H. (1979). Learning lessons. Cambridge, Mass.: Harvard University Press.

Miles, M., & Huberman, A. M. (1980). Qualitative data analysis. Newbury Park, Cal.: Sage Publications.

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

Robinson, W. S. (1951). The logical structure of analytic induction. American Sociological Review, 16, 812818.

Spradley, J. P. (1980). Participant observation. Orlando, Fla.: Harcourt Brace Jovanovich.

Znaniecki, F. (1934). The method of sociology. New York: Farrar & Rinehart.