DOCUMENT RESUME

ED 318 759 TM 014 796

AUTHOR Carifio, James; And Others
TITLE Towards a Conceptualization and Theory of Sampling Large, Complex, and Dynamic Populations.
PUB DATE Apr 90
NOTE 44p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Data Collection; Evaluation Methods; *Research Problems; Response Rates (Questionnaires); Sample Size; *Sampling; Simulation; Statistical Analysis; *Statistical Bias; *Testing Problems

ABSTRACT

Possible bias due to sampling problems or low response rates has been a troubling "nuisance" variable in empirical research since seminal and classical studies were done on these problems at the beginning of this century. Recent research suggests that: (1) earlier views of the alleged bias problem were misleading; (2) under a variety of fairly well-specified conditions, allegedly biased samples are in fact random; and (3) "de facto" biased samples can and will "jump" from being biased to being random samples. These conditions occur when the population is fuzzy and dynamic and when the number of factors structuring the population and those influencing responding or non-responding are large, interact with each other, and have a trivial effect, at best. Researchers and theorists may be rejecting and disregarding what might be reasonable and highly valid data. Actual data from a number of state-wide studies with response rates in the 15% to 30% range illustrate this view. A variety of simulation studies is needed to evaluate these views on bias. The Catastrophe Theory of non-linear functions, family-wise error rates, and fuzzy subset theory are also considered. (Author/SLD)

TOWARDS A CONCEPTUALIZATION AND THEORY OF SAMPLING LARGE, COMPLEX, AND DYNAMIC POPULATIONS

James Carifio, University of Lowell
Allen Shwedel, Boston Public Schools
William Kermis, Northwest Missouri State University

TOWARDS A CONCEPTUALIZATION AND THEORY OF SAMPLING LARGE, COMPLEX, AND DYNAMIC POPULATIONS

James Carifio, University of Lowell
Allen Shwedel, Boston Public Schools
William Kermis, Northwest Missouri State University

Abstract

Possible bias due to sampling problems or low response rates has been an extremely troublesome "nuisance" variable in all areas of empirical research ever since seminal and classical research studies were done on these problems at the beginning of this century.

A wide variety of work, data, and new theories in many different areas, however, strongly suggest that these classical formulations and views of the alleged bias problem are essentially incorrect, misleading, and overly simplistic, and that under a variety of fairly well-specified conditions allegedly biased samples are in fact random (most probably), and that _de facto_ biased samples can and will "jump" from being biased to being random samples, in the manner outlined and described by Catastrophe theory (Zeeman, 1976). These conditions are when the population is fuzzy and dynamic, and the number of factors "structuring" the population (or sample) and those influencing responding or non-responding are large, interact with each other, and net out to zero, or the statistical equivalent of zero (namely, a trivial effect at best). Researchers and theorists in many areas, therefore, may be rejecting and disregarding what might be reasonably and highly (if not perfectly) valid data and studies. Actual data from a number of recent state-wide studies of large, complex, and dynamic populations which had response rates in the 15% to 30% range empirically support this view.

A wide variety of simulation studies of these newer views, which are outlined in this paper, are needed. Some of these needed simulation studies, as well as the nature of the problem(s), are also outlined in this paper.

The very strong suggestions of these newer views, if they prove to be essentially correct upon empirical simulation, would be a very important discovery, not only for the areas of research methodology and statistics, but also for the areas of cognitive psychology and machine intelligence.

Paper presented at the Annual Conference of the American Educational Research Association (Division D, Research Methodology), Boston Ma., April 16-20, 1990.

Increasingly in the social sciences, researchers have been attempting to conduct important surveys of large and complex real-world populations on critically important issues and questions. One problem that has continually plagued these research efforts, as well as the evaluation and acceptance of findings of these surveys has been the low response rates obtained from subjects nominally considered to be part of the population being surveyed. Typically, response rates to surveys of large, complex, real-world populations tend to be 25% to 40% of the sample drawn for participation. Typically, the sample drawn tends to be 10% to 20% of the nominal population, so that the resulting sample of responders tend to constitute data on 5% to 10% of the nominal population.

Usually, these low response rates are most interpreted as indicating that there is a de facto bias in the analysis sample. Consequently, the results of the analyses done are usually strongly challenged and deemed unacceptable because of this alleged bias and both the data and results of analyses are unquestionably considered to be highly dubious and suspect at best. The key and fundamental problem, however, is that the accusation and claim of bias is only a de jure, inferred, extrapolated, and "armchair" claim of bias, and is not a de facto proof of bias (even probabilistically), and that the very theory upon which this inferential, armchair conclusion is based is rift with deep core definitional and conceptual problems, contradictions, impossibilities, and several other major flaws which makes this theory extremely dubious at best in terms of providing an adequate or actual model of events and what actually occurs in the real world (rather than in urns with balls that have a single property) and real world situations.

This fundamental fact (and problem) has led many theorists and researchers in all areas of science and social science to strongly question and challenge the general classical theoretical perspective, views and models developed primarily in the first half of this century. These views have been crumbling and stumbling badly in all areas of science and social science in the last half of this century, particularly as more and more data emerges that contradict this general view, and more conceptual and theoretical analyses are done of this view. These analyses have tended to produce newer and very different theories rather than revisions or modification of the general classical theory, which is essentially a strict set theoretic, logical positive, linear, one-dimensional view and model of the world and its events and operations.

Practicing researchers (in particular) and theorists in all areas of science and social science are currently extremely uncomfortable, suspicious of, and ill at ease with the classical models, positions, and views as they are daily turning out to be extreme over-simplifications that break down quickly and do not work (by a long shot) even as a "very broad gauge macro model" in the real world of everyday life and research. The real everyday world is qualitatively very different (see Cronbach, 1988) from the "world" of the the "simple experiment," or the simple "hermetic lab, T-maze, or urn." Compared to these over simplified "worlds," the everyday world of real life and real experiments is "complex" and even "highly complex," and this "complexity" is very different from classical (over) "simplicity," and not an extension of this "simplicity," but rather a new area requiring a new science; namely, a "science of complexity" (see Pagels, 1988). The real everyday world is not an extension of the "simple classical model," but rather, it is a quantum leap from their "simplicity," and a "jump" of the kind characteristized by Catastrophe theory (Zeeman, 1976), which is a general theory of non-linear phenomena and events.

This great and rapidly emerging "discontinuity" between the "simple" and the "complex" has been strongly impacting practicing researchers and theorists in all areas of science and social science in the past five years, but the impact in the area of sampling, statistical and research theory has been minimal at best. This emerging discontinuity, however, has begun to make some impact on these three areas very recently, as several leading theorists have begun to express their extreme discomfort with the classical views on key topics in these three areas, and their concerns with and alienation from classical theories and models in these three areas that are blythely and unquestioningly accepted by the majority of professionals working in these three areas.

Nowhere are these later points more evident than in the area of the topic of this paper.

Angoff (1988) has called the problem of bias the most ill and poorly defined, and poorly understood concept in all of measurement, statistics, and research, and the one concept most in need of clarification, analysis, conceptualization, hypotheses formation, and empirical simulation of the kind finally done by Glass (1972) on the robustness of ANOVA to violation of its many assumptions, the results of which suprised everyone, including Glass. And the surprise was a very pleasant one, particularly for practicing researchers.

Similarly, Wainer (1989) concluded that the bias question and area is currently shrouded in a very thick fog, and what is needed is not more statistical equations, formulae, and so forth, but better conceptualizations, analyses, modeling and thinking about the problem that will indicate some directions that eventually will put a dent in one of the most important and perplexing methodological problems in research today: missing observations.

This paper, therefore, is an effort in the direction and along the lines that Wainer, Angoff, and others have recently suggested, although our views of the questions and problems were developed independently and in parallel to the work of these theorists and in many respects are very different from the views and analyses of these theorists whose work, thinking, ideas, conjectures, (and courage) we respect and admire greatly. Consequently, in this paper, we are going to try to outline our views, analyses, and thinking about this most fundamental of problems as well as the prototype models and ideas that we have, and have explored to date, as well as some of the data that support our views and the kind of activities and studies that need to be done to advance understanding and models in this area and support or deny several of our conjectures which are derived from newer theories and our experiences. Our primary goals in this paper are (1) to share our thinking and work in this area with others, (2) to get others interested in the ideas, work and issues we present, and (3) to encourage and stimulate others to pursue these ideas, models, issues, and conjectures independently as no single effort or work is going to be definitive or conclusive. There are simply too many questions, and too many pieces to the puzzle at this point, and not enough data, and particularly simulation data and studies.

Initially, we must begin by saying that we are quite confident in the essential correctness of the views, theoretical analyses, points, criticisms, models, and conjectures outlined in this paper within the limits, conditions, and boundaries of the particular situations and phenomena we are focusing upon in this paper; namely, large, complex, and dynamic "populations" that are sampled in an attempt to make (probabilistically) valid statements about the population based on the fragment in hand (i.e., the actual sample data). However, based upon the sound advice of a highly valued anonymous reviewer of a summary of this paper, we wish to emphatically state at this point that our work, models, data, and thinking cannot be currently construed at this point in their development as definitive proof by anyone that their "low response rate" sample is in fact a random sample of the population in question, and there is in fact nothing "wrong" with their sample and that they do not have to be careful and cautious in analyzing it or interpreting their results. As will be seen in this paper, our position, emphatically, is that the resulting randomness or non-randomness of any particular sample in fact is a function of several reasonably well-defined factors and events that must be present and operational in the situation, and that the conclusion that the actual sample is essentially a random or non-random one can only be done probabilistically at this point in time, until the appropriate simulation studies are done. Our views may differ, and even differ radically from those of others; however, this difference in no way implies, or should be taken to imply, that we do not strongly adhere to the purposefully conservative position of science as a process with respect to emerging science.

Conundrums

The views presented in this paper began from two extremely simple (but highly anomalous) observations and/or experiences. The first of these two observations or puzzling experiences was trying to explain how human beings could draw valid inferences, make valid deductions, and draw valid conclusions consistently, and in a highly reliable fashion, from the seemingly highly biased, unrepresentative and very small samples (fragments) of reality that they experience daily and unquestionably in the real world. This very simple, basic, and elementary fact replicated millions of times an hour (at a minimum), every hour of every day world wide, cannot be explained in any simple, easy, or straight forward way by classical sampling and statistical theory, and places classical sampling and statistical theory and all of the concepts and conceptual framework associated with it in the same place and in the same logical and empirical position as classical neurobiology was in the fifties when Lashley said that he would logically be forced to conclude that learning did not and could not exist from the framework, viewpoints and the "received" theories of contemporary neurobiology.

This point is neither a specious or trivial point or argument, nor the stretching of a point or example. It is the converse of the above simple point that is always and the only one emphasized by classical (statistical) theory. The "pro" side of the simple point stated above, however, occurs with at least equal frequency and is equally correct and valid, and therefore must be explained by classical or "received " theory as well. One cannot model and explain just "one side" of a coin and say that this "one-sided" coin "exists" and is real and that the other side of the coin does not exist and is not real. This approach just does not pass muster intellectually or scientifically. If one is going to play the "theory game," then one must play it fair and square, and not make "all" claims for "part" models, because the very part one has thrown out just might be the most important part.

The above point and simple conumdrum is a prime example of the selective attention problem, which is one of the most key and fundamental problems currently impeding progress on and resolution of the many issues and disputes associated with the subject matter of this paper. Selective attention is often quietly, subtlely, and fatally biased because it is unweighted and is not paying attention to and open to both sides of the coin simultaneously, and dynamically adjusting perspective (and theory) on the weight of argument and evidence (logical, theoretical and empirical). When the latter is done, the above coin can spontaneously and instanteously "jump" from one side to another right before one's eyes just like one's perception of ambiguous visual (or statistical) figures. h "jumps" are considered to be "catastrophes" by "either/or" views, but are not catastrophes at all by more inclusive "both/and" views.

The coin described above is a "paradox, contradiction, or conumdrum" coin. It is not the coin one reads about in classical statistical theory. The properties, operations, and dynamics of the conundrum coin, moreover, are very different from the classical coin, because there are opposing (contradictory) forces and interactions built into the coin that are operating simultaneously and are going to "net out" at different time points and conditions to one side of the coin or the other being "up." It is the conundrum coin that can explain how an allegedly biased sample due to non-responders can in fact be random, or so close to random as to constitute totally trivial bias at best. It is also the conumdrum coin that can explain how a sample could and can "jump" from being biased to being a random sample under various sets of reasonably well specified conditions. The formal theory that best represents the conumdrum or paradox coin and the exact dynamics and mathematics of its operations is Catastrophe theory (Zeeman, 1976), which will be outlined in more detail below.

The conundrum coin and the fact that human beings draw correct inferences and deductions from allegedly biased samples of everyday reality are very key and fundamental points and part of the very essence of the problem. These two points go to the very roots of the problem, both conceptually and historically. If a theory, conceptual framework, view or position cannot explain very simple, basic, elementary and incontrovertible facts and events that can be observed daily and anywhere by anyone, and are observed daily and everywhere by everyone (and thus large-scale, highly controlled experiments are not even really needed), and the use of this theory, conceptual framework, view, or position forces one to logically conclude that these incontrovertible facts and events do not exist, and cannot exist, and everyday reality is to be denied, then they are undeniable and uniquivocably fatal flaws, paradoxes, and contradictions in the theory, conceptual framework, view, or position and they are not minor, or minor blemishes, or small points or quirks that may be remedied in various ways.

The occurrence and existence of the type of point, condition, and situation outlined above is almost always indicative of very deep seated and basically fatal flaws, errors, and contradictions deep within the inner vital core kernel of the theory, conceptual framework, view, or position. There are many, many examples of this point historically. Behaviorist theories of language being completely unable to explain the production or comprehension of a new sentence by an individual, or the comprehension of or laughter at a joke, is one pertinent example. Classical learning theories not carrying over to the natural learning context of the everyday world and in fact "declining and falling (see McKeachie, 1979)" is another.

All of these points are made to quickly, succicntly, and simply convey the view that the positions to be presented in this paper have a very strong a priori credibility (warrants) basis in simple every day observations, facts, and logic that cannot be denied or dismissed or patched over and concealed, and to emphasize that there is another extremely important reason to pursue the concepts, ideas, new theories, and thinking directions outlined in this paper beyond sampling theory, research, and statistics. Since the ideas, concepts and new theories outlined in this paper are capable of explaining the very simple observation described above, working some of these problems, issues, and ideas out could make an incredibly valuable contribution to cognitive psychology and machine intelligence. So there is an additional and highly valuable incentive and prize and this prize is basically why we are so interested in this set of interrelated problems. If any of the ideas, points, suggestions, and conjectures outlined in this paper prove to be essentially correct upon further empirical investigation, such findings would be important to many areas other than statistics and research methodology, and, this is why we believe, like Wainer (1989), that this "nuisance variable" is extremely important and worthy of intense study. This point is also the reason why we believe that work on this subject and this "nuisance variable" problem needs to be highly interdisciplinary. Concepts and theories from several different disciplines are needed to start making a "real conceptual dent in this problem , which according to Wainer (1989) is our first real order of business. Consequently, we would like to begin with what we believe is one of the essential ingredients to providing some answers to the many conumdrums in this area; namely, Catstrophe theory.

Catastrophe Theory

Catastrophe theory (Zeeman, 1976) is essentially a theory and mathematics of non-linear functions and relationships that is far more specific and precise than Chaos theory (Gleik, 1987), which is a general and all emcompassing theory of non-linear phenomena. Non-linearity is a mathematical and theoretical "catastrophe" for classical linear models and conceptualizations of phenomena. A theoretical or observational "catastrophe" occurs when the linear model predicts the "next data point" or results from previous data points and the predicted result in reality is observed to be in exactly the logically and mathematically opposite place.

Seven very distinct and different general "catastrophe" models have been developed (Zeeman, 1976) which range from simple to complex depending on the number of observable factors and behavioral dimensions that are related together in the model. Each of these general models describes how the "regression surface" or "plane" in the linear view is "folded" into contoured shapes or structures that can be described by mathematical equations which make the data orderly, structured, and predictable, both theoretically and logically, as well as mathematically.

The key feature of and fundamental principal embedded within each of these seven specific Catastrophe theory models is that event points jump from one "logical" state to the logical inverse of the preceding "logical" state due to the interactions of and buildup (net) effects of reciprocal and contradictory factors operating *simultaneously* upon the point in question. The more interacting factors affecting the point, the more complicated the full event space model and its folding is, but as the number of contradictory factors operating increase, they tend to net out to zero net effects, except for very specific and well-defined regions within the event space of the model where "jumps" become "confined." This later point is a very important point for a number of reasons, but most particularly in terms of the selective attention problem (focusing on one factor and the effects of one factor only), and the net effects of many contradictory or opposing factors interacting and operating simultaneously upon an event or event point netting out to zero effects (namely, no resulting bias).

Catastrophe theory is a very specific and very precise theory and mathematics of interaction and non-linear phenomena, which is to say *conditional* phenomena, rather than general, simple and uniform phenomena. Catastrophe theory is not only the other side of the conumdrum coin, but it can in fact explain the condundrum coin and its operations even to predicting when the conundrum coin will land on its side. It should be noted, however, that Catastrophe theory's models of interactions are very specific, very precise, and well bounded models of interactions that are unlike those of ANOVA which are general and almost completely unbounded and do not have theory contradicting states or outcomes like Catastrophe theory does. Catastrophe theory, moreover, is also far more precise and specific than chaos theory, or fractal theory (Mandrebot, 1975), which also suggests that samples from large, complex, and dynamic "populations" that are influenced by many contradictory factors operating simultaneously will be extremely good "exemplars" of that "population." These implications of fractal theory, which address the feature of and the implications and effects of the structure, structural complexity, and the structural effects of a "population" on consequent and result samples (fragments of the population) are incredible ideas that are at first so counter-intuitive as to warrant detailed consideration and investigation in and of themselves.

Catastrophe theory is not so mysterious at its simplest and most fundamental level. The deep implications of Catstrophe theory are that events (samples) can jump from being biased (or ordered) to being random, or vice-versa, depending upon the nature of the factors, conditions, and interactions simultaneously operating in a given event space. This fundamental point and observation is essentially the fundamental point and observation of Chaos theory and Fractal theory. Both of these later theories, however, point out that what may be ordered at one level may be unordered (random) at another and vice-versa. It is our view, however, that only Castrophe theory at this point in time has the precision neccesary for simulation studies and that is why we have a strong preference for Catstrophe theory as opposed to other theories and models.

What makes catastrophe theory so unique and powerful is that it can handle logically contradictory criteria or dimensions with equal ease and facility as logically consistent or uniform criteria, and it is this feature of catastrophe theory that makes it so interesting, and interesting in terms of the uniquely different surfaces and models that it generates depending upon the number and logical type of the input criteria. It is also this feature of catastrophe theory that makes it so powerful and capable of modeling with great precision such contradictory, discontinuous, and non-linear phenomena as aggression, anorexia nervosa, the stock market, cathartic release from self-pity, the buckling of an elastic beam, phase transitions, and a wide variety of other such non-linear phenomena. Catastrophe theory alone, however, cannot explain our data. Fuzzy subset theory (Kaufman, 1979) is also needed as well as several other concepts. Fuzzy subset theory is needed not only to describe the difference between the process of sampling real world (fuzzy) populations in the real (fuzzy) world, as oppose to the classical ball and urn model of these events, but also to clarify the distinction that must be made between the _nominal_ population and the _effective_ population (true members who are truly available to be sampled and/or respond), particularly relative to large, complex, dynamic, and real world populations. However, it is best at this point to discuss our data and data from some other researchers, as Catastrophe theory's main ideas of the net effects of interacting (contradictory) factors and "logical state jumps" are key ideas and concepts relative to the initial discussion and explanation of findings.

<u>The Second Observation</u>

Our second extremely simple observation and/or puzzling experience occurred when we conducted 3 consecutive annual, state-wide, large-scale surveys of graduates from 15 community colleges and eventually their employers. Through computerized records given to us on computer tape by each of the 15 community colleges, we constructed close to a 40 variable database (N=5,781) on various characteristics of the entire (hypothetical) graduate population (see Carifio and Shwedel, 1983). When we found that the background characteristics of the 20% to 30% of the graduates who responded to the surveys were multivariately no different than those who did not respond, we were extremely surprised and tried to develop some initial conjectures as to why. But when the same results happened a second time (population N=5,627) and a follow-up study of a 10% sample of non-responders revealed no differences on any of the 15 dependent variables examined, we did more than conjecture, we started to build a prototype conceptual model that could explain why this could happen (see Carifio et al., 1988).

When the same results occured a third time (population

N=5,129) in a follow up of employers of graduates, we became more confident in our thinking and model (see Carifio et al., 1988), and then when we encountered two state-wide surveys of practicing nurses that began with a population database of subjects characteristics of N greater than 5,000 and found that 10% samples that had only 40% to 50% response rates in terms of the samples survey (making them 5% samples of the population) were random (see Fazekas, 1989, and Hunt, 1989), we became even more confident in our thinking, prototype model and views, and that our results were not just highly improbable flukes. We also began to have a great deal of confidence in older reports of similiar experiences shared with us by other researchers engaged in similar activities in others states (e.g., Boakes, 1981, and Paulson, 1982).

The 5 studies described above all have a large number of features in common that are extremely important to the subject of this paper and in explaining the outcomes observed in these 5 studies. First, all 5 studies were dealing with relatively large and structurally complex "populations" and study spaces as compared to the typical study, but most particularly as compared to the classic "ball and urn simple experiment model" and/or classical "extensions of this model to more complex studies and spaces (see Coombs, 1964 for the most detailed of descriptions of these two points). Next, both the elements (or members) of these complex populations, and a number of complex intervening factors, were distributed across the event spaces of these studies, making the event space and the population it contained in each of these studies very fuzzy, lumpy, and distributed in effects rather than smooth, homogeneous, exact, precise, and certain like the classical event and study space model as is best outlined and described by Coombs (1964).

In all of the 5 studies described above, the nominal population was known in all of its multivariate complexity. A complex and complete database existed on both the multivariate features of the nominal population and the various possible intervening factors and "breakdowns and influencing factors" within it. Sample results could be checked back against the population multivariately in terms of assessing the degree of bias or randomness of the resulting responder sample. This key feature protects and protected these five studies from the selective attention (to armchair identified) factors problem which is the key feature and major flaw in classical theory and classical "views" and studies of bias, sampling and the non-responder problem. This key weakness is the logical equivalent of the family-wise error rate problem in statistical testing and the need for control of this problem in doing statistical tests. This key weakness is one of the major reasons why all of the classical research, conceptualizations and thinking in this area is so severely flawed, highly questionable, and completely dubious. We will elaborate this point more fully at the end of this paper.

All 5 of the studies outlined above had a real world fuzzy and distributed (random) variable embedded in the core of the conduct of each study that the urn of the classical model and hermetic laboratory context does not allow to operate or exist. Each study had to construct a definition of a member and then sieved elements to find the population (if one thinks that it is extremely easy to define a graduate of a community college, then one needs to spend some time talking to registrars at a few colleges). Next, each study had highly dynamic "true members" of the population fitting the constructed membership definitions.

Current street addresses (the major member locator variable in the large, distributed, fuzzy urn) were inaccurate for up to 20% of the cases in some instances with no discernible pattern (bias) in terms of inaccurate cases. This inaccuracy rate in current addresses is not particularly high, according to a wide variety of people to whom we have spoken who work this type of data and problem. Educated professionals are a highly mobile and dynamic population, and only an instance of the many kinds of dynamic populations there are in all areas.

We found that in following up on non-responders that a goodly number had never received the survey we sent them and/or we had not received a goodly number of the surveys they had returned. Again, this was another dynamic variable that had no discernible pattern (bias) to it in terms of the background variables in our database. Again, in talking to many other people in a variety of different areas, we found that this state of affairs is not at all uncommon, and has been in fact occurring more frequently in the last few years.

The effects of all of the above variables in all 5 of these studies were (1) to reduce the size of the nominal response rates to between 5% and 25%, when in fact the actual effective response rates, when all of these factors were taken into account, were between 35% and 55% of the effective population (see Carifio et al, 1987 for details). The responder samples, therefore, were actually very much larger samples than one would believe they were, if one evaluated response rates in terms of the nominal popluation "urn." The net effects of all of the random and contradictory response affecting factors in these 5 studies were statistically zero.

Each of the responder samples in each of the 5 studies were _de facto_ statistically random relative to their populations, although a wide variety of armchair arguments could be made about various (selective) factors that would strongly suggest, if not prove, that the responder samples were most probably biased samples, when the exact opposite was true statistically and _de facto_. Catastrophe theory, in particular, but fuzzy subset theory and Chaos theory also, would have predicted, however, that the responder samples in these 5 studies were most probably random and not biased. These theories, moreover, led to the most simple initial formulation of our views.

This formulation says that as the structural complexity of a population increases in terms of the number of variables characterizing the elements of the population simultaneously, and as <u>the number of interacting factors</u> increase that simultaneously influence the occurence or non-occurence of responding (and as these factors increase in number as contradictory factors), then the greater the probability that the resulting responder sample will be a random sample of the population.

Further, this view says that the probability that the resulting responder sample will be a random sample of the population is far, far greater (and closer to 1) than the probability that the responder sample will not be random; namely, than the probability that it will be a <u>significantly</u> (as opposed to merely statistically) biased sample. This very simple and somewhat obvious question of just <u>how</u> biased any given sample is, and exactly how biased is biased, both quantitatively and qualitatively, is one of the key, critical, and outstanding unanswered questions in this whole area, which we will comment on further below.

We initially thought that this view and our predictions could be construed as a restatement of the "Law of Large Numbers," where the word "Number" is replaced by the words "Number of (interacting) factors involved and number of (interacting) events involved in a given study space." It could be said that our points and views are inherent, but masked, in the classical view as unseen, but inherently present, implications and deductions from the Law of Large Numbers. This particular point does have some validity to it, but it would mean that we would now have to start considering and talking about the "Laws of Large Number," rather than just the Law of Large Numbers. We have no real problem with such an approach, other than the fact that it has taken so long to perceive, and that it could keep us all shackled to the empirical-descriptive rather than making the transition to the conceptual, theoretical, and explanatory.

Stated in the most simple of terms, our view says that the number and nature of factors simultaneously operating in an event (NNFSO) is as important as N (sample size) in determining the resulting state of the event (i.e., the nature of the resulting sample). This point, it should be noted, is also the implication of Chaos theory, Fractal theory, and fuzzy subset theory, but it is only Catastrophe theory that gives the simplest, clearest, and most precise explanation as to why this is and will be so.

The simple Cusp catastrophe theory model, which is essentially a bimodal state probability distribution, predicts that the sample of responders under the types of conditions described above will gradually begin to shift and then eventually "instantaneously jump" from being a biased sample to being a random sample of responders as the sample size increases, and the number of factors and contradictory factors affecting responding increase in number, and then stay random, due to the effects of all of these interacting factors netting out to zero, particularly for large sample sizes. This simple model and the various factors identified above both explain and account for the observed resulting random samples in the 5 studies described. Catastrophe theory's predictions are very precise, specific, and simulatable. Catastrophe theory says that depending upon the nature of the factors operating in a given study space (NNFSO and N), one of two "strange attractors" will begin to operate and prevail eventually in the study space. The first of these two "strange attractor" is bias (which has been somewhat cursorily studied by the classical view), and the second strange attractor is randomness, which will occur and prevail in large, complex, and dynamic populations, whose samples are of some reasonably sufficient percentage size. It is this second "strange attractor" that is and has been the focus of this paper.

It should be noted that "strange attractors" are most often associated with "scale discontinuities," and structural discontinuities between the "simple" and the "complex." The populations and samples in the 5 studies described above were not large and complex in terms of our meaning of the terms large and complex, but rather only sufficiently large and complex to cross the threshold into the domain of the large and complex that is the focus of this paper. The populations in the studies reported above were only large and complex as compared to typical studies where the population and sample N's are usually 1/5 to 1/10 the size. However, it is a fairly well documented fact that when an effect begins to manifest itself in a statistical model, it starts abruptly and then moves quickly, and we only have to plot values on look-up tables in statistical texts to demonstrate this continually recurring fact. The interesting observation is that only Catastrophe theory is really capable of giving a very full and detailed explanation of this abrupt "kick-in effect," which typically tends to be depicted as simply an empirical-descriptive fact. Empirical descriptions are not explanations and empirical descriptions are not theories.

Strange attractors are a key concept of Chaos theory, but they are also inherent and most definitely present in all seven Catstrophe theory models, each of which describes and models, both conceptually and mathematically, the general points, conditions, and areas where "strange attractors" will develop and occur. It is the concrete and specific character of Catastrophe theory that has drawn us to it as a model and as a thinking model for all of these puzzling questions. Also, Castrophe theory has a specific and clear "semantic" and explanatory component to it which these other theories do not.

Other Features

There are several other features of the above studies that need to be noted relative to the remainder of the discussion in this paper. First, the operational definition of what constitutes a "ball" in an "urn" in a model of a simple experiment is rarely a complex, let alone a fuzzy, definition. "Balls" are not structurally complex entities in the classical model, and neither are the typical populations of "balls" studied in "sampling simulations." Next, in conducting a "ball and urn simulation," one is not making a large number of decisions, nor does one have a <u>distributed decision-making model</u> operating where many different individuals (or Selfridge's demons) are making decisions in parallel producing a net result, outcome, or effect (see Zadeh, 1974b and Caudill, 1990).

This distributed decision-making model (and its effects) is operating at every step in the process in the event spaces of the 5 studies described above. The more elements (and/or demons) participating in a distributed decision-making model, and the more often that this model is operating again and again in a sequential process, the greater the probability that the errors produced by this model will be random, because the probability that a particular bias consistently and consecutively operate effectively and homogeneously this many times in a row is very, very, very close to zero. This later point is also the "tails side" of the conundrum coin.

This distributed decision-making model, and its net effects, are operating at every step in the "real time conducting of the study processes" in the 5 studies described above to produce net effects which interact with every prior and subsequent net effect. As the number of factors in this "real time" model increase, and as the number of factors become contradictory, the net effects across the model will be randomness, particularly at the higher levels of aggregation and data-breaks. One of the major factors that affects the "contradiction index or quotient" of the factors operating in the real-time model is the (relevant) structural complexity of the traits of the elements that constitute and comprise the population.

As a population, and the sampling, research process, and design become more heterogeneous and "complex" in terms of the number of traits and the number of factors involved and examined, and their interrelations and interactions, the greater the probability that the resulting responder sample will move towards randomness, due to the net effects of this complexity. Therefore, the resulting sample will be less biased, at a minumum, than we would think, and particularly think from the viewpoint of classical theory. Consequently, our data would be much better than our (classical) beliefs about our data and we would be highly prone to making "beta errors" and also in actually making them in the analysis and interpretation of the data.

All of the above points say that the data from surveys of large, complex, and dynamic populations are most probably far, far better than we believe this data to be. Otherwise, how would we truly (on the average) get along and survive in the everyday world. All of these same points would also hold for experiments that fit this large, complex, and dynamic model, and all of the points that we have made also have a variety of implications relative to meta-analysis. There are, however, also several other reasons to believe and pursue all of the points that have been outlined above.

Family-Wise Error Rates

We know that in a complex set of data or a complex study space, we can find significant differences (by chance) if we selectively focus on a single and only single feature of the complexity. We know that this finding is a false or "phantom" finding because we have not analyzed the complexity multivariately and controlled the family-wise error rate in our analyses. The family-wise error rate problem is the selective attention problem in a nutshell, and the perfect "quick exemplar" of our points and our points about bias and the biased sample problem. The family wise error rate problem is an exemplar of the conundrum coin and how it "spontaneously flips," and why both sides of the coin must be attended to simultaneously by a balanced algorthm and sequential decision making process (see Zadeh, 1974a). One's attentional and analytical model must be full and open initially and then sequentially move through a "controlled" process examining for and against data and arguments to determine which side of the coin should be up or if the coin should be standing on its edge. It is the clarification and simulation of this process, moroever, that is the greatest of outstanding research needs.

We do not even have rough gauge empirical rules of thumb when it comes to the bias question in any shape or form, and this is why simulation studies are so strongly needed in this area. And the type of simulation studies that need to be done are not the scrupulously clean and overly simplistic in the extreme "ball and urn" classical model studies with one and only one (extremely gross) biasing factor operating and affecting seults.

We really do not qualitatively and precisely know how much any given "biasing" factor or set of factors biases a given sample whose population has a particular well-defined set of characteristics. There is no metric, not even a broad gauge relational metric of the orders of magnitude kind to assess any study or any discourse or claim on the bias question in any form. The significant question and significant research question is not biased or unbiased in terms of results, but how biased; in the second decimal place, the first digit, a tenth of a standard deviation, a whole standard deviation, how biased exactly? Answering this question to some ordinal degree is the only real way that a given sample or study can be rejected or accepted using some type of practical and consistent rule of thumb.

We are aware of no sampling simulation studies that really answer the question how biased, and/or how biased as a function of what factors and/or conditions. Claims of bias and biased results spring up like Topsy all over the research landscape daily again and again, all typically the results of long inferential and speculative chains of arguments and small shreds of suggestive data. When the most major of these claims are actually tested and checked out in Monte Carlo studies, they most often are found to be unfounded or grossly exaggerated (e.g., Glass et al., 1972; Dagenais and Marasculio, 1973; Richards, 1972). The weight of evidence of all of these studies should tell us all something; namely, that assertions of alleged bias should be required to be proved before they are even entertained, because the results of the majority of [redacted] experimental bias investigations and simulations tend to be negative and the bias is proved to be non-existent, trivial, or to hold in only highly restricted (and usually extreme) situations. The weight of the somewhat extensive and high quality existing research evidence has to enter into the consideration of the bias question somehow, at some level, at some point, even only as a rule of thumb to evaulate claims and assertions as we work on the problem. The really outstanding unanswered question is why it does not.

No one, we believe, can currently state how biased any given sample will be under any set of real world conditions, and state it in a way that will roughly prove out to be correct and support classical models and views of all of these questions on empirical simulation. All such estimates most professionals currently make, we believe, are subjective (Bayesian) armchair guesses, and subjective (Bayesian) order of magnitude estimates in the long standing behaviorist post-hoc and armchair tradition that is the classical core of the (overlapping) statistics and research methodology disciplines, as viewed and construed by the majority of current practitioners. Consequently, we have assumed in all of our discussions above, and will assume in all of our discussion below, that we have a simple "Turing Device" that generates an estimate for us of how biased a given resulting sample will be under a particular set of conditions. Our simple Turing Device, however, is a fuzzy and not exact Turing Device.

Fuzzy Subset Theory

Classical sampling and survey theory, in the main, are based on classic set theory and linear model views of the world, which in turn drives all of their logic, inferences, deductions and conclusions (e.g., see Scheffe, 1953; Coombs, 1964; Hayes, 1981; Cochran, 1983; and Kerlinger, 1986). Kaufman's (1976) great insight was that neither classical set theory nor its logic could be applied to real world problems in any easy, precise, accurate, or truly meaningful way. Kaufman, therefore, developed fuzzy subset theory to deal with real-life, set theory, problems, situations, and human thinking, statistically. Fuzzy subset theory is a complete revision of classical set theory to incorporate the concepts of uncertainty and probability into set theory. Kaufman, in devising fuzzy subset theory, actually revised classical sampling theory, although the implications of these revisions were not pursued in any significant way by Kaufman or any other statistical theorist, to the best of our knowledge. There is, however, a reasonably well worked out mathematics of defining fuzzy subsets, manipulating fuzzy subsets and manipulating them statistically.

In fuzzy subset theory, there are, unlike classical set theory, only subsets (due to the "fuzz"), and every set has two basic rules. The first rule is the "classical" rule (or procedure) that defines membership or non-membership in a given subset, and the second is the rule (or procedure) that defines the probability of each member in a given subset as being a "true" member of the subset and its probability of being a member of an alternative subset (or set of subsets). The probability of true membership in the logical subsets is the "fuzz" and it is the "fuzz" that is of prime importance. Unlike the well-defined and hermetic urns, balls, and processes of classical set and sampling theory, membership of any element in any subset is not a given and absolute, but rather fuzzy and probabilistic for some percentage of the elements, and this is Kaufman's real world, key insight and point.

The above features of fuzzy subset theory forces a distinction to be made between the _nominal_ set, subset and/or _population_ (which is the classical view), and the _effective_ set, subset, and/or _population_, which are the members of the set, subset, or population that have the highest probabilities of being true members of the set or nominal population in question (which is the fuzzy view). This distinction is the key and core concept of this view and it is capable of explaining how survey response rates of 10% to 15% of the nominal popuIuation can with a very, very high probability easily be random and representative samples of the population, if the initial samples were randomly drawn, particularly for large, complex, and dynamic populations that are "fuzzy" in nature, such as the populations in the 5 studies described above.

From fuzzy subset theory, it can be shown that in a "dynamic" population, there is always an effective subset (or population) that has a higher probability of being randomly drawn (located, accessed and sucessfully "captured as a respondent") than another subset; namely, the effective population is an intervening variable that classical sampling theory does not take into account because it is based on classical set theory and not fuzzy subset theory. This "inability to access and capture" condition was present in all 5 of the studies described above, and a wide variety of factors create large and important differences between the nominal and effective populations in these types of survey studies.

The question, therefore, becomes is the effective subset actually sampled biased or random. The answer depends on two factors, both of which were identified and explained in detail in the discussion on Castrophe theory above. The first of these two factor is whether or the factor creating the effective population is random or not. This "first factor," it should be noted, could also be the effects of several interacting factors netting out to zero and randomness (or not). If this "influencing factor" is essentially a random factor, then the effective population will statistically be a _very_, _very_ large random sample of the nominal population. However, if there is a bias, depending on its degree of severity, and the size of the effective population as a percentage of the nominal population, the effective population for large scale populations will be _very, very close_ to a very, very large random sample of the nominal population. So even with biasing factors, fuzzy subset theory predicts that one comes very, very close to a random sample at large scales, and this most probable fact is a reasonably strict deduction from fuzzy subset theory, and only counter-intuitive from the point of view of classical set and sampling theory. Further, this prediction is why a good metric and scale for "how biased" is so badly needed.

The second factor that affects whether or not the effective population and samples drawn from it will be random or not is the number of influencing factors operating in the situation. This factor and its operations have been previously explained above in detail. When many factors are inte[illegible]ting together to create the effective population and/or samples drawn from it, the net results are going to tend to be random almost all of the time when NNFSO and N are large. The reason that the samples will tend to be random is the same as that for the effective population as the same principle is operating in the second step of the process.

From fuzzy subset theory and all of the points made in all of the discussions above, it is not too difficult to see why in large scale surveys, response rates are always far, far higher than calculated by conventional methods and one has a far, far greater sample size (of the effective population) than one believes and is represented nominally, which means that one's data is much, much better than one would believe by traditional views. Also, in large-scale surveys of large, complex, and dynamic populations, the odds are very high that one's responder sample is a random sample of the nominal population when it is 15% to 20% of the nominal population. A simulation study could be designed to show that this particular outcome would be so most of the time. The value in conducting this simulation study would be the provision of exact points and cutting ratios rather than ... estimated points and rough rules of thumb given above.

Fuzzy subset theory strongly suggests that sample data of a reasonably sufficient size from large, complex, fuzzy, and dynamic populations is probablistically much better than we believe it is, according to the classical views; and this is good news, even if what the data tell us is not good news. People, therefore, may be ignoring and holding suspect perfectly good data. The common wisdom and older interpretations prevail, even when newer theory, simple logic, intuition, experience, and empirical data give clear messages to the contrary.

Interim Conclusions

We believe that the weight of argument, empirical evidence, and newer theories, at this time, is sufficiently strong enough to suggest that challenges of sampling bias based soley on arguments of low response rates are highly suspect and quite dubious at best, but most particularly so under the sets of conditions outlined in this paper. We blieve that under the sets of conditions outlined in this paper that allegedly biased samples are not as biased as those that claim bias believe, and that a great deal of reasonably good real world data is being dismissed and ignored when it should not be.

We have outlined the basis of our views, and studies that support our views, as well as numerous studies and simulations that need to be done to confirm, partially confirm, deny, or partially deny our views. We believe that the onus should be upon those who claim bias, particularly under the study conditions we have described, to prove their claims or disprove ours as vice-versa, which is the current case. It is our view that such claims should no longer be unquestioningly accepted by researchers, or accepted as proven fact until the appropriate evidence is generated to settle the question. The case against these claims, currently, is stronger than the case for these claims, and all newer theory strongly suggests that the clasical views and claims on these questions will not be supported or well supported when the appropriate studies and simulations are done.

The weight of argument, evidence, and newer theories (that all converge in the same place) is sufficient, we believe, to warrant this view, and to cast very serious doubts upon claims of alleged bias based on classical models and views. It is really a matter of which new theory or set of new theories that one wishes to chose to operate from to get to the conclusions we have drawn and the predictions that we have made. Our choice is Catastrophe theory supplemented by fuzzy subset theory, for all of the reasons we have stated in this paper. Catastrophe theory is the only theory (to date) that precisely explains and predicts the conundrum coin and its various operations. The very strong suggestion of these newer views, if they prove to be essentially correct upon empirical simulation, would be a very important discovery, not only for the areas of research methodology and statistics, but also for the areas of cognitive psychology and machine intelligence.

Real world data is much, much better than classical views say it is, or lead people to believe it is, and this is a problem of great importance and a problem that is very significant both theoretically and practically.

Some of us who worked in this world of "classically suspect data" developed an intuitive feeling about the problem and sought out models and theories to formalize our intuitions and guide research and thinking, which is still in its early stages, but all of the empirical data and the theory is supporting our views. And our views and the data are saying that we are essentially correct and empirically observing what we should be seeing, and that the problem is in the "hermetic urn" views which need to be revised, radically.

Classical survey and sampling theory is essentially seriously flawed and fundamentally incorrect and misleading in several important respects because it has not and does not incorporate fuzzy subset, Catastrophe theory and Chaos theory principles into its views. At a very minimum, therefore, this fact means that we just might now be more open and more accepting of large-scale, real world data rather than just dismissing it due to "flawed methodology," which just might save us in a number of different ways in the long run. Consequently, it is the "downside" problems and flaws of the classical views that are the most problematic and flawed aspects of the classical views, which is a model or view adoption cost that is rarely considered or considered in the evaluation of a view, or the decisions and claims that result from it.

As Student pointed the way for statisticians in the early part of the century in terms of understanding that the "very" small and the finite were different from the "typical laboratory model" of the day that was being used, the newer theories outlined in this paper are pointing the way for statisticians now in terms of showing that the large, complex and dynamic are very different from the "typical laboratory model" that is being used today. The large, fuzzy, dynamic and complex, and the small, well-defined and simple are very, very different from each other. We all need to begin to accomodate to and deal with this basic fact.

Catastrophe, Chaos, and Fuzzy Subset theory as theory names does not mean, or in any way imply, unpredictability or incomprehensibity. This misconception is one of the most common misconceptions of these theories currently. Although still under development, these theories are in fact just the opposite of unpredictability and incomprehensibility; namely, they lead to high predictability of events, phenomena, outcomes, and results, rather than paradox and contradiction.

There are two kinds of general phenomena in this world and not one; namely, uniform and conditional phenomena. There are many conditional phenomena in education (learning is but one), and these conditional phenomena need to be modelled with conditional theories, if we are to understand them adequately. Both models and views, therefore, are needed and needed in fully developed formulations to conduct reasonable and good research. Practicing researchers and theorists simply cannot be be expected to or made to play with one hand tied behind their backs. A great deal of new work is needed and has been needed for a very long time.

With just a little bit of theory, conceptualization, legwork, and study design modification, both the quality and certainty of all of large-scale survey data could be improved immmensely, as well as the cost-benefit ratios. And this is the good new, even if what the data has to say is bad news.

References

Angoff, W. (1988). Proposals for theoretical and applied development in measurement. Journal of Applied Measurement in Education, vol. 1, no. 3, 215-222.

Boakes, W. (1981). Personal Communication. California Board of Community Colleges.

Carifio J., and Shwedel, A. (1983). Utilizing federal reporting requirements to generate useful data at the local level. Paper presented at the New England Educational Research Association Conference, Ellensvile, NY, October; ERIC ED240321.

Carifio, J. Shwedel, A., and Biron. R. (1987). A comparison of community college responsers and non-responders to the VEDS student follow-up survey. Paper presented at the American Educational Research Association Conference, Washington DC, April; ERIC ED309798.

Carifio, J., Biron, R., and Shwedell, A. (1988). Results of the VEDS employer follow-up survey for Massachusetts community college students. Paper presented at Eastern Educational Research Association Conference, Miami, Fla, February; ERIC ED309799.

Caudill, M. (1990) Using neural nets, part II: fuzzy decisions. AI Expert, vol. 5, no. 4, April, 59-64.

Cochran W. (1983). Planning and Analysis of Observational Studies. Wiley, New York.

Coombs, A. (1964). A Theory of Data. Wiley, New York.

Cronbach, L. (1988). Playing with chaos. Educational Researcher, Sept., 1988, p 46-49.

Fazekas, N. (1989). Nurses' attitudes towards aids. Paper presented at the New England Educational Research Conference, Portsmouth, NH, April.

Glass, G.V., Peckham, P.D., and Sanders, J.R. (1972). Consequences of failure to meet assumptions underlying the analysis of variance and covariance. Review of Educational Research, Summer, 42, 3, 237-288.

Gleik, J. (1987). Chaos: making sense of a new science. Penguin.

Hacking, S. (1972). The Logical of Statistical Inference. Cambridge University Press, London.

Hayes, W. (1981). Statistics. Holt, Rinehart, and Winston, New York.

Hunt, V. (1989). Nurses' attitudes towards treating patients with aids. Paper presented at the New England Educational Research Conference, Portsmouth, NH, April.

Kaufman, A. (1976) Introduction to the Theory of Fuzzy Subsets. Academic Press.

Kerlinger, F. (1986). Foundations of Behavioral Research. Holt, Rinehart, and Winston.

Mandrebot, B. (1977). The Fractal Geometry of Nature. Freeman, New York.

McKeachie, J. (1979). The decline and fall of the laws of learning. Educational Researcher, October, 7, 8, 12-24.

Pagels, H. (1988). The Dreams of Reason: The Computer and the Rise of the Sciences of Complexity. Bantam Books, New York.

Paulson, S. (1982). Personal Communication. Oaklahoma Board of Regents.

Richards, J. (1972). A simulation study of the use of change measures to compare educational programs. American Educational Researc Journal, 12, 3, 299-311.

Rucker, R. (1987). ..ng Tools: The Five Levels of Mathematical Reality. Houghton Mifflin, Boston.

Scheffe, H. (1953). Sampling Theory. Wiley, New York

Wainer, H. (1986). Drawing Inferences from Self-Selected Samples. New York, Springer.

Wainer, H. (1989). Eeelworms, bullet holes, and Geraldine Ferraro: some problems with statistical adjustment and some solutions. Journal of Educational Statistics, Summer, vol. 14, no. 2, 121-140.

Zadeh, L.A. (1974a). Fuzzy logic and its application to approximate reasoning. Proceedings of the IFIP Congress, August.

Zadeh, L.A. (1974b). A fuzzy algorithm approach to the definition of complex, imprecise concepts. ERL Memo M474, University of California, Berkley.

Zeeman, C.E. (1976) Catastrophe Theory. Scientific American, 234, 4, April, 65-83.