International Statistical Review, 47 (1979) 111–127 Longman Group Limited/Printed in Great Britain

Representative Sampling, II: Scientific Literature, Excluding Statistics

William Kruskal[1] and Frederick Mosteller[2]

[1] *Department of Statistics, University of Chicago, and* [2] *Department of Statistics, Harvard University*

Summary

This paper describes and illustrates the meanings of 'representative sample' and 'representative sampling' as used in the extrastatistical scientific literature. Six categories seem adequate to describe these uses. The first five appeared also in the review of non-scientific uses in 'Representative Sampling, I' (Kruskal and Mosteller, 1979). We define and illustrate these in Sections 1–5 of the paper.

Meaning 1

General, unjustified acclaim for the data. Here the investigator gives the data a pat on the back by using a seemingly scientific term to raise its stature. No grounds, either in study design or empirical information, are given to support this usage, and hence we recommend that it be abandoned in scientific discourse.

Meaning 2

Absence (or presence) of selective forces. If grounds for the absence are given, this meaning may revert to Meaning 6, a vague term to be made specific. Otherwise it reverts to Meaning 1. In proclaiming the presence of selective forces, one is denying Meaning 6.

Meaning 3

Mirror or miniature of the population. The sample has the same distribution as the population. This rare feature carries with it severe constraints. We would avoid calling this a 'representative sample' because it seems much more.

Meaning 4

Typical or ideal case. An item from the population that represents it on the average or modally (or ideally). We call it a representative, but not a 'representative sample'.

Meaning 5

Coverage of the population. Samples designed to reflect variation, especially among strata. A sample containing at least one item from each stratum of a partition of the population is said to cover it for that partition. We would not use the term 'representative sample' for such samples.

The new meaning found in the scientific literature is that of a vague term; we describe it in Section 0.

Meaning 6

A vague term to be made precise. It often takes a good deal of space to describe a formal sampling scheme and it may be a convenience to have a term like 'representative sample' or 'representative sampling' to refer to the sample or the process generating it. We recommend that this usage be preserved for probability samples, and that the sampling method be described in the same work.

Motivation

We explore the meanings of the terms 'representative sampling' and 'representative sample' as they are used in the scientific literature. Our purpose is to discover to what extent these expressions have meaning, what those meanings are, and when these expressions may be profitably used in scientific discourse.

Introduction

Our first paper on representative sampling (Kruskal and Mosteller, 1979) dealt with the concept in non-scientific writing. We set up the following classification for uses of the terms 'representative sample' and 'representative sampling':

1. General acclaim for data
2. Absence of selective forces
3. Miniature of the population
4. Typical or ideal case(s)
5. Coverage of the population.

We illustrated these classifications by examples from non-scientific writing, and we discussed ambiguities and other problems in the concepts. We ended with the hope that the concept of representativeness would not be used in scientific writing unless it could be usefully sharpened in meaning.

The present paper continues the discussion by exemplifying concepts of representativeness and some of its cousins in scientific, but not primarily statistical, writing. The statistical literature itself is special, and we plan to cover it in a third paper.

We use much the same classification scheme as before, but somewhat extended, and we come to similar conclusions. Also, as before, our sample of usages is itself haphazard, but we hope with wide coverage. A new category we find is

6. 'Representative sampling' as a vague term for some specific kind of sampling procedure that is later on made precise.

Often a quotation we offer could fall into more than one category.

0 *'Representative sampling' as a vague term*

Since the use of 'representative sample' and 'representative sampling' as vague terms did not arise in our paper, 'Representative Sampling, I', and is included here, we take up this usage first and then review the other five meanings for the scientific literature. We number this Section 0 so that the remaining sections have their natural numbers.

In the extrastatistical scientific literature, as in the non-scientific literature, 'representative sample' and 'representative sampling' are used in a number of different senses. These uses lead to a high degree of ambiguity and considerable opportunity for confusion. It may be surprising then to argue, as we shall, that using 'representative sampling' as a vague or general term which is later made specific has merit. In statistics, such vague terms are already in fruitful use in other contexts; standard examples follow, each with some possible precise specifics:

Average or location: arithmetic mean, geometric mean, median, midrange, (various) trimmed means, (various) weighted means, the biweight.

Variability or dispersion: standard deviation, inter-quartile range, mean absolute deviation from the median.

Association: Pearson correlation coefficient, index of correlation, Kendall's tau, Spearman rank correlation, various measures of association for cross classifications.

Such general terms as 'average', 'variability', and 'association' help our exposition, provided that their ambiguities are kept in mind and resolved in any particular application. The value of such a term turns to a large extent on how nearly univocal is the intuitive idea behind it.

Average and variability get high intuitiveness marks on our report card. Association does not do so well, for it includes at least two different intuitive ideas,

1. Covariation . . . two variables in some sense move together, as for the correlation coefficient.
2. Departure from independence in some sense, as for the index of correlation and the contingency coefficient.

The distinction is discussed by Kruskal (1958).

In a vein similar to 'average', 'variability', and 'association', we think it may be profitable to use the expressions 'representative sample' and 'representative sampling' as vague terms that will be made specific in the same work, provided they refer to some method of probability sampling. Among these types of samples, popular ones are

random samples
stratified samples
systematic samples with random starts
cluster samples.

In short, we would regard the use of 'representative sampling' as a less awesome and more informative synonym for 'probability sampling', provided that the specific mode of sampling is then specified.

Probability samples are rarely miniatures of the sampled population. The sense in which a probability sample is representative is that it ordinarily can be used as a basis for responsible inference.

We prefer to exclude non-probability sampling methods from the 'representative' rubric. Thus we would exclude quota sampling, systematic sampling with non-random starts, and chunk sampling (whatever data are readily available), although we recognize that such non-probability samples may sometimes be the only available ones and may provide useful information.

Although it is not a precise example of this vague usage of 'representative sample', the following quotation offers a partial illustration: 'This paper reports data from transcriptions of the speech of a representative sample of black first-graders in the inner city. . . . Twenty monolingual children were randomly selected from the eight first-grade classrooms in a public school in Brownsville, New York City. Reflecting the neighborhood population, all the children were black and all had been born in mainland U.S.A. Their families had incomes at the poverty level.' P. 104 of 'Dialect in the Language of Inner-City Children', by Beatrice Levy Kachuck, *The Elementary School Journal*, **76** (November 1975), 104–112.

We have two reservations about this quotation. First, it is not clear to us from the rest of the article that its author uses the term 'randomly selected' in the strict statistical sense of a random sample. More fundamentally, the reader may ask to what extent eight first-grade classrooms from one school can be representative of all 'black first-graders in the inner city'. Nonetheless, this example helps to illuminate the proposed idea of vague usage.

The following quotation from the statistical literature shows the idea of vague usage more completely than does the previous example: 'The HANES is the first program to collect measures of nutritional status for a scientifically designed sample representative of the U.S. civilian, non-institutionalized population in a broad range of ages, 1–74 years. . . . The probability sample design permits estimates to be made for the total population. . . . The sample design provided that data from a representative subsample would become available after the first year and a half of data collection. . . .

'These preliminary findings are based on the examination of 10,126 persons aged 1–74 years in a representative subset of 35 of the 65 PSU's making up the total sample. A sample of

14,147 persons was selected . . . [to make up] a probability sample of the total U.S. population . . . the program succeeded in examining only 10,126 of these sample persons' (p. 202). 'Preliminary findings of the HANES program', *Statistical Reporter*, June 1974, pp. 201–207.

The next example has exactly the spirit we require, but stumbles a bit at the end. 'The problem of securing a sufficiently representative sample of the Army universe covered by a particular study was not an easy one, but its solution did not lead to new mathematical constructs. Because the situations in which sampling was done in the Army were quite different from any which would be encountered in civilian life, it has seemed unlikely that the Research Branch experience in sampling would prove especially instructive for future civilian research. Therefore the description . . . is kept very brief here. . . . Ordinarily, sampling was a two-step process. First, there was sampling of organizations. . . . Second, within an organization there was sampling of individuals. The selection of individuals at a given sample point ideally involved procedures which might be called probability sampling.' Samuel A. Stouffer in Samuel A. Stouffer, Louis Guttman, Edward A. Suchman, Paul F. Lazarsfeld, Shirley A. Star, and John A. Clausen, *Measurement and Prediction*. Princeton, N.J.: Princeton University Press, 1950, pp. 709–710.

There follows after this a description of systematic sampling, every *n*th member from a duty roster, and elsewhere (p. 715) of stratified sampling and random selection. For stratified sampling with random selection within strata, the example is completely satisfactory for our recommended use of 'representative sampling' as a vague term. But choosing every *n*th soldier with a non-random start would violate our recommendation that only probability sampling be used. Indeed, in working for the organization whose sampling methods are being described one of us (FM) found himself sampling every 32nd soldier on roster lists with man after man turning out to be the same special kind of sergeant. What had happened was that after going through lists of many different lengths, we came to a pile of lists for bunkhouses containing 32 men each with identical rank composition ordered by rank. We started near the top line and as long as we counted correctly the same rank would appear each time. A single random start would not have cured this problem.

Just one experience like this cools original ardor for systematic sampling. Even though the difficulty was detected and corrected in the sampling described, in a less blatant situation it might well not have been. Furthermore, had clerks been carrying out the sampling, the difficulty might not have been detected.

Our final example meets all the requirements of our recommendation. 'Both the Harris organization and NORC drew samples alleged to be representative of all English-speaking, non-institutionalized residents of the continental United States aged 18 and over. Selection of individual respondents proceeded through two phases. A first series of procedures selected specific geographic locations; blocks or segments of the country were selected with the probabilities of selection being proportional to the area's population.

'Within each block or segment, clusters of five (NORC) or eight (Harris) individuals were interviewed. [A footnote gives further technical details about the structure of the probability sampling.]' Charles F. Turner and Elissa Krauss, 'Fallible Indicators of the Subjective State of the Nation', *American Psychologist*, Vol. **33**, May 1978, 456–470, p. 460.

1 *Rhetorical usage; general acclaim*

The praising rhetorical use of 'representative sample' appears in the scientific literature as in the general literature. The usage, when favorable, seems to mean something like 'My sample will not lead you astray; take my word for it even though I give you no evidence'. We deplored this use in the non-scientific literature, and we recommend even more strongly that scientists avoid it.

Writers, who desire to say something more friendly about a set of data than, for example,

'these data just happened to come to hand, and I have no notion of the process that led to them or of relations between the target and sampled population' use 'representative sample' without realizing that they are giving a pat on the back to a set of data. By calling attention to this frequent empty usage we may make scientific authors and reviewers more vigilant and thus reduce its frequency. (This usage is an attractive one for a writer and we shall not be surprised if it is found in our own writing!)

We chose our examples to exhibit the diversity of scientific fields using this back-patting meaning, and the many examples reflect its frequency. We begin with one from mathematics. 'It is in the nature of such a brief survey that not all valuable contributions to the theory of Dedekind sums could be mentioned. I have tried to select a representative sample, but wish to apologize to all those mathematicians [whose work might have been mentioned but is not].' P. xi of Preface by Emil Grosswald to *Dedekind Sums*, by Hans Rademacher and Emil Grosswald, The Carus Mathematical Monographs Number 16, The Mathematical Association of America, 1972.

Next we jump to nuclear engineering policy, for an unusually clear-cut example of the rhetorical usage, 'One GAO official familiar with the investigation told *Science* that these plants were not chosen as a representative sample. Nonetheless, Theodore Taylor, a nuclear physicist . . . says that he thinks the security measures described in the GAO report reflected the general level of protection afforded in mid-1972.' P. 1114 of 'Nuclear safeguards: holes in the fence', by Robert Gillett, *Science*, **182** (14 December 1972), 1112–1114.

It is not quite such a long jump to astronomy of the solar system, 'Most investigators believe that the many types of meteorites found on the earth provide a representative sample of the solar system's raw materials'. P. 772 of 'Presolar grains: isotopic clues to solar system origin', by Allen L. Hammond, *Science*, **192** (21 May 1976), 772–773.

In a similar vein, we find an example from biology, 'If blood proteins are a representative sample of proteins coded by structural genes, the most similar species should have the most similar blood proteins. . . . It remains possible that King and Wilson analyzed a non-representative sample of structural genes . . . .' P. 446 of 'Evolution of DNA: changes in gene regulation', by Gina Bari Kolata, *Science*, **189** (8 August 1975), 446–447.

Turning from the natural to the social sciences we cite a 1940 example. '. . . it is not possible, and with our modern knowledge not necessary, to study everyone (in the city). In lieu of this we resort to the analysis of representative samples. . . .' P. 59 of Louis Wirth, 'The urban society and civilization', pp. 51–63 of *Eleven-Twenty-Six*, Louis Wirth (Editor). University of Chicago Press, 1940.

Next, a recent example from a study of the Bar, 'First there will be interviews with a representative sample of 750 practicing attorneys in Chicago, of whom approximately 70 per cent are members of the [Chicago Bar Association]'. P. 20 of the Annual Report of the Russell Sage Foundation, 1974–1975.

A book by the historical demographer D.V. Glass provides two examples, one referring to the eighteenth century, the other Glass's own, 'The eighteenth-century political arithmeticians began with a handicap: none of the materials available to them was as useful as the Marriage Act statistics had been to Gregory King. They tried to . . . [collect] their own data and [to persuade] . . . acquaintances to cooperate . . . . By modern standards the samples were certainly not well-designed [but] . . . "representativeness" was the aim. . . .' P. 67 of *Numbering the People*, by D.V. Glass. Farnborough, Hants: Saxon House, 1973. '. . . [we] were able to have extracted by the General Register Office, national representative samples of the marriages of 1871 and 1951.' P. 191 of the same book.

As a final sociological example, here is one with the term 'statistically representative', 'The inquiry used what we today would call depth interviews with a small but statistically representative sample of the manual workers of Sheffield'. P. 119 of *Marienthal/The Sociology of*

an Unemployed Community, by Marie Jahoda, Paul F. Lazarsfeld, and Hans Zeisel. Chicago–New York: Aldine-Atherton, 1971. This book was first published in 1933 as *Die Arbeitslosen von Marienthal.* The 1971 English translation was done by the original authors, together with John Reginall and Thomas Elsuesser, and it contains fresh material: a new 'Foreword' and a new 'Afterword', the latter by Hans Zeisel. The above quotation, which refers to an early sample survey, published in 1919, appears in footnote 47 of the 'Afterword'.

The next example comes from psychology. It considers the ubiquitous problem that people may behave differently as subjects of an investigation than otherwise. 'Our impression was that these mothers did not alter their behavior in front of the camera and that our samples were representative of the mothers' communication efforts.' P. 402 of 'The development of language-like communication without a language model', by Susan Goldin-Meadow and Heidi Feldman, *Science*, **197** (1977), 401–403.

Anthropology provides a properly skeptical example, 'What the excavator believes to be a representative sample of information about the lifeways of a prehistoric group may, in fact, reflect just one or a few specialized activities. . . . What we have is so unlikely to represent the whole fairly, that every small addition of empirical data can be expected to change these speculations radically.' P. 238 of 'By their words you shall know them: cultural developments in the Paleolithic', by L.G. Freeman, pp. 234–261 of *Hominisation and Verhalten*, Kurth/Eibl-Eibesfeld (Editors), Stuttgart: Gustav Fischer, 1975. Linguistics gives a non-skeptical example, published in 1928 by a physicist! 'Suppose one takes a large representative sample of written English, counts the number of times each word appears and arranges the words in order of decreasing frequency of occurrence.' P. 300 of 'Statistics of Vocabulary', by E.U. Condon, *Science*, **67** (1928), 300 only.

A public policy example provides diversity, 'The Committee's recommendations for improvement of the survey at Dulles . . . include refining the . . . procedures to help ensure the representativeness of the sample . . . Recommendations for . . . the JFK survey deal with selecting a sample population that adequately represents the geographical area . . .'. 'Plans for Concorde survey reviewed', ABASS Newsletter, National Academy of Sciences – National Research Council, April 1977, p. 4.

We include an example from the field of consumer safety, medical subfield, 'Fifteen samples of consumer spackling and patching compounds were purchased at hardware stores in the New York City area . . . (O)ur analysis of [the] fifteen representative samples . . . has shown that five contained appreciable amounts of . . . asbestos minerals.' Pp. 551, 553 of 'Exposure to asbestos in the use of consumer spackling, patching, and taping compounds', by A.N. Rohl, A.M. Langer, I.J. Selikoff, and W.J. Nicholson, *Science*, **189** (15 August 1975), 551–553.

This example has special interest because the individual purchases are called representative samples, not the whole group of fifteen. The article presents no discussion of how the hardware stores were chosen. We end this section with an example of humorous satire. 'I must say, clever men are fun. It struck me afresh, just reading a sample The Institute of Physics sent me in advance of contributions to *A Random Walk in Science.* (Naturally the sample was representative.)' P. v of Foreword by William Cooper to *A Random Walk in Science*, R.L. Weber (compiler) and E. Mendoza (Editor). London: Institute of Physics, and New York: Crane, Russak, 1974.

In a few instances of this usage, the authors may be able to supply more detail in the method of selection that will give the reader some grounds for evaluating a sample and its possible distortions. Naturally, we urge that this be done.

2 *Absence or presence of selective forces*

A frequently found sense of 'representative sample' is that of non-selectiveness, or at least claimed non-selectiveness. The idea is expressed in abstract terms thus 'It is important . . . to

avoid selecting [instances] . . . in a way which will prejudice outcome. If inductive generalizations are to be reliable, they must be based upon representative samples. Unfortunately, we can never be sure that our samples are genuinely representative ones.' P. 57 of *Logic*, by Wesley C. Salmon. Englewood Cliffs, N.J.: Prentice-Hall, 1963.

An excellent ancient example, going back at least to Cicero, is found in the following quotation[1] from him: 'Diagoras, surnamed the Atheist, once paid a visit to Samothrace, and a friend of his addressed him thus: "You believe that the gods have no interest in human welfare. Please observe these countless painted tablets; they show how many persons have withstood the rage of the tempest and safely reached the haven because they made vows to the gods." "Quite so," Diagoras answered, "but where are the tablets of those who suffered shipwreck and perished in the deep?"' P. 330 of *On the Nature of Gods*, by Marcus Tullius Cicero, translated by Hubert M. Poteat. Chicago: University of Chicago Press, 1950, Book 3, 37.

The Kaplan book just mentioned in the footnote goes on to point out how the bias of selection can be subtle, 'Suppose we are interested in . . . the changing size of the American family. We might begin with a fair sample [N.B.!] of the present population, find out from each person in the sample how many siblings he has, then how many his parents had . . . and go back. . . . But this sampling plan is biased in favor of larger families of the past, for the larger the family the more likely it is to have left descendants . . . [and] childless families in [the past] . . . cannot be represented at all.'

Of course that sampling is probably also biased in favor of large families of the present. Aside from the problem of childless families, however, this example is one in which analytic statistical tools might well permit compensating for the family size selection bias: one would give greater weight to the smaller families with at least one child.

A variety of other kinds of selective forces may be found, 'The sample [of university physicists] . . . is not representative: it heavily over-represents eminent scientists [as against non-eminent ones]'. P. 378 of 'Scientific output and recognition: a study in the operation of the reward system in science', Stephen Cole and Jonathan R. Cole, *American Sociological Review*, **32** (1967), 377–390; or again 'The oft-quoted National Science Foundation report . . . provides data on the status of respondents to a questionnaire which was sent to the respondents of a previous questionnaire. . . . It is doubtful whether this kind of sample is a representative one.' P. 242 of 'Unemployed engineers', a letter from Ephraim Weiss to *Science*, **178** (20 October 1972), 240 and 242. 'Ehrenberg was, of course, aware that his sample was not representative of Krupp employees as a whole because of their long time employment with Krupp,

[1] This is almost the oldest admonition we know towards having a control group. The oldest we know appears at the beginning of the Book of Daniel.

Cicero's anecdote about Diagoras was later told in modified form of Diogenes the Cynic by Diogenes Laërtius (1905, p. 239) and still later – most gracefully – by Francis Bacon (1960, p. 50, Aphorism 46). We are grateful to Robert K. Merton for leading us to the history and use of this splendid statistical story, evidently a favorite one of Merton's, who quotes it at least twice in his writings (1934, p. 543 and 1977, p. 34). It must be liked by others as well, for we note its appearance in two methodological publications, one by Abraham Kaplan (1964, p. 241) and by Adolf Grünbaum (1976, pp. 216–217). Apparently for millennia, scholars have been struggling with representativeness as a concept in statistics and in scientific methodology more generally.

Bacon, Francis (1960). *The New Organon*, edited by Fulton H. Anderson on the basis of the 1863 translation by James Spedding, Robert Leslie Ellis, and Douglas Denon Heath. New York: Liberal Arts Press.

Diogenes Laërtius (1905). *The Lives and Opinions of Eminent Philosophers*. Translated by C.D. Yonge. London: George Bell.

Grünbaum, Adolf (1976). 'Is falsifiability the touchstone of scientific rationality? Karl Popper versus inductivism', pp. 213–252 of *Essays in Memory of Imre Lakatos*, edited by R.S. Cohen, P.K. Feyerabend, and M.W. Wartofsky. Dordrecht: Reidel.

Kaplan, Abraham (1964). *The Conduct of Inquiry*. San Francisco: Chandler.

Merton, Robert K. (1934). 'Recent French Sociology'. *Social Forces*, **12**, 537–545.

Merton, Robert K. (1977). 'The Sociology of science/An episodic memoir', pp. 3–141 in *The Sociology of Science in Europe*, edited by Robert K. Merton and Jerry Gaston. Carbondale: Southern Illinois University Press.

their higher average income and their occupational distribution. . . . He was, however, willing to sacrifice this criterion in favor of obtaining precise data.' P. 84 of *Empirical Social Research in Germany 1848–1914*, by Susanna Petra Schad. The Hague: Mouton, 1965. 'Comparing tonight's [election eve's] returns from a county with its previous voting patterns takes into account that the counties reporting first are not a representative sample. Counties with early complete returns may tend, in some states, to be Republican counties; in others, Democratic counties. At any rate, why hope they are typical or representative?' P. 44 of *Data Analysis for Politics and Policy*, Edward R. Tufte. Englewood Cliffs, N.J.: Prentice-Hall, 1974. '. . . we have never managed to explore the solar environment beyond the narrow and totally unrepresentative strip traced out by the earth's orbit. . . . our ability to sample the solar radiation is extremely limited and what we do see is totally unrepresentative of what might be found at other latitudes.' P. 845 of 'Exploratory journey out of the ecliptic plane', D. Edgar Page, *Science*, **190** (1975), 845–850.

A difficulty flowing from selectiveness is described in the following quotation: '. . . in order to find 40 schizophrenics who were matched [with controls] in age, sex, IQ, and family income, over 500 records had to be examined. . . . The question then arises: if 90% of possible subjects are discarded from the study because they can't be matched, how representative can the results be? The use of multiple measures for matching purposes thus creates the problem of the unrepresentativeness of the samples.' P. 83 of 'Problems of multidimensional evaluation', by Robert Plutchik, *Annals, New York Academy of Sciences*, **218** (1973), 78–86.

A major reason for avoiding selective forces, or at least for understanding them well enough to untangle matters during analysis, is that we want our inferences from a sample to be about a population of genuine interest, not about a population of adventitious distortion. This distinction between target and sampled populations has been described by many authors, for example by Donald T. Campbell and Julian C. Stanley (1963) as follows: 'Consider . . . an experiment on teaching in which the researcher has been turned down by nine school systems and is finally accepted by a tenth. This tenth almost certainly differs from the other nine, and from the universe of schools to which we would like to generalize. . . . It is, thus, non-representative. Almost certainly its staff has higher morale, less fear of being inspected, more zeal for improvement than does that of the average school.' 'While sampling representativeness is impossible of perfect achievement and is almost totally neglected in many sciences . . ., it both can and should be emphasized as a desideratum . . .' (p. 19).

A related quotation is 'The major limitation to the practical use of RDA's [recommended daily allowances for nutrition] is that they are based on data from small and possibly unrepresentative samples that have been extrapolated to populations of all types'. P. 61 of 'The requirements of human nutrition', by Nevin S. Scrimshaw and Vernon R. Young, pp. 51–64 of *Scientific American*, September 1976.

Finally, we cite an example in which a newspaper article about a scientific publication shifts the terminology. The scientific publication, dealing with the reputations of medical schools on the basis of a questionnaire to medical faculty members, says 'The response to the questionnaire by faculty members of different ranks was not fully representative of the population. Full professors represented 29 per cent of the respondents, but only a fifth of the sample . . . There are . . . several plausible reasons for this unrepresentative response rate . . .'. From pp. 664–665 of 'The reputations of American medical schools', by Jonathan R. Cole and James A. Lipton, *Social Forces*, **55** (1977), 662–684.

The scientific article then goes on to say that, although responses are differentially selected by academic rank, the actual ratings of medical schools seem not to be affected because of very high rank correlations among the ratings by respondents of different ranks. (Questions might be raised about that rank correlation criterion, but they lie outside the scope of this paper.)

In at least one press report of the study, however, the above argument is compressed in a way that might leave the reader supposing that the authors had no grounds for their view, 'Cole and Lipton asked 2,049 faculty members to rate 40 schools apiece. The researchers got 583 usable responses, giving them what they consider a representative sample.' Victor Cohn, '3 medical schools get low ranking in national study', *Washington Post*, 9 March 1978. Since Cole and Lipton said the response was not representative of the population, Cohn probably means that the authors think the generalization is satisfactory.

3 *Miniature of the population*

Closely related to the concept of non-selective forces or lack of bias is that of the representative sample as a miniature, or perhaps a mirror, of the population from which the sample is drawn. For example, 'Comparisons of demographic characteristics and geographic distributions of respondents with 1970 Census data showed the survey sample to mirror closely the household composition of the city's population'. P. 1 of *BSSR/The Newsletter of the Bureau of Social Science Research*, July–October 1973.

The natural stringent definition of a mirror or miniature would presumably be the following when all variables are categorical: Each item in the population has associated with it a vector whose coordinates are names of categories. The first coordinate might carry one of the two names for sex; the second might give age in years; etc.

Similarly a sample consists of items and their associated vectors of category name coordinates. For a miniature or mirror, the natural stringent definition is identity of the sample's empirical joint distribution with the population's.

For a modest sized sample to mirror a population, the variables would have to be few and have few categories unless the population had a special structure that was easy to represent. If the population were composed of half men and half women, then equal numbers of each would form a sample that was a miniature if no other variables were relevant. But usually they are, and adding them would in general distort the mirror.

If we think of a population with many dichotomous independent variables, the sample size required for a miniature quickly becomes large. For example, suppose that we have three dichotomous independent variables, and suppose that the less frequent categories for the three variables have proportions 0.1, 0.1, and 0.2. Then the proportion in the population falling into the less frequent category on all three variables is 0.002, and we need a sample of at least 500 (= 1/0.002) to mirror the population. At least one item is required for the smallest stratum. If the population is not so easy to describe, much larger samples might be required.

It is often possible to mirror discrete multivariate distributions closely for single variables or pairs of variables, but when we are required to mirror populations with many variables the effort gets out of hand, or else one needs the idea of approximate mirrors.

A drop of well-mixed miscible fluids or a sip from a martini may well mirror the population of liquid from which they are drawn. Mirroring a continuous population with a finite sample would have to be a matter of degree. We shall not discuss it in this paper.

Among probability sampling methods, the stratified sample with stratum sample sizes proportional to strata sizes comes close to ensuring a mirror for the set of variables used in the stratification, but not necessarily for the response variables. A large overall random sample may come close to duplicating proportions, but it is not as sure to mirror the strata because the sample numbers in the strata are random variables.

We turn now to other illustrations of the mirror idea. 'A sample "truly" represents a group when the mode of variation within the sample is the same as the variation within the group at large: this is what is meant when we say we have an "unselected" sample.' P. 337 of 'Mathematical *vs.* scientific significance', by Edwin G. Boring, *The Psychological Bulletin*, **16** (1919), 335–338.

The miniature idea arises naturally in the political science literature, for example (and we quote a bit more extensively), 'The metaphors of portrait, map, and mirror . . . all are renderings of an "original" in a medium different from it . . . [but there are] examples in which the representing is done by something of the same kind as the original . . . a miniature or condensation of the original, or a part of the original that can be used to stand for the rest' (pp. 72–73). 'Some theorists think of [a legislature] as a representative sample of the whole nation, whether chosen mathematically or by lot. And others see it as a body composed entirely of representative men' (p. 73). 'Thus it is easy to suppose that a legislature must be a representative sample in order to represent truly' (p. 73). 'Behind all the applications of the descriptive view . . . hovers the recurrent ideal of the perfect replica, the flawless image, the map which contains everything, including "the map of the map; the map of the map, of the map, and so on endlessly". But that ideal may well be chimerical, and therefore dangerous. Perfect accuracy of correspondence is impossible. . . . so it is always a matter of our purpose: Is it information that we need, and, if so, what kind of information – what features are to be reproduced, what will be significant?' (pp. 86–87). From *The Concept of Representation*, by Hanna Fenichel Pitkin. Berkeley and Los Angeles: University of California Press, 1967.

We end this section with a modestly humorous quotation that deals with the miniature concept. 'The Chairman . . . said, "Are you sure that the parking lot contains a truly random sample of modern society?" "Well, maybe not," said the O.E., "but we did the best we could. We generated a selection list using a table of random numbers and a set of automobile ownership probabilities as a surrogate for socio-economic class. Then we introduced five racial categories, and an equal male–female split. We get a stochastic sample that way, with a kind of 'Roman cube' experimental protocol in a three-parameter space." "It sounds complicated," said the Chairman. "Oh, no. The only real trouble we've had was when we had to find an Amerindian woman driving a Cadillac . . ."'. From 'The Great Foucault pendulum caper', by W. Lewis Hyde, *The Chronicle of Higher Education*, 16 August 1976, p. 16 only.

Since the idea of a sample as a mirror or miniature of the population is rarely appropriate, we would not use 'representative sample' to describe it when it applies, because then so much more is being offered and because special constraints apply to the sample and the population. We do not feel it wrong to use 'representative sample' for the miniature, but rather that it understates the attractive properties of such a sample.

Ordinarily this use of representative sampling could not agree with our vague use described in Section 0, because a miniature is usually constructed purposefully rather than through a process of probability sampling.

4 *Representative as typical; the single case*

One sometimes sees 'representative sample' used to mean a sample of typical units, for example, 'It was the author's [Gules's] intention to study the "average male delinquent" and he felt that the inmates of Flehingen (n = 105) were representative in that respect. . . . He did not mean that they represented [sic] a representative sample in the statistical sense but that in comparison to other delinquents, in terms of type of delinquency, social background etc. the Flehingen group was rather typical. The only two variables which differentiate his population from the average delinquent were religion (almost all of his subjects were Catholics) and degree of incorrigibility (the majority of his population . . . were unfit to return home . . .).' Pp. 138–139 of *Empirical Social Research in Weimar-Germany*, by Susanne Petra Schad. Paris–The Hague: Mouton, 1972.

As the example illustrates, typicalness may be difficult to define. We suppose that such samples first of all are composed of representatives of the target population. After that several possibilities come to mind: a single item chosen close to the centroid, if appropriate, or perhaps from the modal class offers two forms of typicality. Or 'typical' may mean that averages taken

in the sample are close to those from the population. It does not seem to mean a probability sample or a random sample of size one.

Since this idea is unclear, we would not use the expression 'representative sample' to refer to it. Instead, we would urge a careful statement about what is typical of the sample.

The idea of typicalness naturally leads to samples of one, although we have not found pointed examples of this. Here are two almost-examples. 'Table 1 contains data for the dimensions of a number of representative tumors. . . .' P. 153 of 'Hyperbasic hydrogen therapy: a possible treatment for cancer', by Malcolm Dole, F. Ray Wilson, and William P. Fife, *Science*, **190** (10 October 1975), 152–154.

The next example sounds like an average rounded to whole numbers. 'Extensive field observations have revealed that a representative lion pride contains two adult males and seven adult females.' P. 671 of a review by Howard S. Hoffman of *Growing Points in Ethology*, edited by P.P.G. Bateson and R.A. Hinde, New York: Cambridge University Press, 1976. The review appears in *Science*, **195** (18 February 1977), 671–672.

The two examples above do not use the word 'sample', but its use may be found, usually in the plural. Two examples are 'Representative samples of the sternum, right clavicle, . . . were examined. . . .' P. 141 of 'Spongy bones in prehistoric America', by Mahmoud Y. El-Najjar and Abel L. Robertson, Jr., *Science*, **193** (9 July 1976), 141–143. 'The armchair approach generally consists of devising a list of general factors that are important. . . . Representative samples are to be found in . . . [a list of citations follows].' P. 86 of *Studies in the Quality of Life/Delphi and Decision-making*, by Norman C. Dalkey with Daniel L. Rourke, Ralph Lewis, and David Snyder, Lexington, Mass.: D.C. Heath, 1972.

Three further examples from diverse parts of science now follow: '. . . the first three Mariners sampled only a small and unrepresentative fraction of the Martian surface.' P. 107 of 'Mars', by James B. Pollack, *Scientific American* (September 1975), 107–117. '. . . a mere 20 episodes embracing every region of the globe and a span of 16 centuries just cannot be accepted as representative: the span is too great'. P. 43 of a review by J. David Singer of *Military Deterrence in History*, by Raoul Naroll, Vern L. Bullough, and Frada Naroll. Albany: SUNY Press, 1974. The review appeared in *Science*, **189** (1975), 43–44. 'William Fontaine . . . found that Negro scholars . . . tended to make use of strikingly dramatic rather than representative data.' P. 16 of 'Insiders and outsiders: a chapter in the sociology of knowledge', by Robert K. Merton, *Amer. J. Sociology*, **78** (1972), 9–47.

The idea of typicalness is closely related to that of absence (or presence) of selective forces, category 2 in our list. That is, the authors of these examples suggest that they offer results 'close to average'. We have not run across the Emersonian ideal sense in the scientific literature, except for brief descriptive mention in the book by Pitkin quoted above in Section 3. Emerson thinks of the representative sample as the ideal that others should be compared with – the best. Thus if one were able to select from a batch of seedlings one that had all the desirable properties of its kind, it would be the Emerson ideal. (See the discussion of Emerson's *Representative Men* in our 'Representative Sampling, I'.) No doubt this usage occurs in the scientific literature, but we would prefer not to call it a representative sample when we could so much more informatively call it an ideal representative or a perfect example and avoid the concepts of central, average, or middling.

5 *Coverage of the population's heterogeneity*

An important explication of the term 'representative sample' is that the sample include a wide, heterogeneous collection of cases from the population. A good example of this usage comes from the anthropological literature, '. . . we wanted . . . [to draw] samples of communities that would be representative, according to specifiable dimensions, of large groups of Yoruba communities' (pp. 6–7). 'the villages [are] to show a variety of such features as size, level

of prosperity, closeness to main roads and urban centers, and traditional way of life' (pp. 14–15). 'The first step . . . was selection of specific villages which reflected gross differences in factors presumed to relate to integration . . . size, nearness to highways, economic level, religious diversity, and socio-cultural change' (p. 185; see also pp. 186–189). From *Psychiatric Disorder Among the Yoruba*, by Alexander H. Leighton, T. Adeoye Lambo, Charles C. Hughes, Dorothea C. Leighton, Jane M. Murphy, and David B. Macklin. Ithaca, N.Y.: Cornell University Press, 1963.

A second example comes from the public policy area, 'This function peaks sharply around $N = P/2$ in contradiction to the empirical observation that there are far more small alliances in existence than medium or large ones. We thus seek guidance from a representative sample [of] arbitrary functions that decrease monotonically with increasing population.' P. 50 of 'World population, human disaster and nuclear holocaust', by J. Calvin Giddings, *Bull. Atomic Scientists*, Sept. 1973, vol. **29**, pp. 21–24 and 45–50. Giddings looks at four diverse functions of N, but one might question whether so few functions can provide the desired heterogeneity.

An example that might come under this heading arises in military history, 'As Xenophon's comparison of Rhodian and Persian sling missiles indicates, the standard missiles varied widely in size and weight. Measurement of a representative sample of biconical and ovoid stone missiles . . . indicates how broad the range could be.' P. 39 of 'The sling as a weapon', by Manford Korfmann, *Scientific American*, October 1973, pp. 35–42. One could alternatively interpret this to mean that the representative sample (general acclaim sense) shows the heterogeneity of the *population*.

In earlier examples of this section, the samples are said to be representative because they display in some way the heterogeneity of the population. In the following soil testing example, however, although several subsamples are drawn, a mixing operation destroys the identifiable variation among the subsamples. The hope presumably is that the mixture is more typical of the garden (Section 4) than any individual sample. Or perhaps the soil actually tested in the laboratory is an arbitrary small part of the mixture in the containers. 'As the individual samples [from your garden] are taken, place them in a clean pail or container and mix them well to insure a representative sample.' From a 1971 form letter from the Soil Testing Laboratory of Michigan State University to citizens requesting a soil analysis.

The coverage implied in the following example may be less than the author intends because the sense impressions of the student may well be sharply selected or limited. '. . . the raw data of scientific research arise from the contact of an observer with his environment. . . . A student in a brief period of time will have seen, heard, tasted, smelled, and felt a representative sample of the material world.' P. 102 of 'Recent trends in social science', by Beardsley Ruml, pp. 99–111 in *The New Social Science*, Leonard D. White, Editor. The University of Chicago Press, 1930.

In 'Representative Sampling, I' we defined coverage in a special way. Suppose that a population has been stratified (partitioned) in a way that interests the investigator. If the sample contains at least one item from each stratum, then the population is 100 per cent covered. If some strata have no representatives, then the coverage is the percentage of the population whose strata are represented in the sample. Note that the sample need have no proportional relation to the sizes of the strata. A linguistic example is 'The idea that governed the selection . . . of the text . . . was that it should be synchronic, representative of a wide range of styles, and accurate. . . . To insure representativeness, the 500 samples [of text] were distributed among fifteen categories, representing the full range of subject matter and prose styles, from the sports page of the newspaper to the scientific journal and from popular romantic fiction to abstruse philosophical discussion.' Pp. xviii–xix of the Introduction to *Computational Analysis of Present-day American English*, by Henry Kučera and W. Nelson Francis. Providence: Brown University Press, 1967. Here the partition sets are newspaper sports pages, scientific journals, etc.

We jump to the entirely different area of food chemistry for another apt example of coverage. 'We have collected . . . nearly 500 certified samples of honey. . . . For 84 representative samples selected from these, we have obtained $\delta^{13}C$ values. Samples from 34 states included 37 named floral types from 17 plant families. The criteria for sample selection were such that (i) the geographic distribution was proportional to that of honey production, (ii) for several floral types and blends produced in greatest amount, samples were selected from as wide an area as possible, (iii) at least one example of each commercially significant type was included, (iv) natural mixtures . . . were selected from many areas, and (v) honeydew honeys . . . were included.' P. 892 of 'Carbon-13/Carbon-12 ratio is relatively uniform among honeys', by Landis W. Doner and Jonathan W. White, Jr., *Science*, **197** (26 August 1977), 891–892. Notice that in both the above quotations 'representative sample' refers to but one element in the large sample that is supposed to provide adequate coverage. Note also, in the apiarian quotation, the beginnings of a requirement that goes beyond coverage *per se* to coverage that is proportional to population amounts.

A third example of coverage comes from the field of education, 'The preceding analysis has been necessarily brief and has examined only a small sample of Judd's work. It has passed over, for example his . . . [three books are named] and a number of other books. His work with the American Council on Education and other boards and commissions has also been omitted. But the sample taken is fairly representative of the total, as can be seen by the way . . . certain persistent themes . . . [appear]. A larger survey would have produced more statements of these ideas but relatively few new items.' P. 549 of 'Judd's debt to Wundt', by Harold B. Dunkel. *School Review*, **85** (August 1977), 535–551. We observe here a qualified expression 'fairly representative' that may be difficult to make precise.

Finally, we return to the book by Hanna Fenichel Pitkin, who expresses the coverage concept clearly, although perhaps too restrictedly: why only one object of each variety? A good art museum might have more than one French impressionist painting, Renaissance Madonna, etc. 'The difference between being representative and being typical shows up . . . in a collection containing several major varieties of objects. Suppose that we are to pick not a single representative object but a representative group of objects from such a collection. For most purposes we would pick one object of each of the major varieties in the collection, to make up our representative group. If we were interested also in the relative proportions . . . we would have to select different numbers of objects from each variety instead of just one.' P. 80 of *The Concept of Representation*. Berkeley and Los Angeles: University of California Press, 1967.

The ideas of coverage emphasized in this section seem closely associated with producing representatives and not so much with representativeness in the look alike sense we would have for mirrors of the population. Although probability samples might produce some of the variation desired in these ideas of coverage, sometimes surer and more economical coverage could be gotten by direct judgment selection. We must keep in mind that judgment selection can go badly wrong if variables not occurring in the original partition make an important difference.

We would not use 'representative sample' for the ideas in this section because they deserve to be spelled out and justified separately in each instance to avoid confusion.

6^1 *Linguistic cousins of representativeness; other variations*

In our discussion of the representative sample idea in non-scientific usage, we found a few negative opinions about representativeness. No such have turned up in our scientific reading.

[1] Meaning 6 (the vague term to be made precise) was discussed in Section 0. Two principles for numbering sections are in conflict, and we chose to put the new meaning first.

We have in our collection, however, a number of cousins of representative sampling, and we pick a small subgroup for presentation here.

A highly interesting 1942 book by Allport suggests the notion of degree of representativeness, 'To counteract this selectivity of sample it has been proposed that the cooperation of a more representative cross section be solicited'. P. 125 of *The Use of Personal Documents in Psychological Science*, Gordon W. Allport, Social Science Research Council, New York, 1942. The cross-section usage continues, for example, '. . . the editors of such a book can do no more than bring together a representative cross section of available research results. . . .' From Robert D. Martin's review of *Lemur Biology*, Ian Tattersall and Robert W. Sussman, Editors. New York: Plenum, 1975. Review in *Science*, **191** (1976), 1166 only.

Here is a qualified version: '. . . the sample included a fairly representative cross section of the world's geographical areas, racial stocks, and cultural groupings'. P. 1009 of 'Infantile stimulation and adult stature of human males', by Thomas K. Landauer and John W.M. Whiting, *Amer. Anthropologist*, **66** (1964), 1007–1028.

It may not be only a matter of cross sections, but of fractions, '. . . 50 per cent of the energy release was assumed to be due to fission, and that is a representative fraction'. P. 31 of 'Limited nuclear war', by Sidney D. Drell and Frank von Hippel, *Scientific American*, **235** (November 1976), 27–37; or of the ambiguous notion of a fair sample (fair sized? equitable? suitable for good inferences?), '. . . the selected precincts gave a fair sample of the political conditions in half of the fifty wards of Chicago'. P. 183 of 'Personality Studies', by Harold D. Laswell. Pp. 177–193 in *Chicago/An Experiment in Social Science Research*, edited by T.V. Smith and Leonard D. White. Chicago: University of Chicago Press, 1929.

A similar begging of the question is '. . . a representative quantity of loose . . . soil material was collected with a clean brush. . . .' P. 489 of 'Plutonium hazard in respirable dust on the surface of soil', by Carl J. Johnson, Ronald R. Tidball, and Ronald C. Severson. *Science*, **193** (1976), 488–490.

A variety of other rhetorical cousins may be found. For example, 'The survey . . . involved a statistically valid sample of a metropolitan population. . . .' From 'Energy conservation and creditability', a letter to *Science* from Phyllis T. Thompson and John MacTavish, **192** (1976), 1286. 'This appraisal would involve . . . review of a statistically valid sample of the records. . . .' From p. 41 of 'Rx: a peer review system for physicians', by Jacob Fine. *Bull. Atomic Scientists* (Sept. 1977), 38–43.

Still another nuance is shown in the following quotation, where heterogeneity in a large population is said to make a simple ratio unrepresentative: 'In 1970 there were 34,107 doctors in Mexico . . . The ratio of inhabitants to doctors, which is 1423.7, is not a representative index of the actual distribution of resources because there is a great scarcity of health professionals in rural areas and a high concentration in urban areas. . . .' From p. 1131 of Luis Cañedo, 'Rural health care in Mexico', *Science*, **185** (27 Sept. 1974), 1131–1137.

One fascinating atypical use equates non-representativeness with sampling variability. 'Samples may vary because of . . . real . . . differences between their parent populations, or they may vary, although drawn from populations that really do not differ, because . . . they are not representative of those populations. The last mentioned cause of variation between samples is called "sampling error". . . .' P. 162 of 'A Univac analysis of sherd frequencies from the Carter Ranch Pueblo, eastern Arizona', by J.A. Brown and L.G. Freeman, Jr., *American Antiquity*, **30** (1964), 162–167.

A special use of the term 'representative sample' was introduced by the late Egon Brunswik, a psychologist, but the use apparently did not catch on. Brunswik's discussion refers to experiments, primarily psychological, in which one has independent variables in the usual sense and a dependent variable. Ordinarily experimenters adopt values, when possible, for the so-called independent variables in a systematic way quite unrelated to how those independent

variables are distributed in real life. It is more illuminating, according to Brunswik, to take values of the independent variables as they naturally occur, and this is called representative sampling for the values of the independent variables.

Two objections to Brunswik's espousal of natural values for the independent variables are as follows: First, the idea is ill-defined, since natural values may vary wildly in different settings. Second, the use of natural values may lead to inefficient estimation, perhaps because of near-collinearity. Brunswik's approach and writings are to be applauded, however, for reminding us of the distinction between the population sampled and the population of real interest to which we would like to generalize.

Four relevant publications by Brunswik appeared in 1944, 1949 and 1956.

7 *Conclusion and recommendations*

In the scientific, but non-statistical, literature we find six main meanings for the terms 'representative sample' and 'representative sampling', and we make recommendations about scientific usage for each meaning.

Meaning 1

General, unjustified acclaim for the data. Since this meaning gives an author's data a pat on the back without empirical or theoretical support, we recommend that such misleading puffery be abandoned.

Meaning 2

Absence (or presence) of selective forces. The extensive literature on sample surveys encourages us to believe that absence of selective forces is achieved by very careful design. If such a careful design has been executed, then this usage will revert to Meaning 6.

'Non-representative sampling' is used to indicate presence of selective forces. In such a case, we recommend that specific illustrations of forces leading to bias should be cited, preferably with some estimate of the extent of the effect. This negative usage can probably be interpreted as the negation of Meaning 6 – not a probability sample – and so we do not oppose its use.

Meaning 3

Mirror or miniature of the population. When this rare kind of sampling occurs, or nearly does, we recommend that these vivid and precise terms be used rather than the weaker, vague term 'representative sample'.

Meaning 4

Typical or ideal case(s). Although such cases or specimens have a right to be called 'representative(s)', to reduce confusion, we recommend that the words 'sample(s)' and 'sampling' be avoided for this meaning.

Meaning 5

Coverage of the population. This usage, especially the notion of one or more cases drawn from each stratum of the population, differs enough from others that we recommend for it the expression 'coverage' rather than 'representative sample'. If a specific fraction (less than 100 per cent) is covered, and if it can be determined or estimated, that fraction could well be reported.

If the notion in Meaning 5 is to suggest that the sample has heterogeneity comparable to that of the population, we recommend that the sense be precisely explained and that 'representative sample' be avoided in favor of some more specific label appropriate to the study.

Meaning 6

Vague term to be made precise. Although such usage can be over-worked, it may be convenient in exposition to have a vague term like 'representative sample' to describe the result of the sampling, for example, in the summary of the study. We also recommend that the term be reserved for occasions when some form of probability sampling is employed. We recommend that such usage carry the obligation to make the term precise either in the same paper or by a suitable reference.

'Representative sampling' has the advantage of implying an important purpose of the sampling in a way that 'probability sampling' does not.

Our third paper, 'Representative Sampling, III', will review the usage in the statistical literature.

Acknowledgments

We express our appreciation to colleagues who have called our attention to additional examples and raised questions about some interpretations. Among these are Elizabeth Allred, Ariane Auquier, Donald T. Campbell, Churchill Eisenhart, Donald Fiske, Rebecca Fuhrer, Miriam Gasko-Green, John P. Gilbert, D.V. Glass, Eric Hamp, David Hoaglin, Joseph B. Kruskal, Nan Laird, Robert K. Merton, Kenneth Moy, Anita Parunak, John W. Pratt, Nancy Romanowicz, David Schoenfeld, Susan Shott, Stephen Stigler, Michael Stoto, Judith Strenio, Lawrence Thibodeau, and George Wong.

The preparation of this paper was facilitated by National Science Foundation Grants Nos. NSF SOC72-05228 A04 to the University of Chicago and NSF SOC75-15702 A01 to Harvard University.

References

Brunswik, E. (1944). Distal focusing of perception: size-constancy in a representative sample of situations. *Psychological Monographs*, **56**, No. 1, Whole No. 254.

Brunswik, E. (1949). Systematic and representative design of psychological experiments. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* (1945–1946), edited by Jerzy Neyman. University of California Press, Berkeley and Los Angeles.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. University of California Press, Berkeley and Los Angeles.

Brunswik, E. (1956). Historical and thematic relations of psychology to other sciences. *Scientific Monthly*, **83**, 151–161.

Campbell, D.T. and Stanley, J.C. (1963). *Experimental and Quasi-experimental Designs for Research*. Rand McNally, Chicago.

Kruskal, W. (1958). Ordinal measures of association. *Journal of the American Statistical Association*, **53**, 814–861.

Kruskal, W. and Mosteller, F. (1979). Representative sampling, I. *International Statistical Review*, **47**, 13–24.

Résumé

Cet article d'ecrit et explique à l'aide d'exemples les sens des termes 'échantillon représentatif' et 'échantillonnage représentatif' tels qu'ils sont utilisés dans la littérature scientifique extrastatistique. Six catégories semblent adéquates pour décrire ces emplois. Les cinq premières ont également paru dans la revue des emplois non-scientifiques dans 'Echantillonnage Représentatif I', Kruskal et Mosteller (1979). Nous les définissons et les expliquons à l'aide d'exemples dans les parties 1 à 5 du présent article.

Sens 1. Un enthousiasme général et non justifié pour les données. Le chercheur 'gonfle' les données en utilisant un terme apparemment scientifique pour rehausser leur stature. Aucune raison valable, soit dans le plan de l'étude soit dans l'information empirique, n'est apportée pour justifier cet emploi, et c'est pourquoi nous suggérons qu'il soit rayé du langage scientifique.

Sens 2. Absence (ou présence) de forces discriminatoires. Si des raisons valables pour leur absence sont données, ce sens peut revenir au sens numéro six, un terme vague qu'il reste à rendre scientifique. Sinon, il revient au sens numéro un. En soulignant la présence de forces discriminatoires, on dément le sens numéro six.

Sens 3. Reflet ou miniature de la population. L'échantillon a la même répartition que la population. Cette caractéristique rare s'accompagne de contraintes sévères. Nous éviterions d'appeler ceci un échantillon représentatif car il semble être beaucoup plus.

Sens 4. Cas typique ou idéal. Un élément de la population qui la représente en moyenne ou de façon la plus probable (ou idéale). Nous appelons ceci un représentant et non un 'échantillon représentatif'.

Sens 5. Recouvrement de la population. Echantillons conçus pour refléter la variation, particulièrement entre les strates. Un échantillon contenant au moins un élément de chaque strate d'une partition de la population est dit la recouvrir pour cette partition. Nous n'emploierions pas le terme 'échantillon représentatif' pour de tels échantillons.

Le nouveau sens trouvé dans la littérature scientifique est celui d'un terme vague. Nous le décrivons dans la partie 0.

Sens 6. Une terme vague qu'il reste à préciser. La description d'une méthode d'échantillonnage demande de longues explications et il peut être pratique d'avoir à sa disposition un terme tel qu' 'échantillon représentatif' ou 'échantillonnage représentatif' pour se rapporter à l'échantillon ou au procédé qui l'engendre. Nous suggérons que cet usage soit réservé aux échantillons probabilistes et que la méthode d'échantillonnage soit décrite dans la même étude.