discussion

Is There a Fallacy of Small Sample?

THOMAS LEDDY   *San Jose State University*

Critical Thinking texts generally include discussions of two fallacies in the section on statistical induction: the fallacy of small sample and the fallacy of unrepresentativeness. The text that I am using now, Cederblom and Paulsen (CP), does not classify these as fallacies but does include them as important ways that inductive arguments can go wrong. As an example of small sample it gives Woody Allen's inference from two unsuccessful dates that women will generally reject him.[1] Toulmin, Rieke and Janik (TRJ) classify small sample and unrepresentativeness as two types of the general fallacy of hasty generalization. They describe small sample as drawing a general conclusion from too few specific instances; for example, one might conclude that "All Audis are lemons" from a few reports of friends. Another example given by TRJ is when someone argues that Poles are unintelligent "on the grounds that the thirty odd Poles they have worked with over the years have all appeared to them to be on the dull side."[2] TRJ then contrasts this with their version of the fallacy of unrepresentativeness, called the fallacy of "atypical examples." This fallacy occurs when "we take as our evidence examples that are unrepresentative of the given phenomenon."[3] This distinction between small sample and unrepresentativeness is not unique to CP and TRJ: it is found in several critical thinking texts. Yet the distinction is pointless. There is no special fallacy of small sample: all of the work can be done by the fallacy of unrepresentative-

ness. In short, we correctly criticize a sample for being too small only when its smallness tends to make it unrepresentative. If Woody Allen could be sure that the two women he dated were typical of the population available for dating, then he would have reason to project a disposition to incompatibility on the population as a whole. The sample of two is too small because it is likely to be unrepresentative.

What of the individual who concludes that all Audis are lemons from a small sample? Isn't this a case of committing the fallacy of small sample and not just of unrepresentativeness? The example is tricky since the conclusion is a universal generalization and a universal generalization can always be successfully defeated by one counterexample. The problem here is that any sample which is less than 100% of the population will yield results that are inconclusive. Of course we can give the speaker the benefit of the doubt and assume that he or she is only claiming probable truth for the assertion. But if that is the case then the size of the sample is not the main issue. The issue is how reasonable it is to project to the population as a whole a characteristic of the sample. And if the sample is fully representative of the population then the projection is justified regardless of the size of the sample. But let's assume that the speaker said that most Audis are lemons on the basis of his or her sample. The value of this inference would also depend entirely on the representativeness of the sample. Indeed TRJ's second example—

"Poles are unintelligent"—fits this pattern. The claim could either mean that "All Poles are unintelligent" or "Most Poles are unintelligent." Assuming that the speaker means that most Poles are unintelligent the proper approach is to see whether his or her sample of 30-odd Poles is representative. The size of the sample is only one factor which might indicate that the sample is unrepresentative. Other factors might include the situation in which he or she met the Polish Americans (whether in bars or at mathematics conferences), the geographical location of these meetings, etc. In short, although the size of a sample is neither a necessary nor a sufficient condition for determining whether a sample is unrepresentative, the very smallness of a sample may give us good reason to believe that it is probably unrepresentative. This depends, of course, on the context.

Thus I am not denying that size is a factor in representativeness. For instance, in order to get a good stratified sample of the American populace political pollsters use about 3,000 individuals and not 300. They need to have the larger number in order to cover the large number of variables. They need to have not only blacks, Spanish-Americans, etc., but also rich, middle class and poor. Also a rich black man will not be sufficient. There must be a rich black woman, a middle class black man, etc. It may even be important for there to be more than one rich black man in order to insure that the sample of rich black men is representative of the class of rich black men, although it is always possible that one rich black man would be sufficiently representative for the purposes of the poll.

It is interesting in this regard to glance at a textbook in another field. The writers of **Statistical Reasoning in Sociology** discusses sample size in a chapter on "Parameter Estimation." Their answer to the question "What size sample?" is "It depends on the resources available, how the sample is to be drawn and analyzed, the anti-

cipated loss of cases for final analysis, the characteristics of the population being sampled, and the precision required in the results."[4] They also state that "as the variability in the sampled universe increases, sample size must increase to maintain the same level of precision in parameter estimation."[5] So clearly the sample can be too small, but the fact that is it too small is a judgment determined by the variability of the population and the precision needed by the researchers. No fallacy of small sample is mentioned in this text, and that is so because no such fallacy is needed.

In order to determine whether we can eliminate the fallacy of small sample we must ask whether an argument can commit it and not commit the fallacy of unrepresentativeness or, can an argument be considered fallacious because the sample is too small, regardless of our knowledge of its representativeness? Some might argue that a sample of one is clearly always too small to make a statistical generalization. However upon tasting a teaspoon of milk from a gallon carton in my refrigerator and discovering that it was sour, I poured out the entire contents of the carton. Did I commit a fallacy? By no means. I had reason to believe that the teaspoon of milk was representative of the whole carton. Similarly a scientist might reasonably generalize from a sample of one. Archimedes did not vainly cry "Eureka" when he discovered that a crown of gold displaces less water than an equal weight in silver. He seemed to believe that his sample was not only representative of all crowns of gold but of all weighted objects. His experiments probably needed to be repeated in order to convince doubters, but one sample was enough to give him what he considered to be justified belief, given his background knowledge.

A critic might reply that even though there are some cases in which a sample of one is sufficient there are others in which we know that the sample is too small regardless of our knowledge of its probable representative-

ness. For instance my opponent might argue that a farmer who wanted to plant a prune orchard would inevitably be foolish if he based his decision on the success of only one prune tree. It would be argued that we know independently of representativeness that the quality of survival cannot be projected from a sample of one in the case of prune trees. But if all prune trees generally behaved the same then one sample would be sufficient to show that the new orchard would probably be successful. The facts are, however, that prune trees vary, and in order to capture this variability a larger sample is needed. My critic might reply that the problem with a sample of one is that even if prune trees are generally homogeneous with respect to survival in this location a sample of one could still include one of the exceptions, i.e. a super-prune-tree which will survive in climes where the general run of prune trees fail. Yet this only seems to prove that in our imperfect world samples of one are not sufficiently representative when the class is only strongly, and not strictly homogeneous. Once again the issue of sample size (too small or not) is a function of one's purposes, and that is because representativeness, the real issue, is a function of one's purposes. The sample of one sip of the sour milk was sufficiently representative for my purposes because the unpleasantness of taking a second taste was not warranted by the slim possibility that the first taste might not be representative.

If the arguments against believing in a fallacy of small sample independent of representativeness are so strong, why do people insist on believing in it? This can partly be explained in terms of Humean psychology. Constant conjunction may be necessary to give the impression that something does indeed have the property in question. Archimedes may have jumped in and out of his bath several times because he couldn't believe his eyes, thereby increasing the size of the sample even though the first instance was, logically speaking, sufficient. Second,

smaller samples tend to be less representative than larger samples, though this is not always the case. Third, there is a certain ambiguity in the concept of "too small." Something is "too small" relative to various purposes and needs. The term "small" has a somewhat different logic. Although purposes and needs sometimes enter into our determination of whether something is to be called "small," other factors are also important. For instance, fleas are commonly thought to be small, even though they are large in relation to atoms: we think of fleas as small because they are small in relation to humans. Sometimes we say that something is small and imply that it is too small relative to our purposes. This may lead to the view that the relative smallness of a sample is sufficient to determine that it is too small for our purposes. Yet one can have a sample of one which is clearly small but not too small. People might be inclined to believe that "small" implies "too small" because smallness tends to be valued less highly than largeness in our society, although there are nonetheless some things which are thought good because they are small (note the aesthetic terms "dainty" and "petite"). Small people tend to have less authority in our society than large people and small incomes tend to be less desirable than large incomes. Most people would prefer to have a larger house, army, or backyard. Finally there is a persistent feeling that induction is a matter of piling up examples: the more examples the better the induction. This feeling may persist even in those who have rejected the belief that it represents.

It has been established then that although samples can be small or even too small there is no inductive fallacy of too small sample. The criterion for adequacy of a sample is its representativeness. When a particular population is heterogeneous, small samples may generally fail to be representative, but when we are speaking of homogeneous populations a sample of one could be sufficient. Smallness is a pos-

sible factor in unrepresentativeness but the latter is the fallacy.[6] It may be argued that I am making a mere terminological suggestion: i.e. to use the label "sample too small to be representative" instead of "sample too small" simpliciter.[7] The suggestion I am making may be minor but it is hardly terminological. If we are going to make any sense of the term "fallacy," it must refer to a distinguishable entity: otherwise we would never know whether or not somebody is properly attributing a fallacy term to a particular sentence. We should be able to know whether we have two fallacies or one. The point of this paper is that where we thought there were two fallacies there is in fact only one. It is possible to replace all uses of "sample too small" with "sample too small to be representative" and indeed this terminological move is implied by my ontological claim. What turns out to be impossible is for a sample to be both too small and representative, or (surprisingly) the right size and unrepresentative. What could "right size" mean if not "sufficient size to be representative"? Could it mean "sufficient size to be representative if there were no other reasons why it should not be representative"? I think not, since the determination of whether the sample is too small or whether its size is right is based entirely on the presence or absence of other factors which determine representativeness.

Notes

[1] Jerry Cederblom and David W. Paulsen. *Critical Reasoning: Understanding and Criticizing Arguments and Theories*. Wadsworth Publishing, Belmont CA (1982), pp. 119-120.

[2] Stephen Toulmin, Richard Rieke and Allan Janik. *An Introduction to Reasoning*. MacMillan Publishing Co. New York (1979) pp. 158-161.

[3] *Ibid.*

[4] John H. Mueller, Karl F. Schuessler, Herbert L. Costner. *Statistical Reasoning in Sociology*. Houghton Mifflin Company, Boston (1977) p. 407.

[5] *Ibid.*

[6] Morris Cohen and Ernest Nagel say something similar to this in their *An Introduction to Logic and Scientific Method*. Harcourt Brace and Co., New York (1934) p. 281: "While we can never be altogether certain that an examined verifying instance is a fair sample of all possible instances, in some cases the probability that this is true is very high. This is the case when the subject matter of the inquiry is homogeneous in certain relevant ways. But in such cases it is unnecessary to repeat a large number of times the experiment which confirms the generalization. For if a verifying instance is representative of all possible instances, one such instance is as good as another. Two instances which do not differ in their representative nature simply count as one instance."

[7] Question thanks to an unnamed reader.

Professor Thomas Leddy, Department of Philosophy, San Jose State University, One Washington Square, San Jose, CA 95191 □