WRITING ARCHAEOLOGY

ANALYSES AND ARCHAEOLOGICAL ARGUMENTATION

Keith W. Kintigh

Keith W. Kintigh is a professor of anthropology at Arizona State University.

This document presents some ideas about writing that I've developed over several years of reading term papers, theses, dissertations, and professional articles. It highlights some common problems and makes some positive suggestions.

The Sequence of Analyses Mode of Argument

In presenting the results of a quantitative analysis, authors (and especially student authors) sometimes describe, analysis by analysis, the sequence of steps that they went through. In caricature, the model is:

> Here's my problem...My data set is...and here are my variables...First I did...analysis and it showed...Then I did...and that showed...Then...Then...In conclusion, the reader can plainly see...

The underlying idea is that, at the end, the conclusions will rise out of the mist and will seem warranted by the analyses. This approach may be favored (perhaps unconsciously) because it shows the reader how much the writer suffered to achieve these fabulous insights. It also has the appearance of being even-handed, letting the data speak through the analysis rather than having the author's ideas imposed on the data. This dramatic pulling together of all the narrative threads is an excellent approach for a mystery novel, but not for a scholarly article.

I hasten to add that this mode of presentation may well describe a perfectly appropriate intellectual process that ultimately led to a reasonable conclusion. Doing analysis is an inherently messy business. Most analyses (and certainly all of mine) involve lots of trial-and-error, sidetracks, and dead ends. I've never kept track of the numbers, but it would not surprise me if I discarded 100 analyses for every one that appears in an article.

In an article, however, one must display the wisdom of hindsight in presenting the *really interesting results*. This compact form is dictated partly by limitations of publication space but more importantly by the limits of the reader's patience and time. If the reader has to wait until page 23 to get to the *wonderfully interesting result*, the reader will likely never get there. If you have submitted the article for publication and the reader is a reviewer, you can imagine the consequences.

Consider an analogy. As a buyer of an airplane, I'm interested in a machine that actually flies and its performance characteristics. I'm not interested in the design process that led to the airplane or in seeing all the designs that failed to fly. In an article, you are selling a product (an argument for a conclusion), not a design process. As a consumer, it is not particularly important to me whether you went through a lot of design, redesign, and unsuccessful testing, or whether divine inspiration led you directly to a great result. It's the final performance that counts.

I doubt that the sequence-of-analyses mode is ever an effective style of presentation. The problem is similar to the problem with site reports that describe the data room-by-room, hearth-by-hearth, pot-by-pot, bone-by-bone, or whatever. In the first place, my eyes glaze over almost immediately (I suspect that I am not alone in this). Second, even with a serious effort, the reader usually can not process the information for lack of adequate guidance on what to focus on. In a quantitative presentation, no one ever wants to read the results of eight cluster analyses at multiple clustering stages. (I don't even want to read this if you're talking about *my* data.) This style of presentation is difficult to follow because the reader never knows where the argument is going, and thus can't focus on the critical aspects of the analysis. It puts an *enormous* burden on the reader to fully comprehend the data and absorb the results of each analysis, in order to remain poised for the denouement. Most readers will not accept this burden, and they should not be expected to.

Suggested Mode of Argument

Ordinarily in an article (this applies to theses and dissertations as well), the objective should be to argue convincingly one or a relatively small number of points. At the outset, you may not actually know what those points are. However, by the time you are writing the last draft, you *must* have the key points clear in your head, and you must make them clear to your reader. Your objective in writing should be to guide the reader to understand your points and toward being convinced by the argument and data that support them.

The point of which you want to convince the reader may well *not* be your original hypothesis. That is, you might have started a mortuary analysis intending to show that such-and-such a group could be fairly characterized as a complex chiefdom. However, your analysis may lead you to believe something different or to believe that the evidence is ambiguous. A well-presented case that the evidence is ambiguous on an important issue can be an important contribution. In this case, the *conclusion* you are trying to convince people of is that there is a balance of evidence on both sides of an argument or that the data are so ambiguous that they do not permit us to rule out key alternative explanations.

To make a convincing argument is *not* to say that you only present evidence that supports your point. Your responsibility is to *fairly* present the evidence *relevant* to your conclusions. While you never present every possible analysis that bears on your question, your responsibility is to present the critical analyses and let the chips fall where they may. If you want to convince the reader of a particular conclusion despite apparently strong evidence to the contrary, you should present your grounds for dismissing that evidence.

Your abstract or introductory paragraph should first sketch your problem and an outline of your conclusions. The idea is to capture the reader's attention and make them want to read the rest. Ordinarily, you would then state and justify the problem and provide any necessary background. Next, you might present what you are going to conclude in more detail and outline how you are going to get there. The idea, of course, is to assist the reader in following your argument in every way possible.

If you are up front about it, there is nothing wrong with presenting your argument in a quasi-deductive style, even though your thought processes went the other way. The argument will often be of the form: "Here is my conclusion and here is the evidence that supports it." I am not suggesting that you mislead people about what you are claiming. Indeed, I'm happy for you to admit that these conclusions were developed inductively. You might, in a paragraph, sketch the kinds of analysis that you tried in getting to your conclusions. However, if you do this, make sure that when you are done, it really contributes in some way to the paper. If not, lose it. From there on, you should be presenting an argument that directly relates analyses to your conclusion. It just works better this way.

In the course of writing, you will usually find that to really make the point convincingly, you will need to do additional analyses to plug logical holes in your argument or to provide additional confirmation or disconfirmation. This happens to me all the time.

With the suggested mode of presentation, it is almost always the case that you will present a small fraction of the analyses that you executed. Each analysis that you present should be relevant to your argument. If some result looks interesting, but you do not know what to make of it, the reader probably will not know either. By sticking to a few points in an argument with a strongly developed logical structure, the reader is more likely to really follow where you are going. That said, you might include some of the interesting but less-relevant analyses in an accessible electronic archive that could be used and referenced by someone who wished to explore the issue in more detail.

Once you have made your case, you need to attend to obvious objections. For example, let's assume that you want to argue for clear-cut status differences in a cemetery that has considerable time depth. Showing large differences in the richness of graves that cross-cut age and gender would not be sufficient. An obvious objection would be that the division of grave richness is not due to social classes but to changing burial practices over time. Even though you may have looked at that hypothesis and had good grounds for rejecting it, a brief argument with the relevant data would almost certainly be worth presenting. Absent that argument, readers will find your analysis not wrong, but unconvincing. The problem is knowing how many possible objections to address. A lot of that depends upon the limits of the forum in which the argument is being presented. Dissertations can address a substantial range of possible objections. Brief articles can probably only address a few quite obvious ones.

Papers presenting quantitative results that I consider successful tend to follow this general formula. Remember, *the idea is to bring quantitative methods to bear on a substantive problem, not the other way around*. I'm not trying to cram everyone into a writing straightjacket here. There is a great deal of latitude within the general framework I am suggesting. But at each stage, it should be clear why you have done the particular analysis, what you were looking for, and what it in fact shows. While it is appropriate to present analyses that both support and disconfirm your hypotheses, it is not appropriate to present analyses that, however interesting, do not contribute to the specific arguments of your paper.

The Professional Audience

If you are writing to a general professional audience, it is probably reasonable to assume that the reader understands percentages, correlation, simple probability, how to read a contingency table, and perhaps a bit about principal components (or factor) analysis and cluster analysis. When you use less-common but still fairly standard techniques like discriminant analysis, multidimensional scaling, or correspondence analysis, you should remind the reader in one or a few sentences what the technique is doing for you. If you can't state it succinctly, you probably don't understand it well enough to be using it. If you use more obscure methods, such as Local Density Analysis or Koetje's analysis of concentration, you must explain them sufficiently that the reader can follow what you are arguing. If it is important enough to use the method, it's important enough to explain

it fully. My review article on spatial analysis (Kintigh 1990) does this for a number of methods.

No matter what analysis you use, make clear what variant you used (e.g., stepwise or simultaneous multiple regression). Make clear what variables you used and why you selected them and what cases you used and why you selected them. Also make clear any transformations of the data that you did. Note that "standardize" has a marked and an unmarked sense; if you mean percentages, say so and say what they are percentages of. If you mean z-scores, say you did a z-score standardization or something equivalent.

Precision and Completeness

Writing with precision is essential. This is not any more true for quantitative arguments than for nonquantitative ones. Don't use "demonstrated" to discuss arguments that are better characterized by "we have a slight hint that..." Nontrivial claims of "proof" are rarely if ever justified in archaeology.

When you present tables and figures, it is your job to help the reader *as much as possible*. Make sure that table rows and columns and figure axes are adequately labeled. It is never acceptable to have cases or variables numbered from 1 to N in tables and figures only accompanied by a key. Doing this forces the reader to take many extra steps just to follow your argument.

If you are discussing data where percentages are the relevant concepts, do not provide a table of counts. Instead, give a table of percentages with a total N on which each row or column of percents is based. If you are trying to show differences, think hard about what graphic would best make the case. Most people are not good at seeing patterns in tables of numbers. Stacked box or dot plots, histograms, Ford diagrams, and scatter plots are simple devices that communicate well (although box plots still require a brief explanation for some audiences).

In general, and especially in theses and dissertations, if you are presenting significant data, if it is at all possible, provide an appendix with the raw data so others can re-analyze them or use them for other purposes—and by all means, make the data available electronically. You have nothing to hide, right?

Acknowledgments. This article began as a class handout several years ago. George Cowgill's suggestions, some of which I have paraphrased, have greatly improved this article. Over the years, I have incorporated useful suggestions from other colleagues at Arizona State University; I regret that I have lost track of the specifics and cannot credit them individually.

References Cited

Kintigh, Keith W.
1990 Intrasite Spatial Analysis: A Commentary on Major Methods. In *Mathematics and Information Science in Archaeology: A Flexible Framework*, edited by Albertus Voorrips. *Studies in Modern Archaeology* 3:165–200. HOLOS-Verlag, Bonn.

THINGS TO REMEMBER ABOUT STATISTICS (WHATEVER ELSE YOU FORGET)

George L. Cowgill

George L. Cowgill is a professor of anthropology at Arizona State University.

1. Statistical analysis is not a way to arrive at certainty; it is a powerful aid in discerning *what* your data suggest, and *how strongly* they suggest it. This is often done better by an *estimation* approach than by hypothesis testing.

2. *Look* at your data first, through simple tables and pictures. Often this tells you everything important. If not, it will tell you what is sensible or not sensible to do next. Do not rush to apply advanced techniques while overlooking the messages of simple methods.

3. If you must do a hypothesis test, report the actual probability level obtained and don't treat some arbitrary level such as five percent as a talisman that tells you what to think.

4. It's not the sampling fraction that matters; it's the *size* of the sample. For example, a well-chosen sample size of 100 that is one percent of a large population can tell you a lot, but a sample of 10 that is 20 percent of a small population tells you less.

5. One can wring useful suggestions from very small samples. However, at least for artifact collections, it is usually desirable to get 300–400 cases if possible; smaller collections are often too small for satisfyingly accurate estimates of interesting proportions or other properties. Do not trust the maxim that 100 or so cases will generally be enough.

6. Proportions, percents, and ratios represent something *relative to* something else. Proportions are fractions, with a numerator *and* a *denominator*. When you write, always *report* the denominator. When you read, always ask yourself whether you understand what denominator is implied. Often you will find that the denominator is unclear or inappropriate.

7. "Frequency" should always mean the *count* of something, rather than the *ratio* of something to something else. It is often used to mean both.

8. If you are worried about data quality, reducing data to "present/absent" only makes the problem worse unless you are sure that absence in the sample unambiguously implies absence in the relevant population. But a category that is scarce but present in the population will be totally absent in many random samples from that population, and the chance that it is absent in any one sample is strongly dependent on the size of that sample. Together with sampling vagaries, this makes "presence/absence" a very unstable statistic. If you want to be conservative, use something like "way below average," "about average," and "way above average."