forming a particular class to understand the role of that class. Even if one has missed some classes, which needn't be the case, the role can be understood.

Conclusion
It is not surprising that one should find comments in a volume dedicated to justifying the efficacy of full-coverage surveys that overgeneralize the contrast between sample and full-coverage surveys and overemphasize the value of full-coverage data. Neither the sample nor the full-coverage surveys that have been done exhaust the list of potential surveys that could be done, and the various mixes of strategies that could be used. The papers make a convincing case for full-coverage survey on two issues:

1. When one intends a functional and paleoethnographic analysis of the network of relations that existed among the inhabitants of sites in an area, full-coverage data are essential.

2. If we are to fully understand how sample data are best interpreted, high quality full-coverage data are essential to the effort.

A simpler way of stating these two points requires no more than reflection on the very high density of maps in this volume. Sample data produce relatively poorer maps. For those problems requiring such maps, full-coverage data are essential.

The Kowalewski and Fish concluding discussion points the way to a beneficial dialogue on the relative merits of full-coverage and sample survey. Such discussion should lead to improved survey design, combining the merits of both strategies, and to the degree that circumstances warrant. In any case, these papers provide a powerful basis for comparing the relative merits of full-coverage and sample surveys and for exploring the issue of how their results are to be compared.

References Cited

Huntman, Jeffrey

The contributors to this volume vary somewhat in their assessments of the relative merits of full coverage and spatial sampling in regional archaeological surveys. They generally agree, however, that sampling is sometimes necessary. Above all, if circumstances beyond one’s control make it impossible to survey an entire region at a uniform and reasonably high intensity, then a survey of a well-designed probability sample of spatial units will be much better than an opportunistic or haphazard partial survey. Because one is sometimes forced to resort to sampling in surveys, it is important to know about designs for spatial sampling and their strengths and weaknesses. Among numerous publications on this topic, many of the chapters in Mueller (1975) are still relevant, and Plog, Plog, and Wait (1978) is a classic. Nance (1983) provides a useful and extensive review of
more recent developments. Among still later publications, Read (1986) is especially noteworthy.

The main theme of this book, however, is that spatial sampling is generally a very inferior alternative to full coverage at fairly high intensity. To put it another way, it is most unwise to leave unexamined or only lightly examined any appreciable fraction of the region one is trying to understand. Sampling makes good sense when many elements of the sampling universe are highly similar to one another, so that there is much redundancy. But in regional archaeological survey we find it very difficult and risky to try to infer the properties of unexamined tracts; and sampling, rather than being a sensible and cost-effective way to economize, must be seen as an unwelcome expedient sometimes forced on one by limited time or resources.

As the contributors point out, spatial sampling can yield fairly good estimates of some gross properties of the sites in a region, such as ranges and averages of sizes and numbers per unit area in each of the subregional "strata" predetermined by the sampling plan. However, even these estimates can be problematic if, for example, the fraction of the total region covered is too small to yield reasonably large numbers of some important categories, or if the predefined strata turn out to correspond poorly to certain environmental features or other phenomena later recognized to be important.

Furthermore, spatial sampling, unless the sampling fraction approaches unity, is apt to be virtually useless or even misleading, and thus perhaps worse than useless, for questions about the spatial organization of sites and for the discovery of rare and inconspicuous but important occurrences. All these points are repeatedly made and frequently demonstrated by examples in the chapters of this volume. I thoroughly agree, because the arguments and examples are convincing and because experience with the Teotihuacan Mapping Project leads me to similar conclusions (Cowgill, Altschul, and Sload 1984:158–159). Although that project involved very intensive survey of a single site rather than a region, the total area surveyed (over 30 sq km) and the diversity of the remains encountered made many of the considerations of regional survey relevant.

Indeed, the shortcomings of anything less than full coverage are so well argued and illustrated by the other chapters in this volume that there seems to be little need to reiterate these points. It is more important to direct attention toward refinements in the general concept of "full coverage." Many of the issues I will raise are mentioned by several of the contributors to this volume, but they do not elaborate on them very much.

First, of course, is the matter of defining a "region," touched on especially in the chapter by Parsons. The general idea is that our survey area should be large enough to include all the sites occupied by reasonably well-defined social units during the periods of interest. Obviously we cannot realistically expand coverage to include all elements of long-distance trade networks and the like, but we should try to include all the area occupied by people likely to have been in frequent close interaction with one another. Leaving out parts of such an area is sometimes a political or economic necessity, but it amounts to a particular kind of spatial sampling, and opportunistic rather than probability sampling at that.

Given adequate definition of a region, the basic question we should ask ourselves is "With what intensity, and with what sensitivity, should the region be surveyed in order to get good answers to important questions?"

There are two ways in which survey results can be unsatisfactory. The first is that they can be too ambiguous. That is, even though procedures may be unbiased (i.e., all types of occurrences of data relevant to one's questions, in all types of situations, have equal chances of being recorded by the survey), the sample sizes are too small to permit sufficiently accurate estimates of important population parameters. Obviously, full coverage at relatively high intensity will provide larger numbers of recorded occurrences, and hence will improve the accuracy of estimates. I should add that the relevant population is often not simply the finite number of occurrences that actually exist in the

250 Chapter 12

Cowgill 251
region, but the infinite population of potential occurrences implied by whatever theory we (or others) may wish to use survey results to test. Hence, we can rarely get away from the need to make estimates and the need to be concerned with the confidence intervals of these estimates.

The second way in which survey results can be unsatisfactory is that the procedures themselves are biased [i.e., some types of occurrences, or occurrences in certain situations, are considerably less likely than others to be recorded]. Some kinds of bias are hard to deal with, such as burial of occurrences by alluviation, destruction by severe erosion, and masking by dense vegetation. Probably the only ways to deal with alluviation and dense vegetation are through test excavation or at least clearing of selected tracts, as discussed in this volume, especially by P. Fish and Gresham. It seems that these procedures can be reasonably satisfactory, but they unavoidably require sampling, and remind us that, however much we aspire to "full" coverage, there is no escape from the need to make good use of probabilistic sampling theory in some aspects of the overall design of regional surveys. I know of no good ways to deal with the problem of erosion, except to take note of the likely severity of the effects of erosion in each small part of the survey area. The chapters in this volume have very little to say about the effects of erosion.

A further source of bias that is perhaps less obvious is that even "full" coverage can be considerably biased against discovery of visible small occurrences unless the intensity of the survey is quite high.

Three terms should be defined before I go further. I use the word "occurrence" to refer to any tangible spatially localized phenomenon of potential archaeological significance, from a single utilized flake to a great city with massive pyramids. P. Fish and Gresham in this volume use occurrence to refer to finds of 10 or fewer artifacts, and use the word "site" for localized finds of more than 10 artifacts. I will use the word occurrence for the whole range of sizes, partly for convenience and partly to avoid the need, irrelevant for present purposes, to adopt a definition of the term site.

The terms "intensity" and "sensitivity" also need definition. By intensity I mean simply the spacing approximately maintained between individuals during survey. This is very close to the sense in which the term is used by others [e.g., Plog, Plog, and Wait 1978:389-390], although sometimes its meaning is expanded to include some aspects of what I define as sensitivity. By sensitivity I mean the probability of detecting an occurrence of evidence of ancient activity. Sensitivity depends on five things:

1. The nature of the occurrence. Other things being equal, large sites with big mounds are more likely to be detected than small flat scatters of a few artifacts.
2. The nature of the terrain. Heavy vegetation, modern construction, alluviation, and sharply uneven topography impede detection of occurrences. A moderate amount of erosion can aid detection, but extreme erosion can remove most or all of an occurrence.
3. The closest approach of a surveyor to the occurrence. An occurrence is more likely to be detected if an observer's path crosses it or at least runs very close to it than if no observer ever gets close to it.
4. The extent to which observers are sensitized [using the term in its psychological sense] to a certain type of occurrence. In this volume Fish, Fish, and Madsen, for example, discuss a case in which important rock piles were apparently simply not recognized as archaeologically significant. Sumner, using the term "selectivity," refers to the possibility that field workers may be attuned to occurrences dating to particular periods, or certain types of materials, and may ignore, or at least underrecord, other potential information.
5. The extent to which special techniques are used to detect subsurface phenomena, such as shovel tests, trenching, ground penetrating radar, magnetometer surveys, aerial photos using various spectral bands, and other simple or advanced technology.

The connection between sensitivity and intensity, as I have defined the terms, is that, other things being equal, sensitivity to evidence that is visible at all to an observer walking over the sur-
face is increased by increasing intensity, because it increases the probability that at least one observer’s path will come close to or actually traverse the occurrence.

Nevertheless, sensitivity to small, inconspicuous surface occurrences cannot approach 100 percent unless the intensity of survey is very high (that is, the spacing between observers is very close). For example, spacing of 30 m between observers should result in detecting essentially all surface occurrences with a diameter not much less than 30 m. If, however, one has to get within 6 to 8 m of the edge of a sparse artifact scatter without visible architecture in order to have a good chance of spotting it (which I think is often the case), then if the scatters are only a few meters in diameter, observers spaced 30 m apart will have rather low sensitivity to them.

In fact, reflection on this situation leads to the conclusion that, with respect to small inconspicuous artifact scatters, “full coverage” really isn’t full coverage at all. It is a form of systematic sampling with rather closely spaced and narrow transects.

A little armchair model should make these ideas clearer. My reasoning in what follows is very much in the spirit of the discussion by Plog, Plog, and Wait (1978), although the specific assumptions differ slightly. Suppose that, for fairly sparse artifact scatters with no mounds or other visible architecture, an observer’s path has to run at least 1 m inside the perimeter of the occurrence in order for the detection probability to be 100 percent. (This implies that an observer will sometimes walk right over the outer edge of a sparse scatter without noticing it. I think this is realistic.) Suppose that the probability of detection is 95 percent if an observer gets within a meter of the outer perimeter, 75 percent if an observer gets within 3 m, 50 percent if an observer gets within 5 m, and zero if no observer gets within 5 m of the outer perimeter of the occurrence. All these values are sheer guesses. I think they are plausible for light sherd and artifact scatters in places where vegetation and archaeologically meaningless natural rock scatters and other obscuring phenomena are fairly sparse but not absent. However, I have not tested them and I emphasize that they are merely guesses intended for use in an illustrative model (Table 11).

Table 11. Summary of Illustrative Hypothetical Model

<table>
<thead>
<tr>
<th>Closest approach of observer’s path to periphery of archaeological occurrence</th>
<th>Probability that occurrence is detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 m or more inside periphery</td>
<td>100%</td>
</tr>
<tr>
<td>1 m inside to 1 m outside</td>
<td>95%</td>
</tr>
<tr>
<td>1-3 m outside</td>
<td>75%</td>
</tr>
<tr>
<td>3-5 m outside</td>
<td>50%</td>
</tr>
<tr>
<td>more than 5 m outside</td>
<td>0</td>
</tr>
</tbody>
</table>

Accepting these guesses about sensitivity and assuming that approximately circular occurrences of all sizes are distributed randomly throughout the survey area (and ignoring edge effects at the margins of the area), some interesting consequences follow. If observers follow paths spaced 30 m apart then the probability of detection of all occurrences 32 m or more in diameter will of course be unity because the path of at least one observer will intersect at least the outer meter of every occurrence this large. Furthermore, the sensitivity is nearly 100 percent for occurrences 28 m in diameter because it can be shown that 99.8 percent of such occurrences will be detected. In order to have a 95 percent or better chance of detection, an occurrence must have a diameter of about 22 m or more. The probability of detection of smaller occurrences decreases rapidly. For a diameter of 20 m, it is a little under 90 percent, for a diameter of 10 m it is 56 percent, and for a single isolated object (for which perhaps this simple model is much too optimistic) it is about 27 percent. In order to detect 95 percent of the occurrences 10 m in diameter, the intensity of survey would have to be increased so that the

---

1 This and the other numbers I provide can be obtained by drawing pictures and using a little high school algebra and some of the probability theory given in the early chapters of elementary statistics texts. I leave the derivation of these numbers as a challenge to the reader.
The spacing between observers is only about 16 m. To detect 95 percent of isolated objects on the surface, spacing between observers would have to be reduced to 2 m (which leads to the tongue-in-cheek suggestion that observers space themselves by holding hands). Needless to say, according to this model, at a spacing of, say, 50 m, many occurrences even 20 or 30 m in diameter will be missed.

Perhaps readers will think that the assumptions embodied in Table 11 are unduly pessimistic for good survey conditions. Whether this is so or not, all too often survey conditions are poor enough to make Table 11 overly optimistic. In any case, I cannot repeat too often that Table 11, and all the specific numbers derived from it, represent some plausible guesses intended to illustrate the point that a survey intensity which is “100 percent” for large, easily detected sites will miss a great many small and inconspicuous surface occurrences. The specific numbers should not be taken seriously. To generate numbers that might be taken seriously, field tests of the results of resurveying test areas with different intensities would be needed. It is, incidentally, notable that, except for Whalen, most contributors to this volume have little to say about any such tests.

A circular occurrence 32 m in diameter has an area of about 800 sq m (0.08 ha) and one with a diameter of 10 m has an area of about 80 sq m (0.008 ha). Does it really matter if a large proportion of occurrences only 10 m in diameter (or some other minimal size) are missed? This is a crucial question, as I mentioned earlier, and it requires some clear thought about research objectives and the kinds of occurrences that one thinks can safely be ignored. From some of the chapters in this volume and from other publications (e.g., Plog, Plog, and Wait 1978; Thomas 1975) it is clear that, at least in much of the United States, occurrences this small or smaller are taken very seriously. On the other hand, I do not think that surveys in highland Mexico (see the chapters here by Parsons and by Kowalewski) have ever claimed that they were taking account of such small occurrences, even when they happen to be detected.

At early stages of research in a region where there are some sites covering hundreds of hectares, and thousands of sites covering at least a tenth of a hectare, it is tempting to say that smaller scatters of artifacts, without mounding or other special features, can be treated as nondata. However, even in regions occupied by highly urban complex societies, such small occurrences are not without importance. One major reason is that there are usually periods of great interest, especially the earliest periods, in which practically all occurrences were scarce and small. These should not be overlooked simply because the large sites of other periods are so much more evident. A second reason is that, even in the periods with numerous large sites, there may have been significant small occurrences, such as isolated farmsteads or small special activity areas. To get a full picture of even the most complex society, surveys should provide data on these tiny occurrences. To say this is not to detract from the enormous accomplishments of regional surveys at moderate intensity, but we must be aware of what may possibly have been missed, and realize that there remains a great potential for surveys at much higher intensity.

The most obvious way to get relatively good information on even very tiny occurrences (insofar as they are visible at all on the surface) is to increase the intensity of survey by spacing survey team members so closely that nearly all tiny occurrences will be spotted. Less obviously, but perhaps more importantly, there should be a redefinition of what constitutes data. It is likely that many surveys at moderate intensity have not recorded tiny occurrences even when workers were aware of them, because they were considered “nonsites” and too insignificant to be worth recording.

It is not clear to me how much closer spacing and paying attention to tiny occurrences would increase the person-days needed to survey a given area. Obviously if the spacing is reduced from 30 m to 10 m about three times the person-miles must be walked. But I do not know whether closer spacing, in itself, would increase the total time required to cover the area by a factor of three. What seems more important is that if tiny occurrences are recorded, even with relatively wide spacing, workers would stop to record many more occurrences. This means that they must cover the ground more slowly, but it may not add
greatly to the time required because it should take much less time to record everything if you are to be said about a tiny occurrence and to make whatever collection is feasible. Narrowing the spacing will have almost no effect on the total numbers of larger occurrences encountered, which are the most time-consuming to study and record, and will only increase the number of small and quickly studied occurrences. Probably tripling the intensity of survey and defining even tiny occurrences as data will considerably less than triple the person-days required to survey a given area. Without experimental data, it seems hard to say more than this.

The other possibility, if one wants to get information about tiny as well as large occurrences, is to keep the spacing between field workers relatively wide (say 30 to 50 m) and use tiny occurrences observed along the survey tracts to estimate the local density of tiny occurrences in each small segment of the strips between the survey tracts. Such estimates should be reasonably unbiased and should be quite useful for common types of tiny occurrences because their exact locations within small cells of the region will not matter much. However, there is the complication that the effective sampling fraction will be different for each size of small occurrence and will also depend on the type of occurrence and on local topography and surface conditions. It will no longer do to rely on armchair models such as the one I presented above. I see no good alternative to getting empirical data from surveys at very high intensity in a carefully stratified sample of the survey region (probability sampling again!). Even so, estimates will be rather complicated to calculate and subject to considerable uncertainty. Moreover, for small and rare kinds of occurrences, the problem of small absolute numbers will remain.

All in all, it seems that the best strategy is to think carefully in advance about the smallest and most inconspicuous kinds of occurrences that one wants to treat as data, and then survey with intensity sufficient to spot nearly 100 percent of such occurrences. If it is really impossible to survey the entire region at this intensity, then the next best (but much less desirable) alternative would be uniform coverage of the entire region at the highest feasible intensity, and coverage of a stratified sample of tracts at high enough intensity to ensure nearly 100 percent sensitivity to tiny surface occurrences. In either case, it is essential to report the intensity (or intensities) of the survey and also the criteria used to decide what kinds of occurrences should be recorded.

In thinking about the smallest occurrences worth recording, one must remember that the fewer the objects, the less likely they are to be “diagnostic” of a particular period, type of activity, or ethnic or status category. However, this seems to vary considerably, depending on the type of material and what is already known of the region. Perhaps a dozen eroded sherds are uninformative, but sometimes a single object can be well worth recording, and it is difficult to generalize about this matter.

References Cited


